



Publicly Accessible Penn Dissertations


1-1-2014

DNA Conformational Changes and Phase Transitions Induced by Tension and Twist

David E. Argudo

University of Pennsylvania, argude7@gmail.com

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Biophysics Commons](#), and the [Mechanical Engineering Commons](#)

Recommended Citation

Argudo, David E., "DNA Conformational Changes and Phase Transitions Induced by Tension and Twist" (2014). *Publicly Accessible Penn Dissertations*. 1191.

<http://repository.upenn.edu/edissertations/1191>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/1191>

For more information, please contact libraryrepository@pobox.upenn.edu.

DNA Conformational Changes and Phase Transitions Induced by Tension and Twist

Abstract

DNA is a double stranded helical molecule with an intrinsic right handed twist. Its structure can be changed by applying forces and torques in single molecule experiments. In these experiments DNA has been seen to form super-helical structures (supercoils), collapse into tightly condensed states (toroids) and undergo structural changes (phase transitions). Our work focuses on studying all these phenomena by accounting for DNA elasticity, entropic effects due to thermal fluctuations and electrostatics.

First, we study the DNA compaction problem in super-helices and toroidal structures. To do so we combine a fluctuating elastic rod model of DNA with electrostatic models for DNA-DNA interactions. Our models are able to predict the onset of the transition to supercoils and toroids under a wide range of experimental conditions. Next, we address DNA phase changes in the presence of mechanical loads. A phenomenon well known from experiments is the overstretching transition associated with the sudden change of DNA extension at high tensions. Depending on the ionic concentration, temperature and pulling rate, DNA can either transform into a melted state (inner strand separation) or S-DNA.

Motivated by this, we study the equilibrium and kinetics of the DNA overstretching transitions making use of a quartic potential and non-gaussian integrals to evaluate the free energy of the system. We find that the cooperativity of the transition is a key variable that characterizes the overstretched state.

In a separate study we make use of a heterogeneous fluctuating rod model to examine the hypothesis that a newly discovered left-handed form called L-DNA is a mixture of two relatively well-characterized DNA phases - S-DNA and Z-DNA. L-DNA is stable at high tensions and negative twist. We show that if the idea of a mixed state is correct, then the content of S-DNA and Z-DNA varies as a function of the ionic concentration. Finally, we also use our fluctuating rod model to study the mechanical properties of drug-DNA complexes. We show that our methods can predict the results of experiments from various labs if we use only one set of experiments to fit the data to our model.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Mechanical Engineering & Applied Mechanics

First Advisor

Prashant K. Purohit

Keywords

DNA Kinetics, DNA mechanics, DNA Phase Transitions, Statistical Mechanics, Supercoiling, Thermal Fluctuations

Subject Categories

Biophysics | Mechanical Engineering

DNA CONFORMATIONAL CHANGES AND PHASE TRANSITIONS INDUCED
BY TENSION AND TWIST.

David Argudo

A DISSERTATION

in

Mechanical Engineering and Applied Mechanics

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of
the Requirements for the Degree of Doctor of Philosophy

2014

Prashant K. Purohit,
Associate Professor of Mechanical Engineering and Applied Mechanics
Supervisor of Dissertation

Prashant K. Purohit,
Associate Professor of Mechanical Engineering and Applied Mechanics
Graduate Group Chairperson

Dissertation Committee:
Professor John L. Bassani,
Mechanical Engineering and Applied Mechanics
Professor Pedro Ponte Castañeda,
Mechanical Engineering and Applied Mechanics

Acknowledgments

I would like to thank my advisor, Dr. Prashant K. Purohit. I am grateful for the guidance, encouragement and the time he spent training me as a scientist. His enthusiasm, passion and dedication to science have made these five years a very rewarding experience.

I would like to express my gratitude to the members of my Dissertation committee, Dr. John L. Bassani and Dr. Pedro Ponte Castañeda, for their advise and instruction. Their experience and knowledge have been an invaluable resource for my studies.

My grateful thanks also to the members of the Mechanics of Materials Group: Reza, Morteza, Tianxiang Su and Sheng Mao, with whom I shared this wonderful experience.

To my parents and siblings for being the foundation of who I am and inspiring me to become a better person, student and researcher every day of my life. I also want to extend my gratitude to my grandparents for being the pillars of wisdom in my family.

Last but not least, I thank you Anca. Your love and support have been an essential ingredient of this journey.

ABSTRACT

DNA CONFORMATIONAL CHANGES AND PHASE TRANSITIONS INDUCED BY TENSION AND TWIST.

David Argudo

Prashant K. Purohit

DNA is a double stranded helical molecule with an intrinsic right handed twist. Its structure can be changed by applying forces and torques in single molecule experiments. In these experiments DNA has been seen to form super-helical structures (supercoils), collapse into tightly condensed states (toroids) and undergo structural changes (phase transitions). Our work focuses on studying all these phenomena by accounting for DNA elasticity, entropic effects due to thermal fluctuations and electrostatics.

First, we study the DNA compaction problem in super-helices and toroidal structures. To do so we combine a fluctuating elastic rod model of DNA with electrostatic models for DNA-DNA interactions. Our models are able to predict the onset of the transition to supercoils and toroids under a wide range of experimental conditions. Next, we address DNA phase changes in the presence of mechanical loads. A phenomenon well known from experiments is the overstretching transition associated with the sudden change of DNA extension at high tensions. Depending on the ionic concentration, temperature and pulling rate, DNA can either transform into a melted state (inner strand separation) or S-DNA. Motivated by this, we study the equilibrium and kinetics of the DNA overstretching transitions making use of a quartic potential and non-gaussian integrals to evaluate the free energy of the system. We find that the cooperativity of the transition is a key variable that characterizes the overstretched state.

In a separate study we make use of a heterogeneous fluctuating rod model to examine the hypothesis that a newly discovered left-handed form called L-DNA is a mixture of two relatively

well-characterized DNA phases - S-DNA and Z-DNA. L-DNA is stable at high tensions and negative twist. We show that if the idea of a mixed state is correct, then the content of S-DNA and Z-DNA varies as a function of the ionic concentration. Finally, we also use our fluctuating rod model to study the mechanical properties of drug-DNA complexes. We show that our methods can predict the results of experiments from various labs if we use only one set of experiments to fit the data to our model.

Contents

1	Introduction	1
1.1	Background	1
1.2	Overview of this Thesis	7
2	The dependence of DNA supercoiling on solution electrostatics	10
2.1	General Description of The Model	11
2.1.1	Potential Energy of the System	13
2.1.2	Internal Energy: Entropy and Electrostatics	15
2.1.3	Electrostatics	15
2.1.4	Configurational Entropy	16
2.1.5	Variational Formulation	16
2.2	Comparison with experiments and predictions	17
2.2.1	The transition point and jump estimates	19
2.2.2	Coexistence of loops and plectonemes	22
2.2.3	Multivalent Ions	23
2.3	Conclusions	26
3	Competition between Supercoils and Toroids in Single Molecule DNA condensation	29

3.1	DNA supercoiling	30
3.1.1	Straight Configuration	31
3.1.2	Straight-Supercoiled Configuration	31
3.1.3	Interaction energy in high multi-valent salt	32
3.1.4	Solution for the superhelix parameters	33
3.1.5	DNA transitions and formation of supercoils	34
3.1.6	Transition from straight to straight-supercoil co-existence	34
3.1.7	Transition to the fully supercoiled state: pure plectoneme	35
3.2	DNA Globular State	35
3.2.1	DNA Condensation: coil-globule transition	37
3.3	Results: Comparison with experiment	38
3.3.1	Phase Diagram: DNA configurations	40
3.4	Conclusions	41
4	DNA superhelical Structures with non-constant helical pitch	43
4.1	General Description of Model 1: extended ply structures	44
4.2	The Variational Problem	46
4.3	Potential Energy of the System	47
4.3.1	Variational Formulation: Fixed F , R and n	48
4.3.2	Boundary Conditions for the Variational Problem	49
4.3.3	Solution	49
4.4	General Description of Model 2 : Plectonemes	53
4.4.1	The Variational Problem	54
4.4.2	Potential Energy of the System	55
4.4.3	Internal Energy: Entropy and Electrostatics	57
4.4.4	Variational Formulation: Fixed F and n	57

4.4.5	Boundary Conditions for the Variational Problem	58
4.4.6	Constant pitch solutions	59
4.4.7	Variable pitch solutions: External moment $M_{ext} = M_3$ depends on n	61
4.4.8	Variable Pitch – Analytical Solutions: <i>Three</i> cases	62
4.5	Conclusions	65
5	A heterogeneous fluctuating rod under imposed tension and twist	67
5.1	Model Description	69
5.2	Energy of the system	69
5.2.1	Discretized version of the energy	72
5.2.2	Statistical mechanics of the chain	73
5.2.3	General expressions for the heterogeneous chain	74
5.2.4	Results for the homogeneous chain	75
5.2.5	A two state heterogeneous chain in the n-ensemble.	78
5.3	Applications: Comparison with experiment and predictions	79
5.3.1	Small molecule EtBr binding to B-DNA	79
5.3.2	Analysis of L-DNA mechanics	84
5.4	Conclusions	88
6	Equilibrium and kinetics of DNA overstretching modeled using a quartic energy landscape	90
6.1	Model Description	91
6.1.1	Energy of the system	92
6.1.2	Homogeneous Equilibrium Solutions	92
6.1.3	DNA overstretching transitions: changes in temperature and salt concentration	92
6.1.4	Experimental observables: DNA extension z	93
6.2	Statistical Mechanics of the Chain	93

6.2.1	Partition function Z of a segment of length l	94
6.2.2	Force-extension relation for the chain undergoing the overstretching transition	95
6.3	Applications: comparison with experimental data	96
6.3.1	DNA overstretching transitions: B-to-S, B-to-M or both	96
6.3.2	Cooperativity length l predictions	97
6.4	Kinetics of the chain: Sharp interface	100
6.5	Conclusions	105
7	Concluding Remarks and Future Directions	107
A	Appendix - Chapter 2	109
A.1	The end loop	109
A.2	Testing the Internal Energy models	110
A.3	Critical torque $M_{critical}$ and the jump $\delta M = M_{critical} - M_3$	114
A.4	Indirect method for calculating external moment	115
A.5	Multivalent ions	116
A.6	Experimental and Simulation Data	120
B	Appendix - Chapter 3	121
B.1	Solution for the supercoiling parameters	121
B.2	Alternative energy expressions for the toroidal condensates	121
B.3	Competition of Globules vs Plectonemes	123
B.3.1	Energy source contributions	124
C	Appendix - Chapter 4	128
C.1	Link, writhe and twist	128
C.2	The Constrained Rod and Variable Curvature	129
C.3	Convexity of κ^2 as a function of θ	131

C.4	Variable end loop length	131
D	Appendix - Chapter 5	133
D.1	Force-extension and torque-link relations in a F, M_3 ensemble.	133
D.2	Homogeneous chain	133
D.2.1	Special heterogeneous chain: 2-phase model	134
D.3	Stretch, twist-stretch couple and the (F, n) -ensemble	136
D.3.1	Energetics of the chain with stretch and twist-stretch coupling (Cosserat theory of rods):	136
D.3.2	General force-extension and link-torque relations for a two phase heterogeneous chain.	139
D.3.3	Results for Specific cases	141
D.4	Link Lk and supercoiling parameter σ relations	142
D.4.1	Linking number and phase transitions	143
D.4.2	B-DNA to L-DNA transition where L-DNA is an effective single state	144
D.4.3	B-DNA to L-DNA transition where L-DNA = (Z+S)-DNA	145
D.5	Free energy expressions of a Discrete Rod: bending, twisting and stretching	146
D.5.1	E_ω : Twist Hamiltonian	147
D.5.2	E_ρ : Stretch Hamiltonian	149
D.5.3	E_ν : Bending Hamiltonian - Partially clamped	150
D.5.4	Free energy of a homogeneous chain subjected to tension and twist	152
D.5.5	E_ν : Bending Hamiltonian - (Hinged-Hinged)	153
D.6	More Results	155
E	Appendix - Chapter 6	158
E.1	More Results	158

E.2	Empirical expressions for the critical force as a function of temperature and ionic concentration.	161
E.3	Chain Kinetics	161
E.3.1	Appendix: Approximation of the equilibrium value of u for small C	165
E.4	Asymmetric Potentials	165
E.4.1	Appendix: Hamiltonian including the cubic term	166
E.4.2	Piece-wise Potential: Two quartics	166
E.5	Kinetics of the chain with finite size interfaces: Domain Walls	168
E.5.1	Transition Rates Γ and Mean lifetime Θ of the overstretching transition	169
E.5.2	Calculation of the non-uniform solution in the infinite length case	169
E.5.3	Thermodynamics equations of motion: infinite length rod	171
E.5.4	Appendix: Statistical Mechanics of Domain Walls: evaluating K	173
E.5.5	Nondimensionalization of the stochastic Ginzburgh-Landau	173
E.5.6	Appendix: Chain segments with Periodic B.Cs	174
E.5.7	Appendix: Chain segments with Neumann B.Cs	176
E.6	Short dsDNA constructs	178

References **181**

List of Tables

- 2.1 Effective linear charge ν used in our calculations as a function of the monovalent salt concentration c_o [mM]. The third column shows the fraction $\xi = \nu/\nu_{bare}$, where ν_{bare} has been computed as in Refs. [1–4] for $a = 1\text{nm}$. For salt concentrations in the range [30-500]mM, the value of the charge ν_{bare} can be approximated by a linear fit with $R^2 > 0.99$ (linear fit predicts 99% of the variance on the fitted variable). Based on this idea we performed a linear fit to the value of ν we have used and obtained $\nu^{fit} = 2.46 + 2.38 \times 10^{-2}c_o$ with $R^2 > 0.99$ and c_o in mM units. A linear fit to ξ gives $\xi^{fit} = 0.73 - 2.7 \times 10^{-4}c_o$. The fourth column shows the fraction $\hat{\xi} = \nu/\nu_{bare}$, where ν_{bare} has been computed using $a = 1.2\text{nm}$ as in [5]. Note that for large salt concentrations ($c_o \sim 0.32 - 0.5[M]$) $\hat{\xi} \approx 0.42$ is equal to the charge adaptation factor used in Maffeo et al. simulations [5]. 19
- 2.2 Effective linear charge ν for $c_o = 0.2\text{M}$ KCl buffer and added multivalent salt concentration c_{mu} . For only the monovalent salt $c_o = 0.2\text{M}$ KCl we used $\nu^o = 4.12$ [nm⁻¹]. As the multivalent salt concentration is increased the value of ν obtained from the fit decreases. This can be explained by a better screening of the DNA charge by the salt solution and the varying electrophoretic charge value in multivalent ion solutions. The Sp³⁺ values of ν are well fitted by the curve $\nu_{3+}^{fit} = \nu^o(1+c_o/1.07)^{-1/3}$ ($R^2 > 0.97$), and the Sp⁴⁺ values of ν are well fitted by the curve $\nu_{4+}^{fit} = \nu^o(1+c_o/0.07)^{-1/3}$ ($R^2 > 0.99$). The curve-fits were obtained using the least squares method with a fitting function of the form $f(x) = a(b+x)^c$ 24

3.1	K_b, C_A and D_{eq} effects. A variable's increase and decrease are denoted by \uparrow and \downarrow respectively. The line -- means no significant effect.	34
5.1	Properties of B-DNA and EtBr bound DNA. [6] used a WLC model to fit their experimental data of a torsionally unconstrained molecule and obtained a reduced stretch modulus $S_2 \sim 220\text{pN}$	82
5.2	L-DNA properties as a single state or as a mixture of two states: S-DNA + Z-DNA.	87
6.1	Summary of recent DNA overstretching experiments used throughout this study. L is the length in the B-DNA state and \bar{L} is the length of the DNA filament after the transition measured in L units. The abbreviations n-h and h stand for non-hysteretic and hysteretic respectively.	98
6.2	Calculation of the Zimm-Bragg cooperativity σ_F as a function of l using relation Eq. (6.26). As a reference value, Rouzina and co-workers [7, 8] measured $\sigma \approx 10^{-3}$ in DNA overstretching experiments at room temperature and $I = 150\text{mM}$	100
A.1	Internal Energy Models	111

List of Figures

1.1	A ribbon model for illustrating twist and writhe. Figure 1.1 (a) and (b) are reproduced from Figs 24-11 and 24-18 in [9]. The ribbon in (a) represents the axis of a relaxed DNA molecule. Changes in linking number usually lead to the formation of supercoils, also known as plectonemes, which are DNA extruded interwound structures. In (b) it is shown how DNA twisting generates a torsional strain that can either produce changes in the writhe or a change in twist of the molecule.	6
2.1	Sketch representing single molecule experiments, where a DNA molecule is fixed at one end, while the other end is subjected to a pulling force F and twisted by a given number of turns n	11
2.2	Slope $d\Delta z/dn$ as a function of the external force F . Lines represent our predictions and markers the experimental data. The upper graph shows the experimental data in Brutzer et al. [10] and Maffeo et al. [5]. The values of the linear effective charge ν used are 2.83, 3.80, 6.16 and 10.00 [nm^{-1}] for 30mM, 60mM, 170mM and 320mM salt concentration respectively. The lower graph shows the experimental data in Mosconi et al. [11] and the values of the linear effective charge ν used are 3.73, 5.32, 7.71 and 14.31 [nm^{-1}] for 50mM, 100mM, 200mM and 500mM salt concentration respectively.	17
2.3	External moment M_3 and superhelical radius r theoretical predictions for the different salt concentrations in Brutzer et al. [10]. The lines are our predictions and the markers are the data points corresponding to the values of r and M_3 in the Monte Carlo simulations of Maffeo et al. [5]. The values of ν used are shown in Table 2.1.	18

2.4 Critical number of turns $n_{critical}$ as function of the external force. Lines represent our predictions and markers the experimental data. Experimental results taken from Brutzer et al. [10] at $c_o = 320\text{mM}$ and Forth et al. [12] at $c_o = 150\text{mM}$. For $c_o = 150\text{mM}$ we have used $\nu = 5.93\text{nm}^{-1}$. Including the contribution of the end loop to compute $n_{critical}$ gives consistent results with experiment. The thin continuous lines show the solution for $n_{critical} \pm \Delta n$ (including the end loop), where Δn is an estimate of the fluctuations in n during the transition. 20

2.5 Comparison of the extension jump δz from our theory with the experiments in Brutzer et al. [10] for 1.9kbp and 10.9kbp DNA templates at $c_o = 320\text{mM}$ and the data in Forth et al. [12] for 2.2kbp and 4.2kbp DNA templates at $c_o = 150\text{mM}$. Our model predicts that as the DNA length L or c_o increase δz increases too. 21

2.6 Energy per unit turn. Using our model we can get some idea of the preferred state: if $V_{o-turn} > V_{p-turn}$ plectonemes are favored, if $V_{p-turn} > V_{o-turn}$ loops are favored and if $V_{o-turn} \approx V_{p-turn}$ there is a coexistence of loops and plectonemes. We expect that for high-to-moderate salt concentrations, plectonemes will be formed, while for low-to-medium salt concentration there might be a region of coexistence or even formation of only loops. 23

2.7 Dunlap and co-workers performed two series of experiments corresponding to the control set up ($c_o = 0.2\text{M}$ KCl), one for spermidine Sp^{3+} (circles) and the other for spermine Sp^{4+} (triangles). We fit the value of ν to get the slope of the rotation-extension curve at $F = 0.6\text{pN}$ for control set corresponding to Sp^{3+} (circles) and got $\nu = 4.12\text{nm}^{-1}$. Our prediction with $\nu = 4.12\text{nm}^{-1}$ is shown by the solid line. The data from Mosconi et al. [11] shown by the cross markers corresponds to the $c_o = 0.2\text{M}$ NaCl series. 24

2.8	Predictions for the slope of the rotation-extension curves in the presence of multivalent ions. Crosses are data points from Dunlap and co-workers for a mixture of $c_o = 0.2\text{M}$ KCl and different c_{mu} concentrations of spermidine Sp^{3+} . For $c_{mu} = 5\text{mM}$ and $c_{mu} = 10\text{mM}$ as F increases the supercoiling diameter decreases and approaches the interaxial spacing distance ~ 3.0 nm for spermidine [13, 14]. The dot in the bottom two panels shows the point where $2r$ reaches the limiting interaxial distance and from there on the dashed line shows the solution where $r = 1.5\text{nm}$ is assumed to be constant.	25
2.9	Predictions of the slope of the rotation-extension curves in the presence of multivalent ions. Crosses are data points from Dunlap and co-workers for a mixture of $c_o = 0.2\text{M}$ KCl and different c_{mu} concentrations of spermine Sp^{4+} . For $c_{mu} \geq 0.5\text{mM}$ as F increases the supercoiling diameter approaches the interaxial spacing distance ~ 2.9 nm of spermine [13, 14]. As in Fig. 2.8, the dot shows the point where $2r$ reaches the interaxial distance value and the dashed line shows the solution where $r = 1.45\text{nm}$ is assumed to be constant.	26
3.1	Multivalent ions with 3 or more positive charges can cause condensation. In a given experiment only one type of ion is typically used. The DNA molecule is fixed at one end, while the other end is subjected to a pulling force F and twisted by n number of turns. Top: sketch representing the formation of supercoiled structure in single molecule experiments in the presence of condensing agents. Bottom: sketch representing the formation of a toroid.	30
3.2	The curve of $n_{critical}$ shows the transition from straight configuration to supercoiled + straight coexistence. The curve n_{min} shows the minimum number of turns at which supercoiled + straight can coexist. The curve n_{sg}^{twist} (n_{sg} for toroids with no twist) shows the transition from straight to toroidal configurations. Spermine is the condensing agent, and we have used $C_A = 690\text{pN}/\text{nm}^2$, $K_b = 50k_B T$ and $\alpha = 0.013$ ($\beta \approx 2.6$ and $\Gamma d^2/(k_B T) \approx 0.6$).	38

3.3 Same predictions as in Figure 3.2, but using cosep as the condensing agent. We have used $C_A = 1150\text{pN/nm}^2$, $K_b = 25k_B T$ and $\alpha = 0.013$ ($\beta \approx 7.5$ and $\Gamma d^2/(k_B T) \approx 1.1$). The black crosses, green squares and red circles correspond experimental data in Besteman et al. [15] for different molecules under same experimental conditions. The inset show the comparison to the experimental data using protamine as the condensing agent [16]. For protamine we have used $\beta = 1$ and the same values of C_A , D_{eq} and K_b as the ones used for cosep, since both condensing agent have a high charge density. Using K_b as low as $15k_B T$ still gives very good agreement with experiment for both protamine and cosep. 39

3.4 Phase diagrams constructed using the theoretical solution for the description of Besteman et al. experiments in spermine. In (a) we used $\alpha = 0.013$ as in Fig. 3.2 and in (b) $\alpha = 0.0003$. The straight configuration is the most favorable one in the white region and it is marked with S . The straight-supercoiled configuration is marked with $S + P$ and colored yellow, while the entirely supercoiled structure (pure plectonemes) is marked with PP and colored green. T stands for the region where we expect to see toroidal structures. The red color shows the region where we expect to see toroids (with or without twist) and the purple region where we expect to see toroids only without twist. The light blue region labeled U in (a) corresponds to an unknown region where there is no clear favorable configuration. 41

4.1	Sketch representing single molecule experiments where supercoils are formed. Two DNA molecules are fixed at one end, while the other end is connected to a magnetic or optical bead. The beads apply a pulling force F and twist the DNA strands by a given number of turns n . Moreover the beads can be moved in space to apply a given number of rotations R about the helical axis e_3 . If the DNA strands are nicked then the beads can not fix n and the molecules cannot carry twist. This leads to the formation of braids as in the case presented by Charvin et al. [17]. Similarly, in the dual DNA manipulation set up of Noom et al. [18] the beads are free to rotate and can not impose n . So there is no twist in the molecules. But, in general, an experimental set-up using optical or magnetic traps including twist can be realized [11, 12].	44
4.2	Sketch representing single molecule experiments where plectonemes are formed. A DNA molecule is fixed at one end, while the other end is subjected to a vertical pulling force F and twisted by a given number of turns n	45
4.3	Profile of $\theta(s)$ as a function of s for nicked DNA (no twist) in a setup as shown in figure 4.1. In the calculations we have used $L = 1200\text{nm}$ $K_b = 180\text{pNm}$, $f = 2\text{pN}$, $2e = 0.36L$ and the hardcore radius $2r = 6\text{nm}$, as done in a set of MC simulations in [17]. $\theta(s)$ approaches the mid-section angle γ but diverges fast close to the end points s_1 and s_o . The plateau value of θ is a function of the prescribed R . We also computed \hat{R} using the mid-section approximation where we assume that the helical region has a constant $\theta = \gamma$ in the interval $-l_b < s < l_b$, while $\theta = \alpha$ only at the end points s_1 and s_o . The percentage error between R and \hat{R} decreases for longer braids. For $R = 4.1$, the percentage error is approximately 3%.	50

4.4	Normalized end to end extension z/L as a function of catenation number $C_a \approx 3.4\hat{R}/L$ for nicked DNA (no twist). In the calculations we have used the same parameters as in figure 4.3. The lines show the solution up to the point where constraint Eq. (4.10) holds for $\alpha \leq \pi/4$. The opening angle α in the helical region has to be smaller than or equal to the lock-up angle [19]. The end of each line depicts the critical catenation number C_{a_c} at which the DNA is believed to make a transition into more complex structures (supercoiling of braids) [17]. The X marker shows $C_{a_c} \approx 0.045$ from MC simulations using $2r = 4.2\text{nm}$. The inset shows z/L as a function of Ca using $2e = 0.02L$ and $r = 6\text{nm}$ and the markers correspond to MC data [17].	52
4.5	External moment M as a function of catenation number Ca for nicked DNA (no twist). In these calculations we have used the same parameters as in figure 4.4. The inset shows results using $2e = 0.02L$ and $r = 6\text{nm}$ and the markers correspond to MC data [17].	52
4.6	Helical angle as a function of catenation number Ca for nicked DNA (no twist). In these calculations we have used the same parameters as in figure 4.4. Here $\theta \approx \gamma$, the mid-section angle of the braid, while α is the opening helical angle at the end points. The inset shows results using $2e = 0.02L$ and $r = 6\text{nm}$	53
4.7	(a) External moment M as a function of catenation number Ca for twisted DNA. In these calculations we have used the same parameters as in figure 4.4. (b) Twisting moment M_3 as a function of catenation number Ca for twisted DNA. For a given n adding external rotation R to the system is equivalent to allocating part of the link into writhe. For positive R the induced writhe is negative as noted from Eq. (4.6) and Eq. (4.7). Hence, for fixed n , the twist increases to counteract the effect on the writhe effect, which can be seen as an increase in the value of M_3 as a function of Ca	53

4.8 S stands for a completely straight configuration. $S + P$ stands for the configuration depicted in figure 4.2, where straight tails coexist with a helical region. PP represents the state when the entire DNA molecule has been converted into a superhelix. The scenario depicted represent an monovalent ionic solution of 0.1M NaCl. Black circles correspond to the 0.1M experimental data points from [5, 10]. The transition from $S \rightarrow S + P$ is accurately predicted by the model. We have used for the DNA-DNA electrostatic interaction the model provided by Ubbink and Odijk [20]. For details on the calculations we refer the reader to [21, 22]. 61

4.9 Profile of θ'^2 as a function of θ . The value of $[\theta^{*'}(l_p; \hat{n})]^2$ at P_1 is the maximum value of θ'^2 . The convexity of θ'^2 as a function θ is shown in C.3. 61

4.10 (a) The constant helix solution is at point P_o where $\theta' = \theta'' = 0$ and $C_1 = F - G_{flu}^*$. (b) If $C_1 > F - G_{flu}^*$ (with M_3 and r independent of n) in Eq. (4.43), the profile of θ'^2 shifts upward and there is no value of θ for which $\theta'=0$. Consequently, the first boundary condition Eq. (4.44) can not be satisfied. 62

4.11 Profile of θ'^2 as a function of θ . In the red curve A_1 the point $\theta_{min} = 0$ corresponds to the helix solution that can be extended to infinity. The green curve A_2 shows the scenario corresponding to a fixed end loop length such that point P_o has a natural boundary condition. The blue curve A_3 describes the scenario where the end loop length is variable as described in C.4. Black squares represent point P_1 while black circles represent point P_o 62

4.12 Profile of θ as a function of s for case 3 described in section 4.4.8. In the calculations we have used $F = 4\text{pN}$ and $r = 1.5\text{nm}$, together with the internal interactions described by multivalent ions in [22]. $\theta(s)$ is nearly constant close to the midsection but diverges fast near the end points. We have used different values of M_3 with average values of θ_{mid} increasing as a function of M_3 . The inset shows case 2 described in section 4.4.8. Note that the plateau value of θ is only a function of the prescribed M_3 and it is the same for case 2 and case 3. 66

5.1	Partially clamped set-up realized in single molecule experiments using magnetic tweezers. The rod is subject to tension and twist. The rod is discretized so that we view it as a chain of segments of length l . The chain is free to fluctuate due to thermal effects, but the first and last segments (red) are parallel to the \mathbf{e}_3 axis. The first segment is clamped while the last one is free to have transverse displacements.	68
5.2	Comparison of our model with those valid in the high force long chain limit. a. External torque M_3 as a function of $\sigma = 10.5\Delta T w L k / N$, where N is the number of base pairs and σ is normalized linking number (see D.4). Graph corresponds to a short DNA molecule of $N=500$ base pairs subjected to a low tension force $F = 0.1\text{pN}$. b. Normalized extension $\langle x \rangle / L$ as a function of σ for a moderately short DNA molecule of $N=500$ base pairs subjected to a $F = 0.1\text{pN}$. The red solid line represents the full solution given by expressions Eq. (5.53) and Eq. (5.56), the black dashed line corresponds to the high force limit formulae Eq. (5.55) and Eq. (5.57). We have used the moduli corresponding to B-DNA (bending modulus $A = 45k_B T$ and twisting modulus $C = 100k_B T$). We have plotted the solution for $Q_n > 0$. As $Q_n^2 \rightarrow 0$, the solutions predicted by Eq. (5.55) and Eq. (5.57) using the high force (long chain) approximation diverge toward ∞	77

5.3 Duplex RNA behavior as a function of the twist-stretch coupling term g . In (a) we plot the normalized extension $\langle x \rangle / L$ as function of the normalized linking number σ (see D.4) and in (b) we plot the average torque $\langle M_3 \rangle$ as a function of σ . In the calculations we have used $F = 5\text{pN}$ which is a physiologically relevant force, $A = 63k_B T$ and $l_{RNA} = 0.279\text{nm}$ [23]. For the stretch modulus we use $S = 500\text{pN}$ which corresponds to $A = 61k_B T$ at 150mM monovalent salt concentration as documented in [24]. For C we use the value from molecular dynamic simulations $C_1 = 191k_B T$ [25]. The black solid lines in (a) and (b) correspond to a dsRNA chain where $g = 0$. In (a) the blue (dashed) curve ($g = 300\text{pNnm}$) and the gray (dash-dot) curve ($g = 160\text{pNnm}$) show a negative slope of the extension as function of the number of turns with the slope increasing as g increases. The red (dashed) curve ($g = -300\text{pNnm}$) and brown (dash-dot) curve ($g = -160\text{pNnm}$) show a positive slope of the extension as function of the number of turns with a larger magnitude of the slope as g decreases. Double stranded B-DNA has a negative twist-stretch coupling $\approx -85\text{pNnm}$ ([26]). In (b) we see that for positive g the torque $\langle M_3 \rangle$ slightly shifts to the right and for negative g the torque $\langle M_3 \rangle$ shifts slightly to the left. As the magnitude of g increases, the slope of $\langle M_3 \rangle$ as a function of σ decreases. This is expected since the effective torsional modulus is $\tilde{C} = C - g^2/S$ 78

5.4 Normalized extension $\langle x \rangle / L$ as a function of the normalized twist τ . (a) The top graph shows the result of our model using the properties of B-DNA in table 5.1. We plot the solution up to the point where Lipfert et al. saw the formation of plectonemic structures. In the remaining graphs we compare the results of our two-state model (black solid line) with the experiments in [27] (red circles). The experiments were performed using $N = 20.6\text{kp}$ and $F = 0.25\text{pN}$. The fraction a of the DNA where EtBr has bound is presented in each figure as a function of c_o . The values of the properties of each state, B-DNA and EtBr bound DNA, are presented in table 5.1. Using our model we have fitted the shift in internal twist per intercalated EtBr molecule as $\Delta tw_{EtBr} \approx -0.48\text{rad}$ per intercalation (-28°) which agrees with the values presented in [27]. 80

5.5	EtBr effects on rotation extension behavior at low and high forces. The graphs shows the normalized extension as a function the normalized twist τ for a set of different forces. Circles correspond to data from [27]. Forces used going from bottom to top of the graph are $F = [0.25, 0.5, 1, 3, 10]$ pN. We have plotted the solution up to the point where there is no formation of plectonemes as given by [27]. Properties used are presented in table 5.1.	83
5.6	Blue lines correspond to $F = 0.3$ pN and black lines to $F = 0.8$ pN. Green circles and red squares are the data series from [28] for $F = 0.3$ pN and $F = 0.8$ pN respectively. For the solid lines we have used the same properties as in Fig. 5.4. Dashed line calculated using $A_2 = 54k_B T$ which is the value obtained for $c_o = 10\mu\text{M}$ fitted by [28] using a single state WLC model. In (b) the blue solid line representing the torque $\langle M_3 \rangle$ at $F = 0.3$ pN has a slightly smaller slope than the black solid line (torque at $F = 0.8$ pN). Here we have not fitted any parameters; we use the values of a from comparing our model to [27] experiments. The extension $\langle x \rangle$, torque $\langle M_3 \rangle$ and the shift in τ seen in the experiments of [28] are predicted accurately with our model.	83

5.7 Predictions of the effects of netropsin on rotation extension behavior of DNA at low and high forces.

The graphs show normalized extension as a function the normalized twist τ for three different forces. State 1 is B-DNA and state 2 is DNA-netropsin at saturation $c_o = 100\mu M$ [27]. At saturation the fraction of the DNA chain where netropsin has bound is $a_{ne} = 0.2$. The red-solid curves are calculated for state 1 (properties for B-DNA are presented in table 5.1). For state 2 we have used $A_2 = A_1$ and $l_2 = l_1$ since netropsin has no significant effect on the extension and bending persistence length [27]. The blue-dashed line represents state 2 using $g_2 = g_1$, $S_2 = S_1$ and $C_2 = 110k_B T$ slightly larger than C_1 (since we expect the overwound DNA to have a larger twisting resistance). The gray-dot-dashed line represents state 2 using $g_2 = g_1$ and $C_2 = 110k_B T$ but $S_1 > S_2 = 100\text{pN}$. If after netropsin binds to DNA all the properties of the complex do not significantly change, then the prediction is the blue-dashed curves. If the stretch modulus changes (gray-dash-dot curve), experiments at low forces would not be able to capture the effects (blue-dashed curve and gray-dash-dotted curve are almost on top of each other at $F = 0.25\text{pN}$), but at large forces $F \sim 10\text{pN}$ the extension rotation curves would show a significant difference. 84

5.8 L-DNA as a single effective state. Circles correspond to the experimental data in [29]. Red-solid line

corresponds to $F = 8.5\text{pN}$, gray-dashed to $F = 12\text{pN}$, black-dashed to $F = 24\text{pN}$ and blue-dotted to $F = 36\text{pN}$. We plot the solutions from $\sigma_{end} = -1.7$ (critical number of turns at which L-DNA is fully formed after the transition) up to $\sigma^* \approx 2$ where [29] saw the external torque plateauing again (probably the start of a new phase transition). We use only large forces $F > 8.5\text{pN}$, which is the regime where effects of stretch are noticeable and to avoid the formation of plectonemic structures upon the addition of twist. In (b) we only present the experimental torque for $F = 8.5\text{pN}$ and $F = 36\text{pN}$ to avoid over-crowding the graph. The other two experimental curves that are not shown fall within the blue and red lines. 86

5.9 L-DNA as two states: S-DNA + Z-DNA. Circles correspond to the experimental data in [29]. Color

code is the same as in Fig. 5.8. The properties used for S-DNA and Z-DNA are presented in table 5.2. In D.6 we show the results using $A_2 = 11.6k_B T$ (S-DNA bending modulus) as measured by [30]. 87

5.10	Predictions from our two state model of L-DNA as ionic concentration c_s varies. For moderate salt $c_s \sim 0.1\text{M}$ we use S-DNA fraction to be $x_2 = 0.23$ as fitted to experiments in [29, 31] and for high salt $c_s \sim 1.8\text{M}$ we use $x_2 = 0.4$ as fitted to [31] experiments. Red line correspond to moderate salt concentration and blue-dashed line corresponds to high salt. For illustrative purposes we have used one of the large force $F = 36\text{pN}$ used in [29] experiments. For moderate salt $\sigma_2 \approx -1.7$ while for a high salt $\sigma_2 \approx -1.5$. From (a) we see that the $\langle x \rangle$ peak increases as the concentration increases and in (b) we see a slightly smaller slope for higher c_s . The properties used for S-DNA and Z-DNA are presented in table 5.2.	88
6.1	Procedure to fit values of A_2 and A_4 . Data points correspond to Zhang et al. [32] at $I = 3.5\text{mM}$ and $T = 12\text{C}$, where the critical force $F_c \approx 57.5\text{pN}$. Dashed lines correspond to extension z assuming no thermal fluctuation (see Eq. Eq. (6.5)). We fit $A_2 = 93\text{pN}$ and $A_4 = 500\text{pN}$ to the slope (far from the transition point) and to the change in extension δz between the compact and extended states. Inset: Schematic of double-well potential modeling the overstretching experiment. For $F < F_c$ system is in the B-DNA state. As the force increases, for $F > F_c$ the right well of the potential H is deeper as shown in the figure and the molecule transitions into the overstretched state. The right well represents the global minimum u^* in the overstretched form (S/M stands for S-DNA or M-DNA) and the left well corresponds to the local equilibrium of u in the B-form.	94
6.2	Force-extension relation during dsDNA overstretching. (a) Solid ($l = 25\text{bp}$) and dashed ($l = 100\text{bp}$) theoretical predictions for Zhang et al. [32] experiments (same conditions as in Fig. 6.1). (b) Data points correspond to two experiments in King et al. [33], while lines are theoretical predictions from our model using $l = 25\text{bp}$. Here $F_c = 69.5\text{pN}$ at 50 mM and $F_c = 63.5\text{pN}$ at 150 mM	97
6.3	DNA overstretching at $T = 24\text{C}$ for different ionic concentrations. Red Markers correspond to Zhang et al. [32] experiments and solid lines are theoretical predictions from our model. As the ionic strength decreases going from (a) trough (d), A_2 decreases and l increases. We use $A_4 \sim 500\text{pN}$ and F_c as measured in experiments: $F_c = [68.3, 63.5, 58.8, 50.5]\text{pN}$ going from panel (a) through (d).	98

- 6.4 Force-extension curve fit to Zhang et al. [30] experiments at higher salt concentration. We have used $A_2 = 120\text{pN}$ and $F_c = 67.2\text{pN}$. Although $l \approx 22\text{bp}$ provides a good overall fit, using $l \approx 15\text{bp}$ provides better agreement in the upper-right section of the curve as shown in the inset. 101
- 6.5 Variance σ_v^2 as a function of the force. Points correspond to Zhang et al. [30] experimental measurements at $I = 500\text{mM}$. Lines correspond to our theoretical predictions for different values of l using Eq. Eq. (6.24). The red solid line $l = 15\text{bp}$ agrees strongly with the experimental data for $\Delta F > 0$ (right side of the graph) , while on the left side of the graph the blue solid line $l = 30\text{bp}$ provides a better fit. The asymmetric fitting to the variance is in agreement with the results shown for the force extension curves at the same high ionic conditions in Fig. 6.4. Black solid line $l = 22\text{bp}$ is shown as an average fit for both sides of the graph. 101
- 6.6 Relaxation Rate $r = k_L + k_R$ in dsDNA overstretching experiments at $I = 150\text{mM}$. Blue markers correspond to 2pN force-step experiments in [34], where $F_c \approx 66\text{pN}$. Blue dashed line is the fit used in [34], where the authors combined their experimental measurements with Kramer-Bell theory. Black solid lines correspond to theoretical predictions from expressions in Eq. (6.30) using $l = 22$ and 25bp . Here we have used $\gamma \sim 200$ 103
- 6.7 Transient Kinetics: B-to-S and S-to-B. Force-jumps loading and unloading a dsDNA molecule at $T = 10\text{C}$. As before we use $\gamma = 200$, $l = 22\text{bp}$ and $\bar{L} = 1.67L$. Red solid line depicts the exponential evolution of n_s/n in time during loading phase. At $t_1 = 0$, F is instantaneously changed from $F^{(1)} \ll F_c$ to $F^2 = F_c$. The system reaches steady state in $\sim 1\text{s}$. Blue lines depict evolution during the unloading phase. At $t_2 = 2\text{s}$, F is instantaneously dropped applying a force step $S_F = F^{(3)} - F_c$. For $S_F = -2$ and -3pN (blue dashed lines), there is exponential decay behavior, while for $F^{(3)} \ll F_c$ (blue solid) $n_s \rightarrow 0$ rapidly. Inset: Shows lengthening ($\Delta L/\Delta L_e$) as a function of time for a *partially melted molecule* (S+M). Red markers are Bongini et al. [34] data, and black line is our prediction using the same parameters as in Fig.6.6 ($l \approx 25\text{bp}$). 105

A.1	Comparison between the solutions based on different internal energy models under the experimental conditions of Forth et al. [12]. The black solid line uses $U_1 = U_{PB}(r, \theta) + U_{conf-MS}(r, \theta)$, the red dotted line uses $U_2 = U_{MS}(r, \theta) + U_{conf-MS}(r, \theta)$ and the blue dashed line uses $U_3 = U_{PB}(r, \pi/6) + U_{conf-MS}(r, \pi/6)$ which is a function only of r . The value of the effective linear charge used is $\nu=8.06$ [nm ⁻¹].	111
A.2	Comparison between the solutions based on different internal energy models under the experimental conditions of Forth et al. [12]. The black solid line uses $U_1 = U_{PB}(r, \theta) + U_{conf}(r, \theta)$, the blue dashed line uses the internal energy model where the radial fluctuations in the plectoneme are constrained by the electrostatics $U_4 = U_{PB}(r, \theta, d_r) + U_{conf}(r, \theta, d_r)$ [20] and the red dotted line uses a model that neglects the configurational entropy contribution to the free energy $U_5 = U_{PB}(r, \theta)$. The qualitative behavior of M_3 and $d\Delta z/dn$ are independent of the internal energy models we have used, but the quantitative agreement with experimental data strongly depends on the choice of configurational entropy model. The value of the effective linear charge used is $\nu=8.06$ [nm ⁻¹].	112
A.3	Experimental data for the slope of the rotation-extension curve for two DNA templates taken from Forth et al. [12] for a 150mM salt concentration. We have used $\nu = 5.93\text{nm}^{-1}$	113
A.4	Experimental data for two DNA templates taken from Forth et al. [12] for a 150mM salt concentration. The experimental data of the plectonemic torque in Forth et al. [12] agrees qualitatively with our predictions but seems to match quantitatively our predicted values of $M_{critical}$. We have used $\nu = 5.93\text{nm}^{-1}$	114
A.5	External torque M_3 and critical torque $M_{critical}$ as a function of the external force F for a 7.9 kbp DNA template in a $c_o = 150\text{mM}$ salt concentration using the same parameters as in Fig. A.3. The data points correspond to the experimental values for the ‘buckling’ torque reported in Lipfert et al. [27].	114
A.6	Comparison of the the torque jump δM from our theory with the experiments in Brutzer et al. [10] at $c_o = 320\text{mM}$. Our model predicts that as the DNA length L increase δM decreases while as c_o decreases δM decreases. We have used $\nu = 10.00\text{nm}^{-1}$ as in the main text.	115

A.7	External torque M_3 as a function of the external force F for the different salt concentrations in Mosconi et al. [11]. We show the experimental values from Mosconi et al. [11] after adding a shift of +2.5pNm. We have used ν values from Table2.1 in the main text.	116
A.8	Spermidine: M_3 and r . The dashed lines correspond to the limiting value of the interaxial spacing which is $2r = 3.0\text{nm}$ for spermidine. We have used ν values from Table2.2 in the main text. . . .	117
A.9	Spermidine: the black solid curve shows the prediction of the slope $d\Delta z^e/d\sigma$ when $2r$ is larger than the interaxial spacing 3nm . The black dot shows the point where $2r = 3\text{nm}$, and from there on the black dashed line shows the prediction of the slope for $2r$ constant. The red dashed line shows the solution when $2r < 3\text{nm}$ is allowed to vary.	117
A.10	Spermidine: critical degree of supercoiling.	118
A.11	Spermine: M_3 and r . The dashed lines correspond to the limiting value of the interaxial spacing which is $2r = 2.9\text{nm}$ for spermine.	118
A.12	Spermine: critical degree of supercoiling.	119
A.13	Experimental data for the slopes $d\Delta z/dn$ for different salt concentrations. Triangles represent Brutzer et al. [10] data presented in reference [5]. We have denoted Brutzer et al. data with * next to the salt concentration value in the legend of the graph. Circles represent Mosconi et al. [11] data, which have been denoted with ** next to the salt concentration value in the legend. Squares represent Forth et al. [12] data for two different values of the DNA length. The data sets from Brutzer et al. and Mosconi et al. provide consistent slope values for the entire force range. The slopes from Forth et al.[4] are consistent with the rest of the data sets for moderate forces.	120
A.14	Experimental and Simulations data for the torque measurements M_3 for different salt concentrations. Triangles represent Maffeo et al. [5] simulations data. We have denoted Maffeo et al. data with * next to the salt concentration value in the legend of the graph. Circles represent Mosconi et al. [11] data, which have been denoted with ** next to the salt concentration value in the legend. Squares represent Forth et al. [12] data for two different values of the DNA length. The various data sets show disagreement in the torque values.	120

B.1	An increase in C_A : (a) increases the threshold force F_T , (b) decreases M_3 and θ , (c) increases $d\Delta z/dn$, and (d) causes r to deviate less from the constant value $D_{eq}/2$. An increase in K_b : (a) slightly decreases F_T , (b) causes no significant difference in r (lines are almost on top of each other), (c) decreases θ , and (d) increases both M_3 and $d\Delta z/dn$. An increase in D_{eq} : (a) significantly decreases F_T , (b) shifts the solutions for θ and M_3 to the left along the force axis, (c) increases r , and (d) decreases $d\Delta z/dn$. We have used $K_t = 86k_B T$	122
B.2	Toroid dimensions as a function of the applied number of turns n . The solid lines correspond to the solution for toroids with no twist and dashed lines to toroids with twist.	123

B.3 Energy density of competing structures as a function of the number of turns n for surface tension parameter $\alpha = 0.013$. The parameters used correspond to the theoretical solution for the description of Besteman et al. experiments in spermine in Fig. 2 in the main text. The black solid line corresponds to the straight configuration, the blue solid line to the straight + supercoiled coexistence, the blue dashed line correspond to the pure plectonemic phase, the red solid line correspond to toroids with twist and the green solid line correspond to toroids with no twist. Notice that the solution for toroids with no twist is only valid up to the point where the applied number of turns n is smaller than or equal to the toroid winding number N_g^o . The point where $n = N_g^o$ corresponds to the intersection of the red line and the green solid line. We have also plotted the energy density for torsionally unconstrained toroids depicted by the green dashed line. We point out that the torsionally unconstrained toroids which carry no twist are not a realistic solution for this problem. The link constraint on the DNA molecule must be taken into account otherwise toroids (unconstrained) would become energetically more favorable than supercoiled structures as n increases (green dashed line crosses the blue dashed line). The value of $n_{critical}$ (transition $S \rightarrow S + P$) is given by the intersection of the black solid line and blue solid line, the value of n_{pp} (transition $S + P \rightarrow PP$) is given by the intersection of the blue solid line with the blue dashed line. We have plotted the energy density for three representative forces F : (a) $F = F_T \approx 2.4\text{pN}$, supercoils are the favorable state for $n \geq n_{critical}$; (b) $F = 3\text{pN} > F_T$, for larger forces the intersection between the black solid line and blue line moves toward larger values of $n_{critical}$; (c) $F = 1.5\text{pN} < F_T$, there is no solution for the coexistence of straight + supercoiled, toroids are the preferred state up to n_{pp} which is the point where the PP solution starts. 125

B.4	Energy density of competing structures as a function of the number of turns n for surface tension parameter $\alpha = 0.0003$. The rest of the parameters are the same as in Fig 2 in the main text. Lines represent the same energy densities as in Fig. B.3, but here we do not present the solution for torsionally unconstrained toroids. Changing the value of α changes the preferred state of the DNA molecule as a function of the applied force F . Again we have plot the energy density of the competing structures for three representative forces F : (a) $F = F_T \approx 2.4\text{pN}$, toroids are the preferred state up the intersection point of the blue dashed line and red solid line; (b) as the force increases $F = 2.5\text{pN} < F_T$, both the S and $S + P$ decrease its energy density, and at $n = 0$ the S and $T(\text{twist})$ have the same energy; (c) as the force further increases $F = 3\text{pN} > F_T$, the S configuration becomes the preferred state up to $n_{critical}$ and supercoils are the favorable state for $n \geq n_{critical}$	126
B.5	Energetic contribution in the toroidal configuration including twist. We use the same parameters as in Fig. 2 in the main text.	126
B.6	Energetic contributions in the supercoiled-straight configuration. We use the same parameters as in Fig. 2 in the main text.	127
D.1	Normalized extension $\langle x \rangle / L$ as a function of the normalized twist τ . We compare the results of our 2-state chain model (black solid line) with the experiments in [27] (red circles). The experiments were performed using $N = 20.6\text{kp}$ and $F = 0.25\text{pN}$. The fraction a of the DNA where EtBr has bound is presented in each figure as a function of c_o . The values of the properties of B-DNA are presented in table 5.1. For EtBr we use the same properties as in the main text but replace $S_2 = 1220\text{pN}$ and $g_2 = 385\text{pNnm}$	156
D.2	(a) $F = [0.25, 0.5, 1, 3, 10]$ pN and $a_o(F) = [0.52, 0.55, 0.57, 0.60, 0.67]$. (b) Same values but assume linear relationship between applied applied turns (torque) and fraction $a(F, \tau) = a_o(F) - 1.3\tau$. This linear approximation has been also used in [28]. Both in (a) and (b) for each force the unwinding shift Δz_{Etbr} per intercalation event is different. We have used the same properties as in Fig. D.1. .	156

D.3 Blue lines correspond to $F = 0.3\text{pN}$ and black lines to $F = 0.8\text{pN}$. Green circles and red squares are the data series from [28] for $F = 0.3\text{pN}$ and $F = 0.8\text{pN}$ respectively. For the solid lines we have used the same properties as in Fig. D.1. Dashed line calculated using $A_2 = 54k_B T$ which is the value obtained for $c_o = 10\mu\text{M}$ fitted by [28] using single state WLC model. 157

D.4 L-DNA as two states: S-DNA + Z-DNA. Circles correspond to the experimental data in [29]. Color code is the same as in Fig. 5.8. Red-solid line corresponds to $F = 8.5\text{pN}$, gray-dash to $F = 12\text{pN}$, black-dash to $F = 24\text{pN}$ and blue-dot to $F = 36\text{pN}$. Here we use $A_2 = 11.6k_B T$ for the bending modulus of S-DNA as measured by [30], the rest of the properties used for S-DNA and Z-DNA are presented in table 5.2. 157

E.1 Comparison of our theoretical predictions (solid lines) of the force-extension curves with experimental data in Bianco et al. [35] and Bongini et al [34]. Since experimental conditions are very similar to those in Fig 3(a), we use the same parameters $A_2 = 105.6\text{pN}$ and $A_4 = 500\text{pN}$. The predictions with $l = 22\text{bp}$ are in good agreement with experiment. Here we have used $F_c \approx 65.5\text{pN}$ as reported in [34, 35]. 158

E.2 Change in extension ΔL_e at equilibrium (steady state) corresponding to 1pN force-step protocol. Conditions are the same as in Bianco et al. [35] experiments at $T = 25\text{C}$ (See Fig. E.1) . Red crosses correspond to Bianco et al. [35] Monte Carlo simulations using a segmented chain composed of 1936 two-state units, each of length 25bp ($L = 48.4\text{kbp}$). Blue circles are the predictions of our analytical model using the same $l = 25\text{bp}$. Besides the value of l , here we have used the same parameters as in Fig. E.1. 158

E.3 Temperature effects on the force-extension curve at $I = 500\text{mM}$. Red circles ($T = 10\text{C}$) and blue markers $T = 14\text{C}$ correspond to [30] experimental data. Lines are our theoretical predictions using $l = 15\text{bp}$, $A_2 = 120\text{pN}$, $A_4 = 500\text{pN}$ (same as in Fig. 4 in the main text), where we have used the phenomenological relation Eq. (E.1) with $\eta \approx 0.1[\text{pN/K}]$ and $F^o \approx 50\text{pN}$ [30]. In the inset we use increments of 2C , going from 10C (red) up to 16C (green) using same parameters as in main figure. 159

- E.4 Relaxation rate $r = k_L + k_R$ in dsDNA overstretching experiments at $I = 150\text{mM}$. Blue markers corresponds to *square-wave protocol* data in [34], where $F_c \approx 66\text{pN}$. Black solid line correspond to our theoretical predictions using $l = 22\text{bp}$. The *square-wave protocol* consists of applying a large force step $\sim 30\text{pN}$ during loading, waiting until system reaches equilibrium, and subsequently unloading the molecule with a force-step of the same magnitude [34]. As in the main text, where we fitted r to 2pN force-step experiments, we have used $\gamma \sim 200$ 159
- E.5 Effect of A_4 on the average stretch $\langle u \rangle$ and on the probability of segments in the overstretched state (P_s). Here we have used parameters corresponding to the overstretching transition in Fig. E.1, where $A_2/A_4 = 0.2$, $l = 22\text{bp}$ and $T = 25\text{C}$. Both in (a) and (b) the black solid line corresponds to $A_4 = 500\text{pN}$ (as used throughout the main text), blue dashed line corresponds to $A_4 = 1000\text{pN}$ and red-dashed line corresponds to $A_4 = 250\text{pN}$. In (a) we present the solutions for $\langle u(F) \rangle$. As seen from the graph there is almost no difference between the three solutions in the range of forces δF where the transition takes place and δF remains approximately constant. For $|\Delta F| = |F - F_c| \gg 0$ the three solutions have different slopes. In (b) we present the solutions for $P_s(F)$, where it evident that both, P_s and the width of the transition, are unaffected by A_4 . The inset shows a magnification of P_s near $F \in [3, 4]\text{pN}$ 160
- E.6 Potential H_j landscape. We have used $A_2 = 100\text{pN}$, $A_4 = 500[\text{pN}]$ and $l \approx 25\text{bp}$. These values correspond to the overstretching transition fits found in section 4 of the main text at room temperature. At $C = 2\text{pN}$ the left well is deeper and $k_R > k_L$. From the graph it is clear that the energy barrier $E_t^L \gg k_b T$ and $\bar{u}_n \approx u_b = 0$ 162

E.7	Ratio Θ^f normalized by Θ evaluated at $l_j = 10\text{bp}$. Here we use $T = 25C$ and the same parameters as in Fig. 6.6. Blue-dashed line correspond to solution with Neumann B.Cs (N) using Eq. (E.99), black-solid line is the solution with Periodic B.Cs (P) using Eq. (E.92) and red solid line is the solution using expression Eq. (6.30) in the main text. For the red-solid line we use $\Gamma = k_L = k_R$ (since $C = 0$ here). Inset shows the ratio Θ^f computed from expression Eq. (6.30) in the main text at $T = 25C$ and using the same parameters as in Fig.6.3(d). In the inset we have normalized by Θ evaluated at $l_j = 60\text{bp}$. It is clear that as the cooperativity l increases, the time it takes to reach steady state $\propto r^{-1}$ increases rapidly.	170
E.8	Cartoon showing periodic boundary conditions. Depending on the boundary conditions (B.Cs), the domain walls can have a positive slope (kinks) or negative slope (anti-kink).	174
E.9	Comparison of our theoretical predictions (lines) for the fraction of segments in the overstretched state P_s and Bosaeus et al. [36] experimental results (red markers). (a) Theoretical fit using $x \approx 0.5$ ($A_2 = 60\text{pN}$). (b) Theoretical fit using $x \approx 0.7$ ($A_2 = 105\text{pN}$).	178

Chapter 1

Introduction

Since the structure of double stranded DNA was discovered about six decades ago, there have been great developments in the fields of biology, genetics, biophysics and numerous others. It is the objective of our work to carry on with these developments by providing new insights and understanding of DNA behavior/function using mechanics as our primary tool. To achieve this goal we focus on building analytical models for DNA molecules using rod theory and statistical mechanics.

The mechanical properties of DNA have been thoroughly studied since its discovery not only because of their biological significance, but also because DNA is a model system to illuminate the physics of semi-flexible polymers. Indeed, DNA has been studied as an elastic rod by many authors over the last few decades. Single molecule mechanics experiments on DNA have made it possible to probe its mechanical, entropic and electrostatic properties at the scale of a few base-pairs. A key experiment used for this purpose is one in which a tensile force is applied on the DNA while simultaneously twisting it using optical or magnetic tweezers. Through these experiments it has been possible to isolate certain aspects of the DNA behavior revealing fascinating phenomena. For instance, single molecule experiments are used today to study the DNA collapse in the presence of ions leading to the formation of plectonemes and toroidal structures [15, 21, 22, 37], which can play an important role in DNA compaction and storage inside the cell. Furthermore, these experiments have been essential to characterize newly discovered DNA-forms that arise during phase transitions driven by external mechanical parameters such as force and torque [29, 30, 32, 38].

Before we proceed to a detailed description of the topics covered in this Thesis we briefly discuss DNA structure, topology and the tools used to model DNA as a fluctuating filament.

1.1 Background

In this introductory section we do not present any original material, but we give a brief background for our research work. The concepts and ideas presented in this chapter can be found in a number of references in the literature [9, 39–44].

DNA structure: Double stranded DNA (dsDNA) is a polynucleotide composed of two interwinded strands that run in opposite directions [39]. Each strand is a polymer made up of monomers units called nucleotides. The DNA backbone is a structure with an alternating sugar-phosphate sequence, where the phosphate groups and deoxyribose sugar are located on the surface, while the nitrogenous bases are on the inside of the structure [39, 42]. There are four different bases (A-Adenine, G-Guanine, C-Cytosine and T-Thymine). The two chains are held together in the inside through base-pairing that results from hydrogen bonds linking complementary bases (A-T and G-C) of each chain [42]. In B-DNA, which is the most common one appearing in nature, the base-pairs

(bp) are almost completely flat and perpendicular to the long axis of the right-handed helix [42]. The base-pairs are stacked on top of each other like a ladder and they are separated by ~ 0.34 nm [42]. The effective (crystallographic) radius of the double stranded molecule is ~ 1 nm [45]. The geometrical properties of standard B-DNA vary with parameters such as solvent ionic concentration, temperature and base-pair sequence [43]. Under physiological conditions, the averaged B-DNA helical repeat is ~ 10.4 - 10.5 base-pairs [44].

Storage of genetic information requires that long dsDNA molecules be compacted, condensed and folded in order to fit inside the cell [22, 42, 46]. DNA packing is a complex phenomena where mechanical energies due to bending and torsional resistance of the molecule compete with electrostatic interactions [22, 42]. The DNA compacted structure (plectonemes, globules, etc) plays an important functional role since the genetic information needs to be accessed in a robust manner enabling proper DNA functions such as replication, transcription and repair [21, 22, 46]. Recently, Shao et al. [46] presented a set of results from single molecule experiments that show that adding two polyamines (spermine or spermidine) produced significant changes in B-DNA behavior under physiological conditions. The presence of these polyamines decreased the required torque for the appearance of the interwound supercoiled structures in DNA (See Fig. 1.1(a)). Polyamines are organic compounds known to help in regulating cell growth and gene transcription [46], but high concentrations of polyamines are linked with certain cancers [46]. The method by which polyamines act in cellular metabolism to either help normal cell growth or cancerous growth are still not fully understood, nor are the structural changes in DNA associated with the presence of these organic compounds. Shao et al. suggested that for DNA subjected to an applied number of turns, polyamines may favor B-DNA (stabilize the right handed base-pair structure) instead of promoting structural changes. So, although polymerases are known to twist and stretch dsDNA (to the point of denaturation), natural polyamines promote supercoiling, which is essential for B-DNA packing [46].

Understanding how the molecule behaves in the presences of mechanical loads is a stepping stone in the theoretical description of DNA packing and DNA structural changes. Models for this phenomena can potentially help to gain even further insight in complex processes inside a living cell (transcription, replication and recombination). As described in section 1.2, the main tools used in the following chapters to describe the DNA mechanical response are Kirchhoff's theory of rods and statistical mechanics. Next we present some of the key concepts behind the model of DNA as a one-dimensional rod in the presence of thermal fluctuations.

Rod Theory: A starting point to study DNA mechanics as a rod filament is to use Kirchhoff's classical theory for inextensible and unshereable rods. Here we give a brief treatment of the kinematics and mechanics of the rod theory used in this Thesis. We refer the reader to Chouaieb and Maddocks [47] for a detailed discussion of the subject. Within Kirchhoff's rod description, the coordinates of the center-line are given by the position vector $\mathbf{r}(s)$ with s being the arc-length. In the undeformed reference configuration $\mathbf{r}(s)$ is described in a three dimensional orthonormal fixed frame $\{\mathbf{e}_i(s)\}$ for $i = 1, 2, 3$. For convenience we will use the notation $(\dots)' = \frac{d(\dots)}{ds}$. The tangent of the curve is given by the unit vector $\mathbf{t} = \mathbf{r}'(s)$ and the curvature of the curve $\mathbf{r}(s)$ is the scalar valued function $\kappa(s)$ defined by:

$$|\mathbf{t}'| = |\mathbf{r}''(s)| = \kappa(s).$$

As the rod deforms the reference frame $\{\mathbf{e}_i(s)\}$ is not suitable to study the physical twist of a rod. Therefore, it is necessary to introduce a material frame in the deformed configuration that allows to describe the orientation of the material points in the rod's cross section [47]. The space curve $\mathbf{r}(s)$ is then associated with an orthonormal director frame of vectors [47]:

$$\{\mathbf{d}_i(s)\}, \quad i = 1, 2, 3, \quad -l < s < l \quad (1.1)$$

It is conventional to pick $\mathbf{t} = \mathbf{d}_3(s)$, such that $\mathbf{d}_1(s)$ and $\mathbf{d}_2(s)$ span the normal plane to \mathbf{r} , so their directions encode the orientation of the material cross-section. As the rod deforms, the evolution of $\{\mathbf{d}_i(s)\}$ is governed by the director frame equations:

$$\begin{aligned}\mathbf{d}_i &= d_{i1}\mathbf{e}_1 + d_{i2}\mathbf{e}_2 + d_{i3}\mathbf{e}_3 \\ \mathbf{d}'_i &= \mathbf{u} \times \mathbf{d}_i, \quad \mathbf{u} = (u_1\mathbf{d}_1 + u_2\mathbf{d}_2 + u_3\mathbf{d}_3),\end{aligned}\tag{1.2}$$

where u_1 and u_2 are the bending component of the strain and u_3 is physical twist [47].

The mechanics of the problem within Kirchhoff's theory are based on the constitutive assumption that the stresses depend linearly on the strains. In other words, for a Kirchhoff rod there is a convex strain-energy density function W , such that the moment (stress) \mathbf{m} is given by the constitutive relation [47]:

$$\mathbf{m} = \frac{\partial W}{\partial \mathbf{u}}(\mathbf{u} - \hat{\mathbf{u}}),\tag{1.3}$$

where $\hat{\mathbf{u}}$ are the strains in the reference (minimum energy unstressed) configuration. Here, the strain-energy density does not depend explicitly on the arc-length and $\hat{\mathbf{u}}$ is constant, such that in the reference state $\mathbf{m} = 0$ [47]. Therefore W is only a function of the strain \mathbf{u} . Taking the approximate W as a quadratic function of the strain [47]:

$$W(\mathbf{u} - \hat{\mathbf{u}}) = \frac{1}{2}(\mathbf{u} - \hat{\mathbf{u}})\mathbf{K} \cdot (\mathbf{u} - \hat{\mathbf{u}})\tag{1.4}$$

where \mathbf{K} is a positive definite¹ symmetric matrix, the moment \mathbf{m} is given by the constitutive equation [47]:

$$\mathbf{m} = \mathbf{K}(\mathbf{u} - \hat{\mathbf{u}})\tag{1.5}$$

The strain-energy function W is said to be isotropic if W is invariant under rotations of its arguments about \mathbf{d}_3 [47]:

$$\frac{\partial}{\partial \alpha} W(\mathbf{Q}(\alpha)\mathbf{w}) = 0$$

where:

$$\mathbf{Q}(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad 0 < \alpha < 2\pi$$

The isotropic constraint plus equation Eq. (1.4) requires that the only non-zero components of \mathbf{K} are $K_{11} = K_{22} = K_b$ and $K_{33} = K_t$. Thus, equation Eq. (1.5) is reduced to the following constitutive relation:

$$\mathbf{m} = K_b(u_1 - \hat{u}_1)\mathbf{d}_1 + K_b(u_2 - \hat{u}_2)\mathbf{d}_2 + K_t(u_3 - \hat{u}_3)\mathbf{d}_3\tag{1.6}$$

where K_b are the bending moduli and K_t is the twisting modulus.

The Cosserat theory of rods can be considered as an extension of Kirchhoff rod theory in the sense that it allows us to describe extensible and shearable rods [47]. Hence, the kinematics of the Cosserat rod are described by two strain vectors \mathbf{v} and \mathbf{u} by means of equation Eq. (1.2) and the relation:

$$\mathbf{r}'(s) = \mathbf{v}(s).$$

The components v_i are the strain variables of the axial curve, where v_3 is related to the stretch, while v_1 and v_2 are related to the shear. If we impose the inextensibility and unshereability constraint on the Cosserat rod we recover the Kirchoff rod [47]:

- The rod is said to be inextensible if in any configuration $|\mathbf{r}'| = 1$.
- The rod is said to be unshereable if v_1 and v_2 are identically zero.

The inextensibility and unshereability are manifested in the center line equation constraint:

$$\mathbf{r}' = \mathbf{d}_3, \quad (1.7)$$

which means that \mathbf{r}' is a unit vector in any configuration.

Thermal Effects: The description of DNA response to mechanical loads requires more than the classical treatment of rod mechanics, since in the world of molecular biology we need to consider thermal effects. The relevant units for the description of DNA fragments are pico-Newtons (pN) and nanometers (nm), and hence a relevant energetic measure of DNA filaments is that of thermal energy at room temperature $\sim k_B T \approx 4.13$ pNnm (k_B is the Boltzman constant and T is the temperature in Kelvin). DNA filaments are constantly undergoing Brownian fluctuations. Consequently, entropic contributions to the energy of the filament are significant and they can not be neglected in favor of the elastic and electrostatic energies.

The structure of double stranded DNA provides it with a large bending resistance, so the typical length scales ξ_p (persistence length) over which thermal fluctuations bend the filament axis is about $\sim 40 - 60$ nm [44, 48, 49]. The persistence length is defined through

$$\langle \mathbf{t}(s_o) \cdot \mathbf{t}(s_o + s) \rangle = \exp(-s/\xi_p), \quad (1.8)$$

where \mathbf{t} is the unit tangent vector of the filament and the angular brackets indicate an ensemble average. So ξ_p is the length over which correlations in the direction of the tangent vector are lost, and consequently DNA can be modeled as a flexible chain.

Smith et al. in 1992 [50] performed the first study on DNA entropic elasticity and numerous experimental groups have successfully followed Smith et al. since then. The most basic model to capture entropic effects is the freely joined chain FJC model. The FJC model sees a polymer as a discrete chain of uncorrelated N segments, and the energy of the polymer is given by [44]:

$$E = k_B T \sum_i^N Fl \cos \theta_i \quad (1.9)$$

where θ_i is angle between the direction of the force and the tangent vector to each segment i , and l is the length of each segment (called the Kuhn length). The partition function for a segment i is the summation of the Boltzman weight for all possible configurations [44, 51]:

$$Z_i = 2\pi \int_0^\pi \sin \theta_i \exp \left[\frac{Fl \cos \theta_i}{k_B T} \right] d\theta_i = 4\pi \sinh \left(\frac{Fl}{k_B T} \right) \frac{k_B T}{Fl}, \quad (1.10)$$

and the total partition function of the system of independent segments is [51]:

$$Z = \prod_{i=1}^N Z_i = \left[4\pi \sinh \left(\frac{Fl}{k_B T} \right) \frac{k_B T}{Fl} \right]^N. \quad (1.11)$$

The free energy of the system is:

$$G = -k_B T \ln Z = - \left[4\pi \sinh \left(\frac{Fl}{k_B T} \right) \frac{k_B T}{Fl} \right]^N, \quad (1.12)$$

¹A second order tensor \mathbf{T} is said to be positive definite if $\mathbf{v} \cdot (\mathbf{T}\mathbf{v}) > 0$ for all non-zero \mathbf{v} in \mathbb{E}

such that the full 3-D force-extension relation in this model is given by :

$$\frac{\langle x \rangle}{L} = -\frac{1}{Nl} \frac{\partial G}{\partial F} = -\frac{k_B T}{Fl} + \coth\left(\frac{Fl}{k_B T}\right). \quad (1.13)$$

The FJC model reveals that the entropic force due to thermal motion is proportional to temperature. Yet, this simple model fails to account for a realistic description of DNA filaments as it does not account for bending resistance and although it fits DNA elasticity at low forces (~ 1 pN) it has a significant error for larger forces.

To overcome this difficulty Marko and Siggia [48, 52] developed a more sophisticated model (WLC - worm like chain) based on a continuous flexible chain with bending contributions. The bending energy contribution in the WLC is given by [44]:

$$A \int_0^L \left(\frac{d\mathbf{t}}{ds}\right)^2 ds, \quad (1.14)$$

where $A = \xi_p k_b T$ is the bending modulus, L is the length of the chain, \mathbf{t} is the tangent vector to the axial curve describing the chain and s is the arc-length. The energy in the WLC model can be computed analytically at low and high forces. An analytic formula for the force-extension relation that fits the high force regime is [49]:

$$\frac{\langle x \rangle}{L} = 1 - \frac{1}{2\sqrt{AF}}. \quad (1.15)$$

Polymers in which their monomers are connected by single covalent bonds tend not have torsional resistance [11]. This is not the case in dsDNA (in the absence of nicks and breaks), where the base-pairing inside the dsDNA backbone prevents the twist to be relaxed. This particular feature has very important biological implications that are not yet fully understood, and it will be our main objective in the following chapters of this work to explore torsionally constrained DNA. To do so we need to first introduce the reader to some topological concepts used to describe twisted filaments.

DNA Topology: We already mentioned that the dsDNA molecule is double stranded. So when we talk about DNA we will mean the double stranded (dsDNA) structure. We will refer as ssDNA to each of the single strands composing the DNA molecule. Hence, a DNA molecule can be viewed as an assemblage of two continuous curves (each ssDNA strand) [42]. The DNA axial curve passes through the center of the base-pairs between each ssDNA curve [42] (see Figure 1.1). Under certain conditions, dsDNA can be subjected to loads that lead only to deformation of the its axial curve in space, but the ssDNA strands do not break. This scenario arise in dsDNA problems involving knotting, catenation and supercoiling. In the following chapters we will address the problem of supercoiling, therefore it is important to familiarize the reader with the topological concepts of linking number, writhe and twist .

Imagine for a moment that the DNA molecule is closed at its ends (like a ring), such that the two ssDNA strands are in fact two closed curves that are entangled around each other. Then for any two closed curves C_1 and C_2 , the linking number Lk describes quantitatively the fact that the curves are interwound and cannot be separated without cutting one of them [42]. Intuitively, the linking number of DNA can be found by counting the number of passages required in order to unlink the two ssDNA strands. Lk is a topological invariant, meaning it can not be affected by changing the geometry unless externally modified by cutting a strand or inputing more crossings between the strands [42]. The linking number Lk is the sum of two geometrical properties of the curve [42]:

$$Lk = Wr + Tw. \quad (1.16)$$

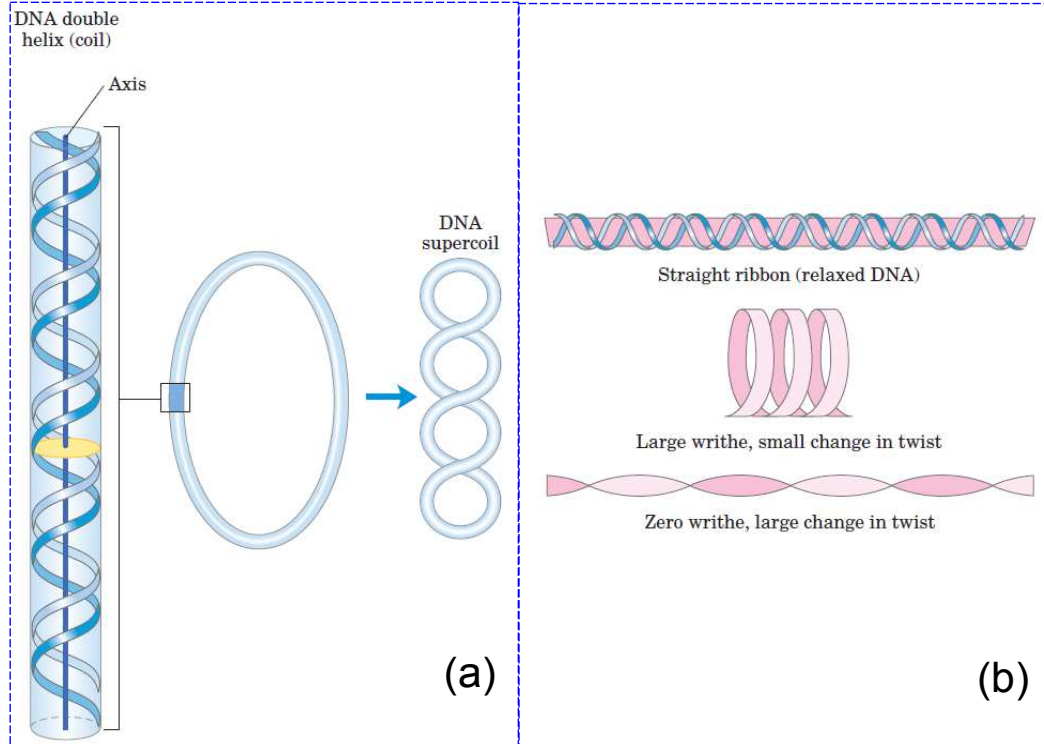


Figure 1.1: A ribbon model for illustrating twist and writhe. Figure 1.1 (a) and (b) are reproduced from Figs 24-11 and 24-18 in [9]. The ribbon in (a) represents the axis of a relaxed DNA molecule. Changes in linking number usually lead to the formation of supercoils, also known as plectonemes, which are DNA extruded interwound structures. In (b) it is shown how DNA twisting generates a torsional strain that can either produce changes in the writhe or a change in twist of the molecule.

The writhe Wr is a global quantity of the curve and it quantifies the chiral deformation of the curve [42]. In the case of DNA, Wr is related to the global path adopted by the axial curve. One can intuitively define the writhe Wr as the number of times the axial curve crosses itself. So when $Wr > 1$ in DNA, this implies the formation of loops, solenoids or plectonemes. The second property is the twist Tw of the curve and it is a local quantity that quantifies the winding of the curves around each other [42]. Note that the definition of linking number can be extended to a linear molecule (not closed physically), by joining its ends with a space curve with zero linking number [53]. Figure 1.1 depicts graphically the concepts of writhe and twist. A DNA molecule without intrinsic curvature and under no torsional stress has zero writhe, so that the natural linking number is $Lk_o = Tw_o = n/h_p$. Tw_o is the natural twist of the DNA molecule and is equal to the total number of base pairs n divided by the average number of base pairs with a helical pitch $h_p = 10.4\text{bp}$ (for B-DNA) [44].

Twisting and Stretching DNA: Magnetic or optical tweezers [10–12, 27] can be used to stretch and change the linking number of a DNA molecule. In experiments the change in extension is studied at a given force F , while the molecule is twisted by controlling the number of turns n . Depending on the value of force F and the sign of n , one can observe different regimes of the DNA molecule in the extension-rotation curves [54]. For instance, for low forces as the number of turns is increased, it is well known that for a critical number of turns $n_{critical}$ there is a sudden decrease in the extension, and from there on the extension drops almost linearly with n [10, 12]. A simple end loop model was proposed by Strick et al. [44] to describe this DNA transition. The model relies on equating the

torsional work done upon adding one turn at the transition point to the energy stored in the loop [44]:

$$2\pi M = \pi \frac{A}{R} + 2\pi RF, \quad (1.17)$$

where $2\pi M$ is the torsional work, A is the bending modulus and F is the applied force on a loop of radius R . This model yield $M^2 = 2AF$ as the transition torque value, which has been shown to overestimate the experimental and simulation values of M [55]. So more sophisticated models haven been proposed to account for this transition [45, 55] in which there is a coexistence of straight B-DNA with supercoiled DNA. These new models have captured the qualitative features of this transition, but the quantitative description requires accounting of the electrostatic and thermal effects.

At low-moderate forces and low-moderate number of turns, B-DNA stays in a straight configuration, but thermal fluctuations reduce its effective length. To study torsionally constrained DNA in this regime, the WLC model can be modified to include the effects of torsion. This description is sometimes referred to as the rod-like chain RLC [44], and its free energy is:

$$E_{RLC} = E_{WLC} + \int_0^L \frac{C}{2} u_3^2(s) ds, \quad (1.18)$$

where u_3 is the local twist and C is the effective twist modulus. A detailed analysis of this model can be found in Moroz and Nelson [40, 56].

It is important to point out that the continuous rod-like chain model can not describe DNA structural transitions. At moderate to large F and positive n B-DNA is known to transform to P-DNA and for negative n it can make a transition into denaturated DNA, Z-DNA or L-DNA [44, 57]. DNA structural transitions driven by torsion is one of the topics we will explore in chapter 5. There we will make use of a more complete model of the elastic energy that includes the effects of bending, twisting, stretching and the coupling between twist-stretch [39]:

$$E = E_{stretch} + E_{bend} + E_{twist} + E_{stretch-twist}. \quad (1.19)$$

With this brief background, we are in position to move into a more detailed description of each problem that we have tackled in our research. In this work we address five specific topics, in each of which we describe the DNA behavior in the presence of mechanical loads. We have organized the theory and results of each of topic as a chapter of the Thesis. In the next section we give a summary of the work contained in each chapter².

1.2 Overview of this Thesis

- In Chapter 2, we develop an elastic-isotropic rod model for twisted DNA in the plectonemic regime. We account for DNA elasticity, electrostatic interactions and entropic effects due to thermal fluctuations. We apply our model to single molecule experiments on a DNA molecule attached to a substrate at one end, while subjected to a tensile force and twisted by a given number of turns at the other end. The free energy of the DNA molecule is minimized subject to the imposed end rotations. We compute values of the torsional stress, radius, helical angle, and key features of the rotation-extension curves. We also include in our model the end loop energetic contributions and obtain estimates for the jumps in the measured torque and extension of the DNA molecule seen in experiments. We find that while the general trends seen in experiments are captured simply by rod mechanics, the details can be accounted for only

²The results presented in the following chapters (2-6) have resulted in the publication of five independent articles. Chapter 2 (Appendix A), Chapter 3 (Appendix B), Chapter 4 (Appendix C), Chapter 5 (Appendix D) and Chapter 6 (Appendix E) are verbatim reproductions of the work in Argudo and Purohit [21, 22, 58, 59] and [60] respectively.

with a proper choice of electrostatic and entropic interactions. We perform calculations with different ionic concentrations and show that our model yields excellent fits to mechanical data from a large number of experiments. Our methods also allow us to consider scenarios where we have multiple plectonemes or a series of loops forming in the DNA instead of plectonemes. For a given choice of electrostatic and entropic interactions we find there is a range of forces in which both regimes can co-exist due to thermal motion.

- In Chapter 3, the condensation of free DNA into toroidal structures in the presence of multivalent ions is described. Recent single molecule experiments have shown that condensation into toroids occurs even when the DNA molecule is subjected to tensile forces. In chapter 3, we show that the combined tension and torsion of DNA in the presence of condensing agents dramatically modifies this picture by introducing supercoiled DNA as a competing structure in addition to toroids. We combine a fluctuating elastic rod model of DNA with phenomenological models for DNA interaction in the presence of condensing agents to compute the minimum energy configuration for given tension and end rotations. We show that for each tension there is a critical number of end rotations above which the supercoiled solution is preferred and below which toroids are the preferred state. Our results closely match recent extension rotation experiments on DNA in the presence of spermine and other condensing agents. Motivated by this we construct a phase diagram for the preferred DNA states as a function of tension and applied end rotations and identify a region where new experiments or simulations are needed to determine the preferred state.
- In Chapter 4, we present an elastic-isotropic rod model for superhelical DNA structures where we relax the constant curvature assumption so that the helical angle is varying as a function of the arc-length. Our motivation for a variable helical angle comes from some experiments and simulations on DNA braids where complex superhelical structures have been observed. The helical solutions are minimizers of a free energy consisting of elastic, entropic and electrostatic terms. These minimizers are obtained within a variational framework where the end-points of the helices are allowed to be variable so that the length of the superhelix is computed as part of the solution. Considering variable curvature solutions brings up the possibility of finding more complex DNA structures because for two (or more) interwound helices there is a geometrical lock-up helical angle which puts a limit on the length of a superhelix. We perform calculations with different ionic concentrations and study the effects of lock up for braided structures. We also extend the variable curvature model to study the formation of plectonemes in the presence of multivalent salts where the supercoiling radius can be regarded as a constant prescribed by the balance of attractive and repulsive forces in DNA-DNA interactions, and provide analytical solutions in terms of elliptic functions for the supercoil parameters.
- In Chapter 5, we discuss the statistical mechanics of a heterogeneous elastic rod with bending, twisting and stretching. Our model goes beyond earlier works where only homogeneous rods were considered in the limit of high forces and long lengths. Our methods allow us to consider shorter fluctuating rods for which boundary conditions can play an important role. We use our theory to study structural transitions in torsionally constrained DNA where there is coexistence of states with different effective properties. In particular, we examine whether a newly discovered left-handed DNA conformation called L-DNA is a mixture of two known states. We also use our model to investigate the mechanical effects of the binding of small molecules to DNA. For both these applications we make experimentally falsifiable predictions.
- It is well known that the dsDNA molecule undergoes a phase transition from B-DNA into an overstretched state at high forces. For long, the structure of the overstretched state remained unknown and highly debated, but recent advances in experimental techniques have presented evidence of more than one possible phase (or even a mixed phase) depending on ionic conditions, temperature and base-pair sequence. In Chapter 6, we present a theoretical model to study

the overstretching transition with the possibility that the overstretched state is mixture of two phases: a structure with portions of inner strand separation (melted M-DNA) and an extended phase that retains the base-pair structure (S-DNA). We model the dsDNA as a chain composed of n segments of length l , where the transition is studied by means of a Landau quartic potential with statistical fluctuations. The length l is a measure of cooperativity of the transition and is key to characterizing the overstretched phase. By analyzing the different values of l corresponding to a wide spectrum of experiments, we find that for a range of temperatures and ionic conditions, the overstretched form is likely to be a mix of M-DNA and S-DNA. For a transition close to a pure S-DNA state, where the change in extension is close to 1.7 times the original B-DNA length, we find $l \approx 25$ base-pairs regardless of temperature and ionic concentration. Our model is fully analytical, yet it accurately reproduces the force-extension curves, as well as the transient kinetic behavior, seen in DNA overstretching experiments.

Chapter 2

The dependence of DNA supercoiling on solution electrostatics

The mechanical and electrostatic properties of DNA affect in a direct way various cellular processes, such as replication, transcription, compaction and protein-DNA binding. This is our motivation behind the study of DNA supercoils which are also known as plectonemes. Plectonemes in DNA molecules are manipulated by several molecular machines during key processes such as transcription and DNA repair [39]. In several scenarios, the action of these molecular machines or enzymes on DNA has been found to depend on the mechanical stress present in the molecules [61, 62]. Consequently, DNA supercoiling remains a subject of study for theorists and experimentalists alike.

Experimentally, DNA supercoiling has been investigated using several biochemical and biophysical methods including single molecule experimental techniques, where individual DNA molecules can be stretched and twisted under physiologically relevant conditions [11, 12, 27, 63, 64]. In these experiments, it is possible to apply a force and/or moment parallel to the filament axis of a DNA molecule, and to measure the elastic response in terms of elongation and angle of twisting about the filament axis. In the rotation-extension experiments the vertical extension of the DNA filament and the external moment are recorded as a function of the number of turns. It is a well-known feature of the experimental curves that there is a regime, corresponding to the formation of plectonemes, where there is almost a linear relationship between the DNA extension and the applied number of turns. Also, as shown in recent experiments of Forth et al. [12], Lipfert et al. [27] and Mosconi et al. [11] the external moment is approximately constant in the plectonemic regime.

Theoretically, plectonemes have been studied as elastic rods by many authors [65–70]. In order to interpret single molecule experiments, Purohit [71, 72] accounts for the effects of thermal fluctuations as well as electrostatics in the plectonemes and straight portions of the DNA and shows that many features seen in recent experiments of Forth et al. [12] can be qualitatively reproduced using an elastic rod model. Furthermore, as seen in figure 5 in Purohit [72], his theoretical results for the slope of the linear region in vertical extension of the DNA *vs.* number of turns of the bead is around double the value of those found in experiments by Forth et al. [12]. One of the goals of this chapter is to address this problem and get more quantitative agreement with single molecule experiments. Our approach follows van der Heijden et al. [70] and Clauvelin et al. [45, 73], who use a variational formulation to solve for the geometry of the plectoneme. The analysis in van der Heijden et al. [70] considers only the elastic energy of the filament, but Clauvelin et al. [45, 73] and other authors [4] consider electrostatic interactions together with the elasticity and are able to reproduce some features of the rotation-extension experiments. In agreement with Purohit [72], Clauvelin et al.

[73] reach the conclusion that electrostatics plays a minor role compared to the elasticity of the DNA in these experiments. Contrary to this conclusion, recent single molecule experiments and molecular simulations have shown that the results of the rotation-extension experiments depend strongly on the salt concentration of the solution [5, 10]. For this reason we carefully consider electrostatics in this chapter and present an analytical model that captures the behavior of DNA in rotation-extension experiments and simulations for a variety of DNA lengths, applied loads and salt concentrations. We also apply our model to a novel set of DNA experiments with a mixture of monovalent and multivalent salts and show that we can predict the results of these experiments.

Other key variables that are affected by the salt concentration are the discontinuities in extension and torque during the supercoiling transition [10]. These discontinuities have been studied recently by Forth et al. [12] and Daniels et al. [74]. Purohit's models [71, 72] capture these discontinuities or jumps qualitatively, but he does not comment on the salt dependence of the jumps. We use our model to provide estimates for the number of turns at which the DNA makes a transition from straight to a supercoiled configuration and for the jump in the extension and moment as a function of DNA length and salt concentration. Furthermore, we contemplate the possibility of the formation of multiple plectonemes and other forms of DNA compaction (loops and plectonemes coexistence) due to energetic reasons.

2.1 General Description of The Model

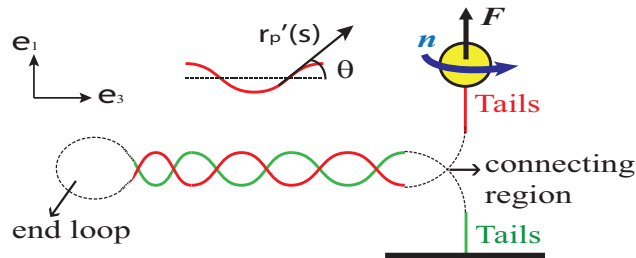


Figure 2.1: Sketch representing single molecule experiments, where a DNA molecule is fixed at one end, while the other end is subjected to a pulling force F and twisted by a given number of turns n .

We proceed with a model of the plectonemic region of the DNA molecule based on the framework of Clauvelin et al. [73], but we account for thermal fluctuation effects, confinement entropy and an end loop model. The DNA in the experiments is modeled as a Kirchhoff-inextensible-elastic rod of length $2l$ (with $-l \leq s \leq l$, where s is the arc-length along the centerline of the rod). The Kirchhoff theory of rods models the center-line as a curve in space $\mathbf{r}(s)$ endowed with mechanical properties which are assumed to be suitable averages over the cross-section of the rod [75, 76]. The configuration of an inextensible, unshereable rod is defined by $\mathbf{r}(s)$ and an associated right-handed orthonormal director frame $\mathbf{d}_i(s)$, $i = 1, 2, 3$. For convenience the vector $\mathbf{d}_3 = \mathbf{r}'(s)$ is taken to be tangential to the rod. The kinematics of the frame are encapsulated in the director frame equations $\mathbf{d}_i' = \mathbf{u} \times \mathbf{d}_i$, where the components of $\mathbf{u} = u_i \mathbf{d}_i$ are measures of the strain; u_3 describes the physical twist; u_1 and u_2 are associated with bending such that the square of curvature is given by $\kappa^2 = u_1^2 + u_2^2$. We assume a linear constitutive relation between the stresses and the strains, so that the moment $\mathbf{m} = K_b u_1 \mathbf{d}_1 + K_b u_2 \mathbf{d}_2 + K_t u_3 \mathbf{d}_3$, where K_b is the bending modulus and K_t the twisting modulus. The rod is made up of three regions (see Fig. 2.1):

- In the linear regions the tails are, on average, aligned with the vertical axis. The tails are not completely straight and the center line follows a writhed path due to thermal fluctuations of the DNA molecule. The analysis of fluctuating polymers subjected to tension and twist in the

straight regime has been carried out in detail by Moroz and Nelson [40, 56], where expressions for the twist and writhe have been provided. In our model we will use their expressions.

- In the plectonemic region the position vector $\mathbf{r}_p(s)$ and the tangent vector $\mathbf{r}'_p(s)$ describe the superhelix. Note that each helix is itself a piece of double-stranded DNA molecule. So in the literature, DNA plectonemic geometrical variables (angle and radius) are often referred to as supercoiling or superhelical, to distinguish them from the intrinsic helical nature of the base pair structure. Due to the symmetry of the problem, it is convenient to introduce cylindrical coordinates (r, ψ, z) for the position vector:

$$\mathbf{r}_p(s) = \chi r \mathbf{e}_r + z \mathbf{e}_3, \quad (2.1)$$

where \mathbf{e}_3 is the axis of the helix that wraps around the cylinder and $\mathbf{e}_r = \cos \psi \mathbf{e}_1 + \sin \psi \mathbf{e}_2$. The tangent to the position vector is:

$$\mathbf{r}'_p(s) = \sin \theta \mathbf{e}_\psi + \cos \theta \mathbf{e}_3, \quad (2.2)$$

$$\psi' = \chi \frac{\sin \theta}{r}, \quad z' = \cos \theta, \quad 0 < \theta < \frac{\pi}{2},$$

where the chirality $\chi = \pm 1$ stands for the handedness of the helix; $\chi = 1$ for a right-handed helix and $\chi = -1$ for left handed helix [45]. The other filament of the plectoneme is obtained by a rotation of π about the helical axis \mathbf{e}_3 . The plectonemic region is characterized by the helical radius r and the helical angle θ , which are assumed to be independent of the arclength s . The complement $\pi/2 - \theta$ of the helical angle is often referred to as the pitch angle. Both r and θ may depend on the loading. Geometric impenetrability of the helices implies that $\theta \leq \pi/4$ [19, 77]. Note that the external moment M_{ext} applied in the upper tail of the DNA molecule is equivalent to a total moment M_3 about \mathbf{r}'_p at the beginning of the plectonemic region. By the arguments of conservation of torque about the body axis of an isotropic rod, $\mathbf{m} \cdot \mathbf{d}_3 = K_t u_3 = M_3$ is a constant [76], implying that the twist u_3 is constant in the helical region¹. One consequence of the use of the expressions given by Moroz and Nelson [56] is that the twist u_3 in the tails is different from that in the plectoneme even though the twisting moment $M_{ext} = M_3$ is the same, since the effective twist modulus is different in each region.

- In the end of the plectonemic region there is a loop. The end loop is formed in a transition from the straight configuration into the plectonemic configuration. In order to model the loop we propose an approximation based on the localizing solution of an elastic rod [78, 79] neglecting thermal fluctuations [80]. For details we refer the reader to section A.1 in Appendix A.

The molecule contour length spent per tail is denoted by l_t ($L_t = 2l_t$), the contour length in the loop is denoted by L_o and the contour length per helix is denoted by l_p ($L_p = 2l_p$). The sum of the length of all regions is given by $L = L_p + L_t + L_o$. The equilibrium configuration of the rod is fully specified by the center-line, through the variables r , θ and M_3 . In what follows we compute these parameters as a function of the loading (pulling force, F , and the number of turns, n) by minimizing the free energy of the system.

The experiments are performed under imposed end rotations; therefore, the energy minimization will be performed under the constraint that the number of turns n imposed on the bead at one end

¹At the transition point going from an initially straight state to a plectonemic state there is a jump in the external torque. We define $M_{ext} = M_{critical}$ as the twisting moment in the straight configuration right before the transition (no plectonemes formed), while $M_{ext} = M_3$ is defined as the twisting moment when plectonemes (helices) are present, and $\delta M = M_{critical} - M_3$ as the jump in the twisting moment at the transition (see section 3.1). We use the notation M_{ext} in section 2 for the external torque. When plectonemes are present, the equations describing the DNA tails can be used by replacing M_{ext} with M_3 . When there are no plectonemes in the straight state right before the transition, the equations describing the DNA tails can be used to describe the entire molecule by replacing M_{ext} with $M_{critical}$.

of the DNA is equal to the excess link Lk_p of the DNA molecule in the helical region, the excess link Lk_t in the tails and the excess link Lk_o in the loop:

$$n = Lk_p + Lk_t + Lk_o, \quad (2.3)$$

where the link Lk_p in the helical region corresponds to the classical partition into twist Tw and writhe Wr [53]:

$$Lk_p = \left[\frac{M_{ext}}{2\pi K_t} - \chi \frac{\sin 2\theta}{4\pi r} \right] L_p. \quad (2.4)$$

At this point we note that clockwise rotations n about the \mathbf{e}_1 axis, corresponding to a positive external moment M_{ext} generate a left-handed helix with $\chi = -1$, while a negative external moment generates a right-handed helix with $\chi = 1$. We also note that in the presence of thermal fluctuations there is a writhe contribution from the tails which can be accounted for by using the results of Moroz and Nelson [56]:

$$Lk_t = \frac{M_{ext}(L - L_p - L_o)}{2\pi} \left(\frac{1}{K_t} + \frac{1}{4K_b K} \right) + O(K^{-3}), \quad (2.5)$$

where,

$$K = \frac{\sqrt{K_b F - M_{ext}^2/4}}{k_B T}, \quad (2.6)$$

k_B is the Boltzmann constant and T is the absolute temperature. The link in the end loop can be approximated as (see section A.1 in the Appendix):

$$Lk_o = \frac{M_{ext} L_o}{2\pi K_t} + Wr_o, \quad (2.7)$$

where $Wr_o \approx 1$ is the writhe present in the loop.

2.1.1 Potential Energy of the System

It is convenient to express the total potential energy of the DNA filament as:

$$V = \int_0^L \Gamma(s, \theta, q_i) ds = V_{tails} + V_{loop} + V_{helices}, \quad (2.8)$$

where q_i are variables like M_3, r, \dots independent of s . The free energy in the case of fixed-force F and fixed-torque M_{ext} in the tails (straight portion) is given by [56]:

$$E_t = \left(-F - \frac{M_{ext}^2}{2K_t} + G_{flu}^* \right) L_t, \quad (2.9)$$

where the last term is a correction due to thermal fluctuations:

$$G_{flu}^* = \frac{(k_B T)^2}{K_b} K \left(1 - \frac{1}{4K} - \frac{1}{64K^2} \right) + O(K^{-3}), \quad (2.10)$$

where K is given by Eq. (2.6). The extension with thermal fluctuations taken into account is given by $\partial E_t / \partial F = \rho L_t$ ², where

$$\rho = 1 - \frac{1}{2} \frac{1}{\sqrt{\frac{K_b F}{k_B^2 T^2} - \frac{M_{ext}^2}{4k_B^2 T^2} - \frac{1}{32}}} + \frac{K_b k_B T}{L_t (K_b F - \frac{M_{ext}^2}{4})}. \quad (2.11)$$

²The given formula for ρ includes corrections as detailed in Moroz and Nelson [40].

The slope of the rotation-extension curve after the formation of plectonemes can be obtained from constraint Eq. (2.3) together with Eq. (2.11). The extension of the filament is given as $\Delta z = \rho(L - L_o - L_p)$. Noting that L is constant, L_o is approximately constant and ρ is independent of n , altogether the slope of the rotation-extension curve is given by:

$$\frac{d}{dn}(\Delta z) = -\rho \frac{d}{dn}(L_p). \quad (2.12)$$

In what follows we drop the negative sign and simply refer to the slope as $\rho(dL_p/dn)$. The end rotation conjugate to M_{ext} is given by $-\frac{\partial E_t}{\partial M_{ext}} = 2\pi Lk_t$. To get the free energy of the system under imposed end rotations Lk_t , we apply a Legendre transform:

$$V_{tails} = E_t + 2\pi Lk_t M_{ext}. \quad (2.13)$$

The free energy of the loop will be approximated under the assumption that the bending and twisting energy decouple (see Appendix section A). The twisting moment M_{ext} is a constant along the molecule and therefore the twist energy per unit length is a constant too. The expression for the bending energy per unit length E_{o-bend} and the length of the loop L_o are obtained from the expressions given by Kúlic et al. [79] in the absence of twist.

$$V_{loop} = \left(\frac{M_{ext}^2}{2K_t} + E_{o-bend} \right) L_o, \quad (2.14)$$

where $L_o = 4\sqrt{K_b/F}$ and $E_{o-bend} = F$.

The free energy of the plectonemic region can be divided into elastic energy $V_{el}^{helices}$ and the energy due to internal interaction $V_{int}^{helices}$. The elastic energy is given by :

$$V_{el}^{helices} = \left(\frac{K_b}{2} \kappa^2 + \frac{M_{ext}^2}{2K_t} \right) L_p \quad (2.15)$$

where $\kappa = \sin^2 \theta / r$ is the curvature of a uniform helix [76]. Eq. (2.15) captures the elastic behavior of the rod in response to the applied loadings; it is zero in the straight and twist less (natural) configuration of the rod. The electrostatic and entropic interactions present in the plectonemic region $V_{int}^{helices} = U(r, \theta, x_i) L_p$ will be described in more detail later; here x_i represents any auxiliary parameters or internal variables that may appear in the free energy of the system depending on the model picked to describe the electrostatic and entropic parts of the energy. The potential energy can be written by separating the terms that contribute along L , and the ones that only contribute along L_p and L_o . We introduce a Lagrange multiplier λ to account for the constraint Eq. (2.3) and define:

$$\begin{aligned} I(\theta, r, M_3) = & \frac{K_b \sin^4 \theta}{2 r^2} + F + U(r, \theta, x_i) - G_{flu}^* - \frac{M_{ext}^2}{4K_b K} \\ & + \frac{\lambda}{2\pi} \left(\frac{M_{ext}}{4K_b K} + \chi \frac{\sin 2\theta}{2r} \right), \end{aligned} \quad (2.16)$$

such that the final expression for the potential energy of the system subject to the constraint Eq. (2.3)

is given by:

$$\begin{aligned}
V = & I(\theta, r, M_3)L_p + \left(\frac{M_{ext}^2}{2K_t} - F + G_{flu}^* + \frac{M_{ext}^2}{4K_bK} \right) L \\
& + \left(E_{o-bend} + F - G_{flu}^* - \frac{M_{ext}^2}{4K_bK} \right) L_o \\
& + \lambda \left[n - \frac{M_{ext}L}{2\pi} \left(\frac{1}{K_t} + \frac{1}{4K_bK} \right) + \frac{M_{ext}L_o}{8\pi K_bK} - W r_o \right].
\end{aligned} \tag{2.17}$$

2.1.2 Internal Energy: Entropy and Electrostatics

In the previous section we introduced the term $U(r, \theta, x_i)$ as a general expression to account for the internal interactions and configurational entropy cost in the plectonemic region. The term $U(r, \theta, x_i)$ represents the undulation-enhanced free energy per unit length plus electrostatic energy [20]. We split the internal energy of the plectoneme $U(r, \theta, x_i)$ into the configurational entropy cost contribution $U_{conf}(r, \theta, x_i)$ and the purely electrostatic interactions between the charged helices in ionic solution $U_{el}(r, \theta, x_i)$, such that $U = U_{conf} + U_{el}$.

2.1.3 Electrostatics

At moderate length scales electrostatic interactions between phosphate groups in two different molecules and between phosphate groups and counterions (positively charged) and coions (negatively charged) are present in the solution. Theoretical analysis of electrostatic interactions between polyions in solution have been done by Kornyshev et al. [81] and Parsegian and coworkers [82, 83]. Up to date, DNA-DNA interactions are still not clearly understood. We studied the effects of some variants of the internal energy models available in the literature that have been used to model DNA single molecule experiments in section A.2 in the Appendix. One of the electrostatic models shown in section A.2 is the Ubbink and Odjik [20] model derived for supercoiled DNA. This analytical model is based on the leading asymptotic contribution of the Debye-Huckel potential around two charged line segments (helices). The other model shown in the supplementary data section A.2 is the Marko and Siggia [52] electrostatic model. This model corresponds to a uniform approximation of the superposition of the potential in two limiting forms – the electrostatic potential independent of θ and the electrostatic potential independent of r . As noted by Ubbink and Odjik [20], this is less accurate. More importantly, since in our problem we are minimizing the free energy of the system, we are interested in the derivatives of the electrostatic potential. But, the Marko and Siggia superposition model underestimates the value of the derivative [20]. Finally, the Ubbink and Odjik model includes undulation enhancement effects due to thermal fluctuations. For these reasons and the results summarized in the supporting section A.2, we have used the expression given by Ubbink and Odjik [20]:

$$\begin{aligned}
U_{el} = U_{PB}(r, \theta, d_r) = & \frac{1}{2} k_B T \nu^2 l_B g(\theta) \sqrt{\frac{\lambda_D \pi}{r}} e^{2\frac{d_r^2}{\lambda_D^2} - \frac{2r}{\lambda_D}}, \\
g(\theta) = & 1 + 0.83 \tan^2(\theta) + 0.86 \tan^4(\theta),
\end{aligned} \tag{2.18}$$

where d_r represents the small undulations of the helix in the radial direction and leads to a correction in the electrostatic interaction energy due to the thermal fluctuations. The Bjerrum length l_B (nm) is defined as the length scale at which thermal energy is equal to Coulombic Energy and it is approximately 0.7nm in water at 300 K[54]. The Debye length λ_D (nm) and the effective linear charge ν (nm⁻¹) depend on the monovalent salt concentration. It is important to note that no consensus has been reached on the exact value of ν [5, 73]. The Debye screening length in water

can be obtained from $\lambda_D = 0.305[nm]/\sqrt{c_o[M]}$, where $c_o[M]$ is the monovalent salt concentration in molar units [54].

2.1.4 Configurational Entropy

The fluctuation free energy is $k_B T$ per correlation region [52] and the free energy of entropic confinement per unit length of the strand in the plectonemic supercoil may be written approximately as a superposition of two fluctuating modes due to radial (d_r) and longitudinal ($p\pi$) displacements [20, 84]:

$$U_{conf}(d_r, \theta) = \frac{k_B T}{A^{1/3}} \left[\frac{c_p}{(p\pi)^{2/3}} + \frac{c_r}{d_r^{2/3}} \right], \quad (2.19)$$

where $A = K_b/(k_B T)$ is the persistence length of the fluctuating rod. The term $2\pi p$ is the pitch of the helix and is given by $p = r \cot \theta$. The terms c_r and c_p are in general unknown constants. For a worm-like chain confined in a harmonic potential $c_r = c_p = 3(2^{-8/3})$ in one dimension [20], but as noted by van der Maarel [84] it is not clear that these values can be adopted for the supercoiled configuration. In our calculations in section 3 we will use $c_r = c_p = 2^{-8/3}$ which are empirically optimized constants [84].

2.1.5 Variational Formulation

Once the DNA has transitioned from the straight configuration into the plectonemic state the external moment M_{ext} plateaus. Recall that we define $M_{ext} = M_3$ as the external moment present in the molecule in the plectonemic state. To minimize the energy we need to equate the following partial derivatives to zero [4, 20, 45, 72, 73]:

$$\left\{ \frac{\partial V}{\partial M_3}, \frac{\partial V}{\partial r}, \frac{\partial V}{\partial \theta}, \frac{\partial V}{\partial L_p}, \frac{\partial V}{\partial d_r} \right\} = 0,$$

which yields:

$$\lambda = 2\pi M_3 + O(K^{-3}), \quad (2.20)$$

$$\left(\frac{K_b \sin^4 \theta}{r^3} - \frac{\partial U(r, \theta)}{\partial r} + \chi M_3 \frac{\sin 2\theta}{2r^2} \right) L_p = 0, \quad (2.21)$$

$$\left(K_b \frac{2 \sin^3 \theta \cos \theta}{r^2} + \frac{\partial U(r, \theta)}{\partial \theta} + \chi M_3 \frac{\cos 2\theta}{r} \right) L_p = 0, \quad (2.22)$$

$$I(\theta, r, M_3) = 0, \quad (2.23)$$

$$\frac{\partial U(\theta, r, d_r)}{\partial d_r} = 0, \quad (2.24)$$

where $I(\theta, r, M_3)$ is given by Eq. (2.16). Note that we minimize with respect to the external moment M_3 which is constant along the DNA molecule, instead of minimizing with respect to the twist u_3 [45, 73] which is different in the tails and helices depending on the magnitude of thermal motion. Because of the manner in which we treat fluctuations in the energy expressions our results for the equilibrium supercoiling variables θ, r, M_3 do not depend on the value of K_t , unlike the case in the full solution in Neukirch and Marko [4]. We are interested in the non-trivial solution $L_p \neq 0$ which corresponds to the minimum energy configuration when $n > 0$. The results obtained in this section for the plectonemic state of the DNA molecule could be obtained also under the assumption that $L_o \ll l$, when the loop size is neglected in comparison to the length of the tails and plectoneme.

2.2 Comparison with experiments and predictions

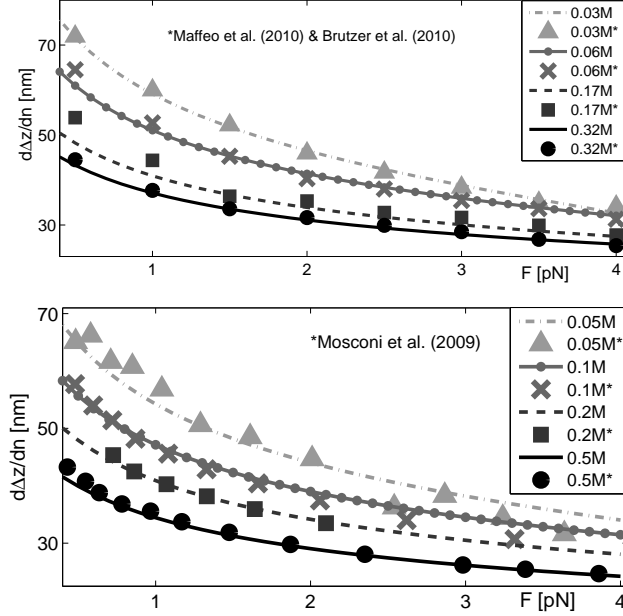


Figure 2.2: Slope $d\Delta z/dn$ as a function of the external force F . Lines represent our predictions and markers the experimental data. The upper graph shows the experimental data in Brutzer et al. [10] and Maffeo et al. [5]. The values of the linear effective charge ν used are 2.83, 3.80, 6.16 and 10.00 [nm^{-1}] for 30mM, 60mM, 170mM and 320mM salt concentration respectively. The lower graph shows the experimental data in Mosconi et al. [11] and the values of the linear effective charge ν used are 3.73, 5.32, 7.71 and 14.31 [nm^{-1}] for 50mM, 100mM, 200mM and 500mM salt concentration respectively.

We begin with a short review of the experiments. In Brutzer et al. [10], Forth et al. [12] and Mosconi et al. [11] the response of single DNA molecules to externally applied forces and torques was directly measured using an angular optical trap or magnetic tweezers. The end-to-end extension of the DNA molecule was monitored as a function of the number of turns n applied at the unconstrained end. Maffeo et al. [5] completed the data sets of the slopes in the experiments in Brutzer et al. [10] for: 30mM, 60mM, 170mM and 320mM monovalent salt. Forth et al. [12] reported direct measurements of the external torque M_3 using optical traps, while Lipfert et al. [27] used a novel method to directly measure the torque in single molecule experiments using magnetic tweezers. Both Brutzer et al. [10] and Mosconi et al. [11] have provided indirect measurements of the external torques M_3 . Besides the experimental results, Maffeo et al. [5] also provided the external torque, radius and slopes of the rotation-extension curves from Monte Carlo simulations. The experimental data of the slopes from Brutzer et al.[10] matched quantitatively with the Monte Carlo results. Thus, Maffeo et al. [5] concluded that within a cylinder approximation, DNA-DNA interactions can be described only by a significantly reduced DNA charge. Maffeo et al. [5] derived a simple model which neglected the entropy due to thermal fluctuations in the DNA molecule and accounted for the electrostatic interactions using the Debye-Huckel equation for a point charge over two line segments (helices), where the effective linear charge ν is fitted to be 0.42 times the bare DNA charge (see Supplemental Material in Maffeo et al. [5]).

The data sets from Brutzer et al.[10], Maffeo et al.[5] and Mosconi et al.[11] provide consistent slope values over the entire force range. The slopes from Forth et al.[12] are consistent with the rest of the data sets for moderate forces, but differ in magnitude at low forces as shown in Fig.A.2 and Fig.A.13 in the Appendix A. In general, the experimental results confirmed that at higher

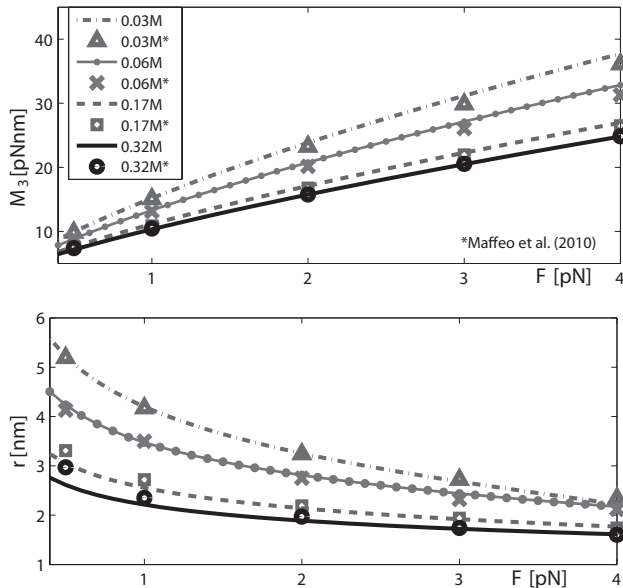


Figure 2.3: External moment M_3 and superhelical radius r theoretical predictions for the different salt concentrations in Brutzer et al. [10]. The lines are our predictions and the markers are the data points corresponding to the values of r and M_3 in the Monte Carlo simulations of Maffeo et al. [5]. The values of ν used are shown in Table 2.1.

salt concentrations, the slopes of the rotation-extension curves and the torques M_3 are both lower. But, the various data sets show greater disagreement in the torque values (see Fig.A.14 in the supplementary data A). The torque data sets from Forth et al.[12], Lipfert et al.[27] and Brutzer et al.[10] provide rather high M_3 values that do not agree quantitatively with the Monte Carlo simulations as functions of the salt concentration c_o . For instance, the indirectly measured torque M_3 for the 320mM series reported by Brutzer et al. [10] differs by about 20% from the ones obtained in the Monte Carlo simulations carried out by Maffeo et al. [5]. The torque measurements from Mosconi et al.[11] are the lowest and can be made consistent with the Monte Carlo simulations with a 1.5 pNnm offset. The indirect torque measurements of Mosconi et al.[11] and the force dependence of the slopes satisfy the 'Maxwell' type relation derived by Zhang and Marko [85], as do the simulations and our theory (see supporting section A.4). The values of M_3 reported by Forth et al. [12] at 150mM are larger than the 50mM series reported by Mosconi et al. [11] by more than 20% at low forces. This contradicts the general trend that at larger ionic concentrations the external torque should be lower [5, 10] and provides an opportunity to determine what trends are predicted by theory.

In this section we compare our theoretical predictions to the different sets of data mentioned above. Since the electrostatics in DNA-DNA interactions is not completely understood and the different models in the literature have not reached consensus on the value of the effective linear charge ν , we will let it be a fitting parameter. Based, on the work of Stigter [1–3], values ranging from 0.42 to 1 of the bare DNA charge (of a uniformly charged rod with radius $a \in [1, 1.2]$ nm) are found in the literature [4, 5, 52]. Besides Stigter, Ubbink & Odjik[20] and Vologodskii & Cozzarelli [86] have also provided ν values for a charged cylinder with $a = 1.2$ [nm]. Our values of ν for each salt concentration lie within the range of values used by other authors. The effective linear charge ν used in our calculations (for each salt concentration) is presented in Table 2.1.

We show the results of our model including the effect of undulations along the radial direction in the internal energy $U(r, \theta, d_r) = U_{PB}(r, \theta, d_r) + U_{conf}(\theta, d_r)$. The values of M_3 , r and θ as functions of the external force F are obtained by solving the system of equations given by Eq. (2.20)-Eq. (2.24).

The slope can be obtained combining expressions Eq. (2.3) and Eq. (2.12):

$$\frac{d\Delta z}{dn} = \rho \left[\frac{\sin 2\theta}{4\pi r} - \frac{M_3}{8\pi K_b K} \right]^{-1}. \quad (2.25)$$

We have obtained solutions for F in a range of 0.4 pN to 3.5 pN, for which the Moroz and Nelson [56] formulae apply. In Fig. 2.2 and Fig. 2.3 we show the results of the present model under the conditions of the experiments in Brutzer et al. [10] and Maffeo et al. [5], for a DNA template of 1.9kbp. We use in our calculations the bending modulus $K_b = 50k_B T \text{nm}$ as used by Brutzer et al. [10]. As seen in Fig. 2.2 and Fig. 2.3 the quantitative predictions of our model for M_3 , $d\Delta z/dn$ and r match consistently with the Monte Carlo simulations and experimental data in Maffeo et al. [5]. For low salt concentrations and high forces the predictions of our theoretical model overestimate the external moments by only about 1pNnm. It is reassuring that our theoretical model matches almost exactly the three variables M_3 , $d\Delta z/dn$ and r with only one fitting parameter ν . We also found that the ratio of the undulations d_r to the superhelical radius r is about 30% which is consistent with the ratios reported by Ubbink and Odijk [20]. Similarly, in Fig. 2.2 we compare our theoretical predictions for the slope of the rotation-extension curves to some of the experimental data series reported by Mosconi et al. [11] for a DNA template of 15.9kbp. As before, we pick ν to be a fitting parameter and use $K_b = 50k_B T \text{nm}$. Fig. 2.2 shows excellent quantitative agreement between the direct measurement by Mosconi et al. [11] and our theoretical predictions. Our fitting values of ν (see Table 2.1) are consistent with each other and follow the expected trend by increasing as the salt concentration increases. Using the present internal energy model with the configurational entropy coefficients $c_r = c_p = 2^{-8/3}$ produces theoretical predictions for M_3 that follow the qualitative trend of the indirect measurements by Mosconi et al. [11], but differ quantitatively by about 2.5pN (a possible reason for this discrepancy is given in the supporting section A.4).

c_o [mM]	ν [nm^{-1}]	$\xi = \nu/\nu_{bare}$ ($a = 1$ nm)	$\hat{\xi} = \nu/\nu_{bare}$ ($a = 1.2$ nm)
30	2.83	0.70	0.61
50	3.73	0.75	0.63
60	3.80	0.71	0.59
100	5.32	0.75	0.62
150	5.93	0.67	0.53
170	6.16	0.62	0.50
200	7.71	0.71	0.54
320	10.00	0.64	0.46
500	14.31	0.60	0.42

Table 2.1: Effective linear charge ν used in our calculations as a function of the monovalent salt concentration c_o [mM]. The third column shows the fraction $\xi = \nu/\nu_{bare}$, where ν_{bare} has been computed as in Refs. [1–4] for $a = 1\text{nm}$. For salt concentrations in the range [30-500]mM, the value of the charge ν_{bare} can be approximated by a linear fit with $R^2 > 0.99$ (linear fit predicts 99% of the variance on the fitted variable). Based on this idea we performed a linear fit to the value of ν we have used and obtained $\nu^{fit} = 2.46 + 2.38 \times 10^{-2} c_o$ with $R^2 > 0.99$ and c_o in mM units. A linear fit to ξ gives $\xi^{fit} = 0.73 - 2.7 \times 10^{-4} c_o$. The fourth column shows the fraction $\hat{\xi} = \nu/\nu_{bare}$, where ν_{bare} has been computed using $a = 1.2\text{nm}$ as in [5]. Note that for large salt concentrations ($c_o \sim 0.32 - 0.5[M]$) $\hat{\xi} \approx 0.42$ is equal to the charge adaptation factor used in Maffeo et al. simulations [5].

2.2.1 The transition point and jump estimates

It is known that at the transition from the straight to the plectonemic state there is a jump in the value of M_3 and the vertical extension of the DNA molecule [10, 12, 74]. The jump in the vertical extension means that a section of the initially straight DNA becomes a writhed supercoiled structure right after the transition. We define δn as the amount of twist from the straight configuration

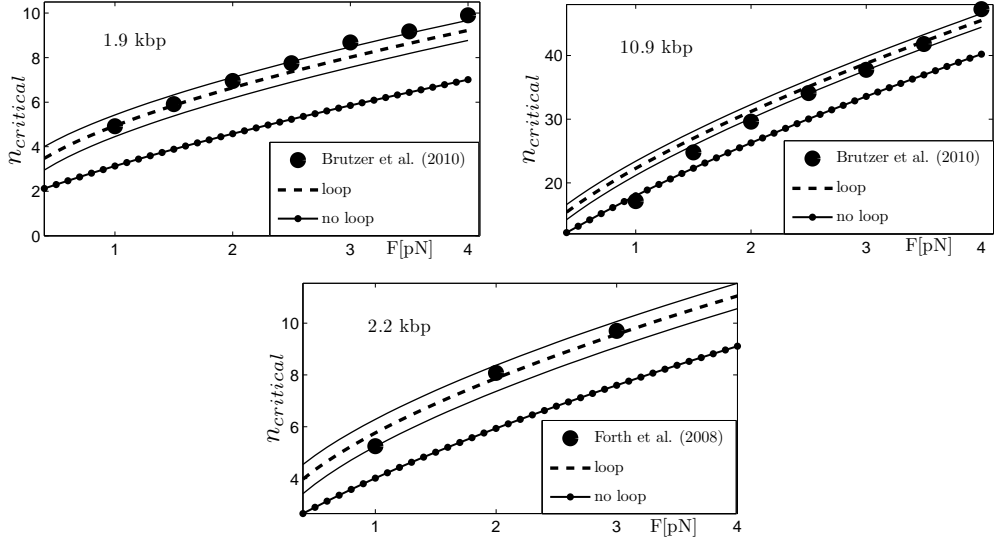


Figure 2.4: Critical number of turns $n_{critical}$ as function of the external force. Lines represent our predictions and markers the experimental data. Experimental results taken from Brutzer et al. [10] at $c_o = 320\text{mM}$ and Forth et al. [12] at $c_o = 150\text{mM}$. For $c_o = 150\text{mM}$ we have used $\nu = 5.93\text{nm}^{-1}$. Including the contribution of the end loop to compute $n_{critical}$ gives consistent results with experiment. The thin continuous lines show the solution for $n_{critical} \pm \Delta n$ (including the end loop), where Δn is an estimate of the fluctuations in n during the transition.

(right before the transition) which is transferred into writhe in the supercoiled configuration (after transition). Brutzer et al. [10] suggest by using a simple model to fit their data that in the transition the amount of twist δn which is transferred into writhe is larger than $Wr_o \approx 1$ by a significant amount (at $F = 3\text{pN}$ and $c_o = 0.32\text{M}$ $\delta n = 1.6 \pm 0.1$ turns for the 1.9kbp DNA template and $\delta n = 3.4 \pm 0.2$ turns for the 10.9kbp DNA)[10]. This conclusion would suggest that the jump in their data would correspond to the formation of an initial loop and some helical turns. Hence, the jump in the end to end distance is not just the size of the end loop [10, 74]. Strick et al.[44] show a measurement of the critical torque at the transition point based on the minimization of energy in an initial loop model. This calculation of $M_{Strick} = (2K_b F)^{1/2}$ is approximate since it neglects the thermal fluctuations in the loop and assumes that the loop is circular. As noted by Marko [55], the value of M_{Strick} overestimates by 25% the plectonemic torque data extracted from their MC simulations. Here we propose a different approach. The jump in the external moment is denoted by the difference $\delta M = M_{critical} - M_3$. We can estimate the critical number of turns $n_{critical}$ for which the transition occurs, the size of the jump in the end to end distance δz , and the jump δM by noting (a) that at the transition the energy of the straight configuration and plectonemic configuration are equal, and (b) that the linking number $n = Lk$ is a topological invariant that must be continuous at the transition between the two configurations. The energy of the straight configuration just before the transition is given by Eq. (2.13) replacing $M_{ext} = M_{critical}$ and $L_t = L$:

$$\hat{V}_s = \left(-F + \frac{M_{critical}^2}{2K_t} + G_{flu-s}^* + \frac{M_{critical}^2}{4K_s K_b} \right) L, \quad (2.26)$$

where K_s and G_{flu-s}^* are given by Eq. (2.6) and Eq. (2.10) evaluated at $M_{ext} = M_{critical}$. The energy of the plectonemic configuration just after the transition is given by Eq. (2.17) replacing

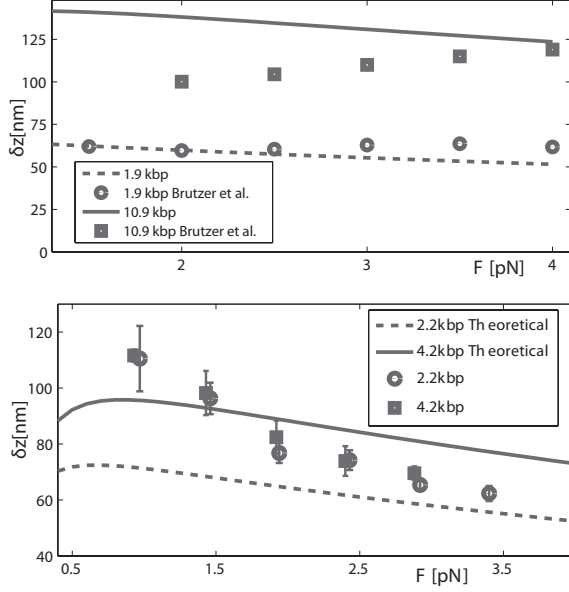


Figure 2.5: Comparison of the extension jump δz from our theory with the experiments in Brutzer et al. [10] for 1.9 kbp and 10.9 kbp DNA templates at $c_o = 320$ mM and the data in Forth et al. [12] for 2.2 kbp and 4.2 kbp DNA templates at $c_o = 150$ mM. Our model predicts that as the DNA length L or c_o increase δz increases too.

$$M_{ext} = M_3:$$

$$\begin{aligned} \hat{V}_p = & \frac{M_3^2}{2K_t} L + \left(\frac{K_b \sin^4 \theta}{2r^2} + U \right) L_p^* \\ & + E_{o-bend} L_o + \left(G_{flu-p}^* + \frac{M_3^2}{4K_b K_p} - F \right) (L - L_p^* - L_o), \end{aligned} \quad (2.27)$$

where K_p and G_{flu-p}^* are given by Eq. (2.6) and Eq. (2.10) evaluated at $M_{ext} = M_3$. The length eaten by the helices during the transition due to the dynamic jump is L_p^* . Setting $\hat{V}_p = \hat{V}_s$ we get an equation with two unknowns $M_{critical}$ and L_p^* . We get a second equation by using the continuity requirement of $n = Lk$. In the straight DNA configuration the critical number of turns n_{c-s} before the transition is given by:

$$n_{c-s} = \frac{M_{critical}}{2\pi} L \left(\frac{1}{K_t} + \frac{1}{4K_b K_s} \right). \quad (2.28)$$

In the plectonemic configuration the critical number of turns n_{c-p} just after the transition is:

$$n_{c-p} = \frac{M_3 L}{2\pi K_t} + \frac{M_3 (L - L_p^* - L_o)}{8\pi K_b K_p} + Wr \quad (2.29)$$

where $Wr \approx 1 + \sin(2\theta)L_p^*/(4\pi r)$ accounts for the writhe present in the loop and the helices. Our second equation to solve for $M_{critical}$ and L_p^* is given by $n_{c-p} = n_{c-s}$. The amount of link (twist) that is converted into writhe is readily given from Eq. (2.28) and Eq. (2.29) as the writhe after the transition minus the writhe before the transition:

$$\delta n = \left[\frac{M_3 (L - L_p^* + L_o)}{8\pi K_b K_p} + Wr \right] - \left[\frac{M_{critical} L}{8\pi K_b K_s} \right] = \frac{\delta M L}{2\pi K_t}, \quad (2.30)$$

From Eq. (2.30) if $\delta n \approx Wr \approx 1$, then we can conclude that only an end loop is formed and $L_p \approx 0$. Otherwise the jump in the end to end distance would correspond to the formation of an end loop and a plectonemic region of length L_p^* . Finally, the jump in the vertical extension is given by:

$$\delta z = \rho_s L - \rho_p [L - (L_p^* + L_o)] = \rho_p (L_p^* + L_o) + (\rho_s - \rho_p) L, \quad (2.31)$$

where ρ_s and ρ_p correspond to Eq. (2.11) for ρ evaluated at $M_{ext} = M_{critical}$ and $M_{ext} = M_3$ respectively.

Next we show the results for the transition variables obtained using $K_t = 95k_B T$ which is an accepted value of the twisting modulus [56, 71]. Our theoretical model predicts that the size of the jump at the transition strongly depends on the length of the DNA molecule and the salt concentration c_o . We find that the jump in the external moment δM decreases with increasing DNA length and the jump in the end to end extension δz increases with increasing DNA length. We conclude that as c_o decreases δM and δz decrease too. The experimental data in Forth et al. [12], Daniels et al. [74] and Brutzer et al. [10] agrees with our conclusion. Fig. 2.4 shows the comparison between the theoretical predictions and experimental measurements of $n_{critical}$ as a function of F , where we plot the solution for $n_{critical}$ accounting for an end loop. We also plot the solution obtained by neglecting the loop by setting $Lk_o = Lk_p = 0$ in Eq. (2.3) and $L_p = L_o = 0$ in Eq. (2.5) such that the critical number of turns is given by $Lk_t = M_3 L [K_t^{-1} + (4K_b K_p)^{-1}] / (2\pi)$ similar to Clauvelin et al. [73]. The predicted $n_{critical}$ in the end loop model agrees very nicely with the data points from the experiments of Forth et al. [12] and Brutzer et al. [10], while neglecting the loop underestimates the values of $n_{critical}$. In Fig. 2.5 we show the comparison between the experimental measurements in Brutzer et al. [10] and Forth et al. [12] with our predictions for the jump in the end to end extension δz . Our qualitative predictions for the transition jump in the extension agree with experimental data, meaning that as the DNA length L or c_o increase δz increases too. As seen in Fig. 2.5 the experimental data from Brutzer et al. [10] and Forth et al. [12] show different trends as a function of the applied force F . We note that our theory predicts a relatively constant value of δz as function of F for $c_o = 320\text{mM}$ (qualitatively similar to experiment) and describes qualitatively the decrease of δz as a function of F for $c_o = 150\text{mM}$. In the supporting section A.3 we present a comparison of the predicted values of $M_{critical}$ with experimental data and also estimates of the torque jump δM .

We point out that the numerical calculation of $n_{critical}$, L_p^* and $M_{critical}$ when comparing the energy of the two states neglects the fluctuation due to thermal kicks. An estimate of the fluctuations of n can be obtained within the Einstein approach for fluctuations[87, 88]:

$$\langle \Delta n^2 \rangle = \frac{k_B T}{2\pi} \left. \frac{\partial n}{\partial M_{ext}} \right|_{T, F}. \quad (2.32)$$

By doing so, we can approximate the change in the number of turns $\Delta n \approx \sqrt{\langle \Delta n^2 \rangle}$ due to the thermal kicks. Therefore the transition for a given force F takes place over $n_{critical}^{\pm k_B T} \approx n_{critical} \pm \Delta n$. For the cases presented in Fig. 2.4, $\Delta n \sim 0.5$ to 1 turns.

2.2.2 Coexistence of loops and plectonemes

Our methods also allow us to consider scenarios where we have a series of loops forming in the DNA instead of plectonemes. When only loops and no superhelical structures are present, the applied number of turns $n = Lk$ is distributed in the form of twist among the entire molecule, writhe due to thermal fluctuation in the straight regions [56], and writhe in the loops ($Wr_o \approx 1$ per loop formed). This happens when the energetic cost of forming a loop is lower than that of forming a writhed superhelix, and leads to a different slope of the rotation-extension curve. However, for a given choice of electrostatic and entropic interactions we find there is a range of forces in which both regimes can co-exist due to thermal motion.

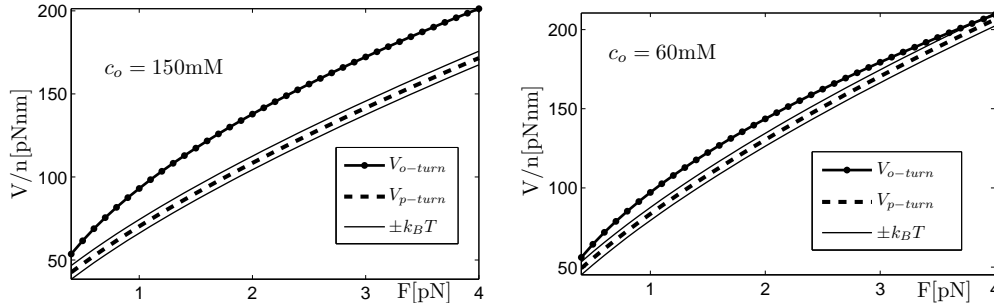


Figure 2.6: Energy per unit turn. Using our model we can get some idea of the preferred state: if $V_{o\text{-turn}} > V_{p\text{-turn}}$ plectonemes are favored, if $V_{p\text{-turn}} > V_{o\text{-turn}}$ loops are favored and if $V_{o\text{-turn}} \approx V_{p\text{-turn}}$ there is a coexistence of loops and plectonemes. We expect that for high-to-moderate salt concentrations, plectonemes will be formed, while for low-to-medium salt concentration there might be a region of coexistence or even formation of only loops.

The free energy per unit turn (excess link) in the plectonemic regime (see Eq. (2.17)) is:

$$V_{p\text{-turn}} = \frac{dL_p}{dn} \left[\frac{K_b}{2} \kappa^2 + F + U - G_{flu}^* - \frac{M_3^2}{4K_b K_p} \right] = 2\pi M_3, \quad (2.33)$$

where $dL_p/dn = \rho^{-1}(d\Delta z/dn)$ given by Eq. (2.25). Since $W\tau_o \approx 1$, the free energy per unit turn for a series of loops using the model described in the section 2.1 is approximately:

$$V_{o\text{-turn}} \approx L_o \left[2F - G_{flu}^* - \frac{M_3^2}{4K_b K} \right]. \quad (2.34)$$

Fig. 2.6 shows the regimes in which the free energy analysis would lead to the formation of either plectonemes, loops or both. For large to moderate salt concentrations $V_{p\text{-turn}} < V_{o\text{-turn}}$ for a range of external force $[0.4, 4]$ pN. As the salt concentration decreases $V_{p\text{-turn}} \approx < V_{o\text{-turn}}$ and due to thermal fluctuations there could be coexistence of both states. We have plotted the results for $c_o = 150$ mM and $c_o = 60$ mM. The lines on either side of the lower curve show the range $\pm k_B T$ at $T = 300$ K. If the upper curve is within $\pm k_B T$ of the lower curve then transitions between loops and plectonemes could happen. We recall that in our end-loop model we neglect self-contact, electrostatics and twist stored in the loop. Consequently, we expect that our analysis of the free energies per unit turn will give us only an estimate of the coexistence state of loops and plectonemes. We expect that for high-to-moderate salt concentrations, plectonemes seem to be the favorable state, while for low-to-medium salt concentration there might be a region of coexistence or even formation of only loops. Our predictions regarding the transition between the two states agree with the qualitative conclusion of Brutzer et al. [10].

2.2.3 Multivalent Ions

In this section we extend our plectonemic DNA model to make predictions for DNA single molecule experiments in the presence of multivalent ions. We consider the limiting case of high concentration of monovalent ions and a low concentration of multivalent salt as this is the case in several experimental studies on DNA aggregation [14, 89, 90] and more recently in DNA single molecule experiments (private communication with Qing Shao, Sachin Goyal, Laura Finzi and David Dunlap at Emory university). These experiments show that the addition of small quantities of multivalent salt, such as spermidine Sp^{3+} or spermine Sp^{4+} , to a solution with a high monovalent salt concentration (0.2 M KCl) can significantly modify the pitch and twist of the DNA plectoneme. When the polyions are added to the solution the experiments yield more compact plectonemes that start forming at lower

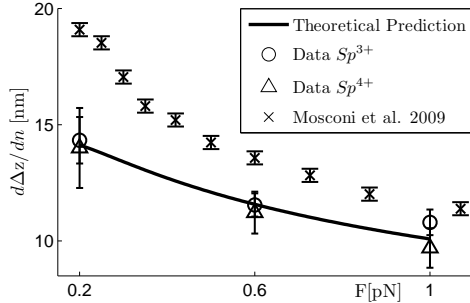


Figure 2.7: Dunlap and co-workers performed two series of experiments corresponding to the control set up ($c_o = 0.2\text{M KCl}$), one for spermidine Sp^{3+} (circles) and the other for spermine Sp^{4+} (triangles). We fit the value of ν to get the slope of the rotation-extension curve at $F = 0.6\text{pN}$ for control set corresponding to Sp^{3+} (circles) and got $\nu = 4.12\text{nm}^{-1}$. Our prediction with $\nu = 4.12\text{nm}^{-1}$ is shown by the solid line. The data from Mosconi et al. [11] shown by the cross markers corresponds to the $c_o = 0.2\text{M NaCl}$ series.

values of the supercoiling density $\sigma_{critical} \approx 3.54(n_{critical}/L)$. The Debye length accounting for the different salts is given by [14]:

$$\lambda_{D-M}[nm] = 0.435 [(\zeta^2 + \zeta) c_{mu}[M] + 2c_o[M]]^{-1/2}, \quad (2.35)$$

where $c_{mu}[M]$ and $c_o[M]$ are the multivalent and monovalent salt concentration in molar units respectively and ζ stands for the multivalent ion's valence. In the experiments of Dunlap and co-workers the control corresponds to a $\sim 3\text{kbp}$ DNA template at room temperature in a 0.2M KCl salt solution. The experiments were performed with different concentrations of Sp^{3+} or Sp^{4+} added to the control. We have used $K_b = 55k_B T$ as measured in the experiments and $K_t = 95k_B T$. To compare with the experimental data we fit the effective linear charge ν to the 0.6 pN point for each salt concentration and use it to predict the results for other values of the force. Note that the experimental slopes of the rotation-extension curves from the 0.2 M KCl series of Dunlap and co-workers and 0.2 M NaCl series in Mosconi et al. [11] do not agree quantitatively (see Fig. 2.7). In DNA molecular dynamic simulations by Savelyev and Papoian [91] qualitative differences in Na^+ and K^+ condensation patterns were observed, suggesting that ion-specific modeling is required to describe electrostatics at short distances. In our plectonemic DNA model we account for the effects of ion-specific differences by the fitted value of ν . Table 2.2 shows the effective linear charge ν used in the calculations for a set up consisting of $c_o = 0.2\text{M KCl}$ buffer with added multivalent salt (Sp^{3+} or Sp^{4+}) concentration c_{mu} .

c_{mu} [mM] (Sp^{3+})	ν [nm^{-1}]	c_{mu} [mM] (Sp^{4+})	ν [nm^{-1}]
0	4.12	0.2	2.72
1	3.35	0.5	2.02
2	2.75	0.75	1.76
5	2.20	1	1.66
10	2.10	2	1.34

Table 2.2: Effective linear charge ν for $c_o = 0.2\text{M KCl}$ buffer and added multivalent salt concentration c_{mu} . For only the monovalent salt $c_o = 0.2\text{M KCl}$ we used $\nu^o = 4.12 [\text{nm}^{-1}]$. As the multivalent salt concentration is increased the value of ν obtained from the fit decreases. This can be explained by a better screening of the DNA charge by the salt solution and the varying electrophoretic charge value in multivalent ion solutions. The Sp^{3+} values of ν are well fitted by the curve $\nu_{3+}^{fit} = \nu^o(1 + c_o/1.07)^{-1/3}$ ($R^2 > 0.97$), and the Sp^{4+} values of ν are well fitted by the curve $\nu_{4+}^{fit} = \nu^o(1 + c_o/0.07)^{-1/3}$ ($R^2 > 0.99$). The curve-fits were obtained using the least squares method with a fitting function of the form $f(x) = a(b + x)^c$.

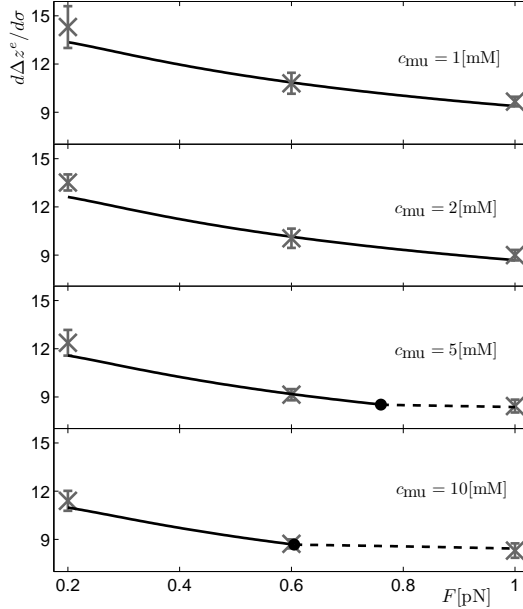


Figure 2.8: Predictions for the slope of the rotation-extension curves in the presence of multivalent ions. Crosses are data points from Dunlap and co-workers for a mixture of $c_o = 0.2\text{M}$ KCl and different c_{mu} concentrations of spermidine Sp^{3+} . For $c_{mu} = 5\text{mM}$ and $c_{mu} = 10\text{mM}$ as F increases the supercoiling diameter decreases and approaches the interaxial spacing ~ 3.0 nm for spermidine [13, 14]. The dot in the bottom two panels shows the point where $2r$ reaches the limiting interaxial distance and from there on the dashed line shows the solution where $r = 1.5\text{nm}$ is assumed to be constant.

Fig. 2.8 shows the comparison of our theoretical model and the experiment for Sp^{3+} , where we have plotted $d\Delta z^e/d\sigma$ as a function of the applied force F . $d\Delta z^e/d\sigma$ is the slope of the graphs showing the effective extension $\Delta z^e = \Delta z/L$ as a function of the degree of supercoiling $\sigma \propto n$. As the multivalent salt is increased the plectonemes are more compact. This can be explained by better screening of the DNA charge by the salt solution (smaller ν value). The reduction of the value of ν with increased polyvalent salt has also been explained by the reduction of the electrophoretic charge value. The effective linear charge is proportional to α (electrophoretic charge value) as given by Stigter and coworkers [2, 3]. For monovalent salt solutions the value of α remains constant for a large range of concentrations [3], but this is not the case in polyvalent ions [92, 93] and mixtures of multivalent ions with monovalent salts [90].

For $c_{mu} = 5\text{mM}$ and $c_{mu} = 10\text{mM}$, we found that as the force F is increased the supercoiling diameter approaches the interaxial distance $\sim 3\text{nm}$ found in hexagonally packed Sp^{3+} -DNA aggregates [13, 14]. In aggregation and condensation experiments the DNA formed closed packed hexagonal arrays where the interaxial distance corresponded to an equilibrium spacing due to competition of attractive and repulsive forces [13] which arise due to effects such as hydration, van der Waals forces, London-like dispersion forces and counter-ion fluctuations [13, 92]. We expect that as the polyvalent salt increases and the supercoiling diameter approaches 3nm these effects would become important and dominate the interactions, leading to compact DNA plectonemes with a diameter approximately equal to the interaxial spacing. In Fig. 2.8 the dot shows the point where $2r = 3\text{nm}$ and from there on the dashed line shows the solution for a constant $r = 1.5\text{nm}$. Remarkably, our predictions with r constant match the experimental point at $F = 1\text{pN}$ for larger c_{mu} concentrations.

Fig. 2.9 show the results obtained when using spermine Sp^{4+} . For Sp^{4+} we have only fitted the value of ν to the experimental point $F = 0.6\text{pN}$ for the control set up and $c_{mu} = 0.2 - 0.75\text{mM}$ concentrations and obtained a curve for ν as a function of c_{mu} . For $c_{mu} = 1\text{mM}$ and $c_{mu} = 2\text{mM}$ we have extrapolated the value of ν from the curve obtained from the previous fitted values. As

before, the dot shows the point where $2r$ is equal to the interaxial spacing $\sim 2.9\text{nm}$ for Sp^{4+} (Todd et al. [13], Raspaud et al. [14]) and from there on the dashed lines correspond to the solution with $2r \approx 2.9\text{nm}$. Our results for both types of polyvalent ions show good quantitative agreement with the experimental values. In the supporting section A.5 we show that our results are in good agreement

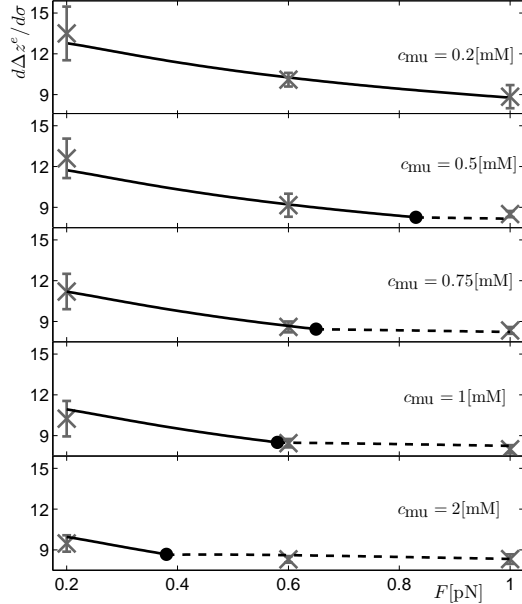


Figure 2.9: Predictions of the slope of the rotation-extension curves in the presence of multivalent ions. Crosses are data points from Dunlap and co-workers for a mixture of $c_o = 0.2\text{M}$ KCl and different c_{mu} concentrations of spermine Sp^{4+} . For $c_{mu} \geq 0.5\text{mM}$ as F increases the supercoiling diameter approaches the interaxial spacing distance ~ 2.9 nm of spermine [13, 14]. As in Fig. 2.8, the dot shows the point where $2r$ reaches the interaxial distance value and the dashed line shows the solution where $r = 1.45\text{nm}$ is assumed to be constant.

with experimental measurements of the supercoiling density $\sigma_{critical}$ at which the DNA molecule makes the transition from the straight to the supercoiled configuration. The theoretical predictions of the plectonemic moment M_3 and supercoiling radius r can also be found in the supporting section A.5.

2.3 Conclusions

We have analyzed the mechanics of plectoneme formation, where a twisted DNA molecule in the plectonemic regime has been modeled as an elastic-isotropic rod. Here we give a short summary of all the results we have obtained. We have used a variational approach to solve the energy minimization problem that corresponds to angular optical trap (or magnetic tweezers) experiments on a DNA molecule attached to a substrate at one end, while subjected to a tensile force and twisted by a specific number of turns n at the other end [10–12, 27]. Our model description is symmetric in that over-twisting or under-twisting the rod under tension gives the same results. However, this is not the case in DNA for a large number of turns n (or supercoiling density σ). Stretching and under-twisting DNA at low to moderate values of σ leads to denaturation as is known from experiment [44] and atomistic simulations [94]. Therefore, our model is valid in the over-twisting regime only for moderately large values of σ (before a structural transition into P-DNA [57]), where the Moroz and Nelson [56] formulae are valid and the DNA can be modeled with constant elastic properties along the entire molecule. We have minimized the energy with respect to the dependent variable M_{ext} rather than its conjugate n , since we are modeling rotation controlled experiments. We do

not minimize with respect to the twist u_3 , since M_{ext} is constant along the DNA molecule while u_3 is different in the tails and the helices depending on the magnitude of thermal motion. An interesting, and possibly experimentally verifiable result of minimizing with respect to M_{ext} is that the equilibrium supercoiling variables θ, r and M_3 are only functions of the bending modulus K_b , but are independent of the twisting modulus K_t .

In our one-dimensional continuum description of the DNA molecule we account for DNA elasticity, DNA-DNA interactions, fluctuations and configurational entropy in the tails and helices. As mentioned before, there is no consensus on the electrostatic models in the mechanics of DNA. So, we have used our theoretical framework to test several models of DNA-DNA electrostatic interactions and configurational entropy in the plectonemic region (see support in section A.2). Understanding the effects of each of the models and approximations ultimately lead us to pick the $U(r, \theta, d_r)$ model of Ubbink and Odijk [20] with the entropic parameters $c_p = c_r$ given by van der Maarel [84]. The electrostatic contribution to K_b is rather small for the physiological range ([0.1-0.5]M) of salt concentrations [95, 96]. Therefore, both the bending and configurational entropy energetic costs, are independent of the salt concentration. Hence, for a given a monovalent salt the plectonemic configuration as a function of F is dictated by ν . So, as noted by Maffeo et al. [5], single molecule experiments can be used to determine the appropriate effective linear charge ν for plectonemic DNA. Here we give simple analytical formulae for ν as a function of salt concentration for both monovalent and some multivalent salts (in low concentrations) that result in strong agreement of our analytical model with the different sets of experimental data and Monte Carlo simulations over a wide range of forces. Our qualitative results agree with the conclusions obtained in previous works [4, 5, 45, 71–73] and the values of ν are within the range previously obtained by others. As the salt concentration increases, the charge adaptation factor ν/ν_{bare} decreases (see Table 2.1), and the reduced effective linear charge ν approaches the values used in Maffeo et al. simulations [5].

In our model we have also accounted for the presence of the end loop. This allows us to compare the energy of the straight DNA configuration and the plectonemic DNA and leads to a method to obtain analytical estimates of the jumps in the external torque δM and end to end extension δz of the DNA molecule at the transition. Our predictions of the jump variables and the critical number of turns at which the transition occurs agree with those observed in experiment. If the energetic cost of forming a loop is lower than that of forming a helix then we will have a series of loops and this will lead to a different slope of the rotation-extension curve. But, there is a range of ionic concentrations and forces at which both regimes can co-exist due to thermal motion. We have concluded that at high-to-moderate salt concentration the most favorable state is the plectoneme, but as the salt concentration decreases the energy difference between a loop and a plectoneme also decreases. In our model we have assumed that the plectonemes can be modeled as uniform helices with constant radius and curvature. This does not have to be the case and softening the constraints in the model might lead to a better understanding of the problem. For instance, allowing the helical axis of the plectoneme to bend can lead to more complicated structures. Further, since constant curvature solutions require special boundary conditions we consider it important to analyze the more general case of variable curvature solutions [21]. Variable curvature solutions can provide theoretical insight in understanding the formation of multiple plectonemes because for two (or more) interwound helices there is a geometrical lock-up helical angle [19].

Finally, we have shown that our model for plectonemic DNA including the end loop can reproduce experimental data from single DNA molecule experiments in the presence of polyvalent ions. The theoretical estimates of the slopes and critical number of turns ($n_{critical} \propto \sigma_{critical}$) match experiments (private communication with David Dunlap) where low concentrations of multivalent salts are added to a high concentration of monovalent salt solution. In the presence of multivalent ions, it is well known that DNA forms toroidal condensates in bulk [97] and more recently toroids have been suspected to form when DNA is subjected to a tensile force [98]. A potential field of study is complex DNA condensates due to polyvalent ions in the presence of forces and torsional constraints, where there could be formation of plectonemes and toroids alike.

Acknowledgments We thank Qing Shao, Sachin Goyal, Laura Finzi and David Dunlap at Emory University for providing us their experimental data and for discussions.

Chapter 3

Competition between Supercoils and Toroids in Single Molecule DNA condensation

Several experimental studies on DNA aggregation and condensation [14, 15, 89, 99–101] have been carried out under a wide range of salt concentrations because of its physiological relevance, especially during the cell cycle [97, 102]. The counterion-induced condensation of DNA is reproduced without difficulty in experiments and yet it is still not fully understood although it is a fundamental and crucial process for our very existence [81]. Multivalent ions promote DNA bending since they neutralize the negative charges on DNA phosphate, which facilitates enhanced protein-DNA interaction, and they may play an important role in facilitating transcription [102]. Protamine-DNA assemblies closely resemble those of DNA condensed by multivalent ions that have less charge and are much smaller in size [103]. Protamine binds and condenses DNA into compact configurations in the sperm of most vertebrates; it inactivates and packages centimeters of DNA in the sperm head (size of a micron approximately) until DNA is re-activated after fertilization [81]. In the present study we extend the results obtained in chapter 2 to make predictions for DNA single molecule experiments in the presence of high concentration of condensing agents and low concentration of monovalent salt in order to gain a deeper understanding of more complex forms of DNA packing.

Most DNA condensation experiments at high multivalent ionic concentrations have been focused on free molecules in bulk solution [104–107], where toroidal structures have been observed. Besides the experiments in bulk, the force-extension curves in single molecule experiments under controlled force/extension in high multivalent ionic solutions have been measured [101, 108]. It has been demonstrated that DNA is incorporated into the collapsed condensate in discrete steps [37, 101, 109, 110] and it has been seen that DNA condenses into a single collapsed structure (torus) [37]. More sophisticated single molecule experiments have recently been performed by Besteman et al. [15] in which DNA molecules in high multivalent salt solutions were subjected to an applied tension and end rotations. They suggested a simple model to explain their experimental findings proposing the idea that the DNA compaction starts with the formation of a loop. But it is still not clear what would be the final compact DNA structure (toroids or supercoils) in the presence of high concentrations of multivalent ions under imposed end rotations. Through our model we can provide an answer to this question by mapping the different DNA states in a phase diagram constructed by computing and comparing the energies of each state as a function of the external force and number of turns.

We have used a variational approach to solve a constrained energy minimization problem that corresponds to angular optical trap (or magnetic tweezers) experiments on a DNA molecule attached to a substrate at one end, while subjected to a tensile force and twisted by a specific number of turns

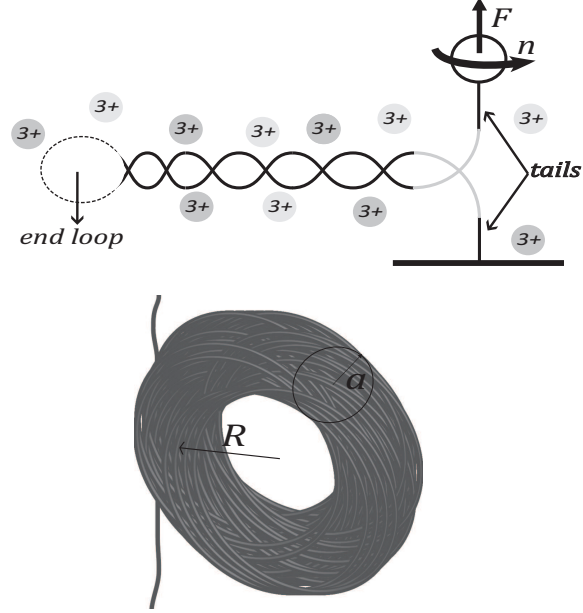


Figure 3.1: Multivalent ions with 3 or more positive charges can cause condensation. In a given experiment only one type of ion is typically used. The DNA molecule is fixed at one end, while the other end is subjected to a pulling force F and twisted by n number of turns. Top: sketch representing the formation of supercoiled structure in single molecule experiments in the presence of condensing agents. Bottom: sketch representing the formation of a toroid.

n at the other end in the presence of condensing agents (see Fig. 3.1). Our model is valid in the over-twisting regime only for moderately large values of n (before a structural transition into P-DNA [57]), where the results of Moroz and Nelson [56] for the entropic elasticity of a twisted polymer are valid and the DNA can be modeled with constant elastic properties along the entire molecule. In our one-dimensional continuum description of the DNA molecule we account for DNA elasticity, DNA-DNA interactions, fluctuations and configurational entropy in the tails, helices and toroidal loops. For the description of the toroidal structures we have added the twist effects to a well established model [111, 112]. And for the description of the supercoiled structures we follow the treatment in [21] including the effects of multivalent cations in the internal energy.

3.1 DNA supercoiling

In this section we study supercoils. We describe the straight-supercoiled configuration of the DNA, which corresponds to the co-existence of a helical region (plectonemes) and straight portions of the DNA molecule (tails). The analysis of the transition from the straight configuration into the straight-supercoiled configuration is known to be a dynamic process [10, 12, 74] and a simplified analysis based on equilibrium energetics [21] will be done in section 3.1.5. But first we briefly describe the DNA straight configuration before the supercoiling transition and the straight-supercoiled configuration after the transition.

3.1.1 Straight Configuration

As shown in [21], the energy of the straight state under imposed end rotations and tension is given by a Legendre transform of the energy under controlled torque given by Moroz and Nelson [56]:

$$V_s = \left(\frac{M^2}{2K_t} + G_{flu}^* + \frac{M^2}{4K_bK} - F \right) L, \quad (3.1)$$

where F is the external force applied, M is the external moment in the straight configuration due to the end rotations,

$$K = \frac{\sqrt{K_b F - M^2/4}}{k_B T}, \quad (3.2)$$

and the term G_{flu}^* is a correction to the energy due to thermal fluctuations [56]:

$$G_{flu}^* = \frac{(k_B T)^2}{K_b} K \left(1 - \frac{1}{4K} - \frac{1}{64K^2} \right) + O(K^{-3}). \quad (3.3)$$

where k_B is the Boltzmann constant and T is the absolute temperature. The bending modulus of the DNA filament is denoted by K_b and the twisting modulus is denoted by K_t .

The external torque M in the straight configuration can be obtained from the imposed number of turns n_{c-s} , which is equal to the linking number of the DNA molecule [56]. The link corresponds to the classical partition into twist and writhe [67, 113]. The contribution of the thermal fluctuations to the total writhe can be accounted for by using the results of Moroz and Nelson [56]:

$$n_{c-s} = \frac{M}{2\pi} \left(\frac{1}{K_t} + \frac{1}{4K_b K} \right) L + O(K^{-3}) \quad (3.4)$$

3.1.2 Straight-Supercoiled Configuration

In the straight-supercoiled state we will model DNA as consisting of three distinct regions: tails, helices and end loop. The helical region is characterized by the helical radius r and angle θ , which are assumed to be uniform. This assumption has been previously used in several theoretical studies of plectonemes [4, 5, 71, 73] giving results agreeing with experiment.

The molecule contour length spent in the tail is denoted l_t , the end loop length is denoted l_o and the contour length in the helices is denoted l_p . The sum of the length of all regions is equal to the total length of the DNA chain L . The equilibrium configuration of the rod is fully specified by the center-line, through the variables r , θ and the external moment in the presence of supercoils $M_3 \neq M$ [21]. In what follows we compute these parameters as functions of the loading (pulling force F and the number of turns n) by minimizing the total energy of the system.

The experiments are performed under imposed end rotations therefore the energy minimization will be performed under the constraint that the number of turns n imposed on the bead is equal to the excess link Lk of the DNA molecule in the tails, helices and end loop [21]:

$$n_{c-p} = Lk_p + Lk_t + Lk_o, \quad (3.5)$$

where Lk_p is the link in the helical regime:

$$Lk_p = \frac{M_3 l_p}{2\pi K_t} - \chi \frac{\sin 2\theta}{4\pi r} l_p, \quad (3.6)$$

where χ stands for the chirality of the helix [73]. Note that rotations n inducing a positive external moment M_3 generate a left-handed helix $\chi = -1$, while a right-handed helix $\chi = 1$ corresponds to negative external moments. The contribution of the tails to the total link of the straight-supercoiled

configuration is given by Eq. (3.4) replacing the total DNA length L with the tail's length $l_t = L - l_p - l_o$:

$$Lk_t = \frac{M_3 l_t}{2\pi} \left(\frac{1}{K_t} + \frac{1}{4K_b K_3} \right) + O(K^{-3}), \quad (3.7)$$

where K_3 is given by Eq. (3.2) replacing M with M_3 . The link in the end loop can be approximated as done in [21]:

$$Lk_o \approx \frac{M_3 l_o}{2\pi K_t} + 1, \quad (3.8)$$

where $l_o \approx 4\sqrt{K_b/F}$ (homoclinic loop [79]) and we have used the approximation that the writhe present in the end loop is equal to 1.

An approximation of the potential energy in the straight-supercoiled configuration can be written by separating the terms that contribute along L , l_p and l_o [21]:

$$\begin{aligned} V_{p-s} &= \left(K_b \frac{\sin^4 \theta}{2r^2} + F + U(r, \theta) - G_3^* - \frac{M_3^2}{4K_b K_3} \right) l_p \\ &+ \left(\frac{M_3^2}{2K_t} - F + G_3^* + \frac{M_3^2}{4K_b K_3} \right) L \\ &+ \left(2F - G_3^* - \frac{M_3^2}{4K_b K_3} \right) l_o \\ &+ \nu (n - Lk_p - Lk_t - Lk_o), \end{aligned} \quad (3.9)$$

where we have introduced the constraint Eq. (3.5) via the Lagrange multiplier ν . The term G_3^* is the correction to the energy due to thermal fluctuations [56] and is given by Eq. (3.3) with K_3 (evaluated with M_3 instead of M) replacing K . The term $U(r, \theta)$ in the energy Eq. (3.9) describes the internal energy interactions in the helical regime and it will depend on the type of ions that dominate the salt solution. The appropriate interaction model for multivalent ions will be described in the next sections.

Minimization of the energy Eq. (3.9) with respect to the unknown variables l_p, r, θ, M_3 yields:

$$\nu = 2\pi M_3 + O(K^{-3}), \quad (3.10)$$

$$\frac{K_b \sin^4 \theta}{r^3} - \frac{\partial U(r, \theta)}{\partial r} = -\chi M_3 \frac{\sin 2\theta}{2r^2}, \quad (3.11)$$

$$K_b \frac{2 \sin^3 \theta \cos \theta}{r^2} + \frac{\partial U(r, \theta)}{\partial \theta} = -\chi M_3 \left(\frac{\cos 2\theta}{r} \right), \quad (3.12)$$

$$K_b \frac{\sin^4 \theta}{2r^2} + U(r, \theta) + F - G_3^* = -\chi \frac{M_3 \sin 2\theta}{2r}. \quad (3.13)$$

The system of equations above gives the solution to the variables r, θ and M_3 that describe the energy per unit length in the straight-supercoiled configuration as a function of the applied force F .

3.1.3 Interaction energy in high multi-valent salt

When multivalent ions dominate the solution, the electrostatic potential is known to be repulsive-attractive [100].

We assume that the electrostatic interaction between the helices is approximately equal to the interaction between DNA strands in a hexagonal array. The validity of our assumption will have to be borne out of the comparison with experimental data. It was certainly adequate for studying forces during DNA packing in viruses [114, 115]. Therefore, we prescribe the electrostatic contribution

$U_{es}(r, \theta) \approx U_{es}(r)$, where $U_{es}(r)$ is obtained by relating the osmotic pressure to the total energy of a hexagonal array of parallel DNA filaments [13, 100]. In the repulsive-attractive regime Todd et al. [13] propose a phenomenological model for the total pressure between the two DNA strands:

$$\Pi = -C_A \exp\left(-\frac{D}{\lambda}\right) + C_R \exp\left(-\frac{2D}{\lambda}\right), \quad (3.14)$$

where $D = 2r$ is equal to the spacing between the strands, $\lambda \approx 4.8 \pm 0.5 \text{ \AA}$ is the characteristic decay length. The parameters C_A and C_R describe the attractive and repulsive interactions respectively. As stated in [13] at the equilibrium spacing $D = D_{eq}$ the attractive and repulsive forces are equal such that $\Pi = 0$ and consequently $C_R = C_A \exp(D_{eq}/\lambda)$. The electrostatic potential per unit length is given by [13]:

$$\begin{aligned} U_{es}(r) &= - \int_{\infty}^{2r} \sqrt{3} \Pi D dD \\ &= \frac{\sqrt{3} \lambda}{e^{4r/\lambda}} \left[\frac{C_R}{4} (4r + \lambda) - C_A (2r + \lambda) e^{2r/\lambda} \right] \end{aligned} \quad (3.15)$$

The equilibrium spacing D_{eq} for some condensing agents has been measured experimentally and is $D_{eq} \approx 2.8 \text{ nm}$ [13]. The factor of $\sqrt{3}$ enters because each DNA strand is surrounded by 6 others in a hexagonal array.

The total internal energy is the result of adding the effects of the electrostatic potential U_{es} and the cost of configurational entropy $U_{conf-helix}$. The free energy of entropic confinement per unit length of the strand in the plectonemic supercoil is given by [20, 84]:

$$U_{conf-helix} = \frac{k_B T}{A^{1/3}} \left[\frac{c_p}{(p\pi)^{2/3}} + \frac{c_r}{r^{2/3}} \right] \quad (3.16)$$

where $A = K_b/(k_B T)$ is the persistence length of the DNA chain. The terms c_r and c_p are in general unknown constants fitted to experiment [84]. In our calculations we will use $c_r = c_p = 2^{-8/3}$ which are empirically optimized constants as given by van der Maarel [84]. The term $2\pi p$ is the pitch of the helix and is given by $p = r \cot \theta$. In the expression given for $U_{conf-helix}$ we assume that the undulations in the radial direction of the helix are not restrained by electrostatics but only by the structure of the plectoneme. When using $c_r = c_p = 2^{-8/3}$, the free energy of entropic confinement per unit length is small compared to the mechanical and electrostatic energetic contributions. So, small errors in the values of c_r and c_p can be subsumed into the value of C_A which we treat as a fitting parameter.

3.1.4 Solution for the superhelix parameters

Using $U(r, \theta) = U_{es} + U_{conf-helix}$ with C_A and D_{eq} values found in the hexagonal array experiments of Todd et al. [13], it is interesting to find that the system of equations Eq. (3.11) - Eq. (3.13) has a nontrivial solution only for forces larger than a threshold force F_T . Table 3.1 summarizes the qualitative effects of the bending modulus K_b and the electrostatic parameters (C_A, D_{eq}) on the supercoiling variables and the slope of the hat curves, as a function of the applied force F .

In Fig. B.1 in the Appendix B we present a more detailed analysis of the effects of the parameters K_b, C_A and D_{eq} on the supercoiling variables. As a general trend the helical radius r is approximately constant and close to $D_{eq}/2$. As $F \rightarrow F_T$ the slope increases sharply, $2r \rightarrow D_{eq}$, and the values of M_3 and θ decrease, which means that the helices are getting less tight (see Fig. B.1). The solution to the system of equations Eq. (3.11) - Eq. (3.13) show that the equilibrium values of θ and M_3 approach a positive nonzero value as $F \rightarrow F_T$ because of the direct dependence of the internal energy U on

Parameter	F_T	M_3	θ	r	$d\Delta z/dn$
$C_A \uparrow$	\uparrow	\downarrow	\downarrow	\uparrow	\uparrow
$K_b \uparrow$	\downarrow	\uparrow	\downarrow	--	\uparrow
$D_{eq} \uparrow$	\downarrow	\uparrow	\uparrow	\uparrow	\downarrow

Table 3.1: K_b , C_A and D_{eq} effects. A variable's increase and decrease are denoted by \uparrow and \downarrow respectively. The line -- means no significant effect.

$(\cot \theta)^{-2/3}$.

In the next section we will analyze the DNA ‘phase-transition’ into supercoiled structures.

3.1.5 DNA transitions and formation of supercoils

When multivalent ions dominate the solution, the results from the transition have some similar features to the ones found for monovalent ions including the possibility of a transition straight \rightarrow straight-supercoil and the transition straight-supercoil \rightarrow pure plectoneme (for sufficiently large number of turns). But the dominance of multivalent ions leads to new features of the transition such as the possibility of collapse from a completely straight DNA configuration into condensed structures such as toroids, due to the attractive-repulsive nature of the DNA-DNA interactions (see section 3.2.1). Next we outline the method to calculate the transition between the Straight and Supercoiled configurations that is described in detail in [21].

3.1.6 Transition from straight to straight-supercoil co-existence

The transition from the straight to the straight-supercoil co-existence is characterized by a critical force $F_{critical}$ as a function of the number of turns or equivalently a critical number of turns $n_{critical}$ as a function of the applied force F . As noted in [10, 21, 74], at the transition point there is a jump in the torque $\delta M = M_{critical} - M_3$ together with a jump in the vertical extension δz . The jump in the extension corresponds to the appearance of the end loop and a sudden formation of a supercoil with $l_p^{critical} > 0$, where $l_p^{critical}$ is the length of the helical region right after the transition. To find $(n_{critical}, F_{critical}, \delta M, \delta z$ and $l_p^{critical})$ we follow the method described in [21]. We compute the transition variables by recognizing (a) that at the transition the free energy of both configurations is equal, and (b) that the linking number $n = Lk$ is a topological invariant that must be continuous at the transition. Comparing the energies Eq. (3.9) and Eq. (3.1), together with $n_{c-s} = n_{c-p}$, readily yields $l_p^{critical}$, $M = M_{critical}$, $n_{critical}$ and the other transition variables.

Using proper parameters in the interaction energy $U(r, \theta)$ as calculated in [13], we found that as F decreases the length of the supercoil that appears just after the transition $l_p^{critical}$ increases and the torque just before the transition $M = M_{critical}$ decreases. Furthermore, for the force F_T the solution to the minimization of the energy V_{p-s} yields $M_3, \theta, r > 0$ and $l_p^{critical} < L - l_o$. Since M_3, r and θ have non-zero values, by the use of equations Eq. (3.5) - Eq. (3.8), it is clear that the number of turns at which the DNA molecule transitions into supercoiling is larger than zero. The number of turns $n_{critical}(F_T)$ is in fact equal to the minimum number of turns n_{min} for which supercoiled structures can coexist with straight DNA. We point out that the numerical calculation of F_T based on comparing the energy of two states neglects the presence of thermal kicks which can lower or raise the energy barrier for the transition between states. In general for a given force $F \geq F_T$, the transition would take place over a range of values of n . Consequently, the computed values of F_T and $n_{critical}$ at the transition are estimates. Furthermore, we know that the transition is a dynamic process that needs further study including adequate kinetic analysis.

3.1.7 Transition to the fully supercoiled state: pure plectoneme

So far we have neglected in our description of the problem the geometric constraint $l_p \leq L - l_o$. As more turns are added, beyond $n_{critical}$, the straight l_t portion decreases and l_p increases. As a consequence of the above constraint, holding F constant, as n increases it reaches $n_{critical}$ for which supercoils and tails coexist. As n is further increased it reaches a value $n = n_{pp} > n_{critical}$ where $l_p = L - l_o$ and the DNA has been entirely converted into a pure plectoneme. The number of turns as a function of F at which the transition into the pure plectoneme takes place is given by:

$$n_{pp}(F) = n_{critical}(F) + \frac{L - l_p^{critical}(F) - l_o(F)}{dl_p(F)/dn}, \quad (3.17)$$

where dl_p/dn is the change in the length of the helices as a function of the number of imposed turns [21]:

$$\frac{dl_p}{dn} = \left[\frac{\sin 2\theta}{4\pi r} - \frac{M_3}{8\pi K_b K} \right]^{-1}. \quad (3.18)$$

So far we have theoretical estimates of $n_{critical}(F)$ and $n_{pp}(F)$ in n vs. F space describing DNA transitions in single molecule experiments. This would prove to be useful in section 3.3.1 to understand the description of the complete phase diagram of DNA configurations when we include the possibility of formation not only of supercoils, but also toroidal structures. So, next we consider the possibility of DNA condensation into toroidal structures under imposed tension and end rotations.

3.2 DNA Globular State

Experimental studies of DNA condensation (induced by multivalent ions) in the bulk have shown that DNA condenses into toroidal structures [104–107]. Consequently the globular state is modeled as a torus. We will denote by a the radius of the tube and R the distance from the center of the tube to the center of the torus (See Fig. 3.1). The free energy in the globular state in the absence of twist can be found in refs. [111, 112]. Here we include the possibility of applying moments to the DNA chain. The free energy of the globule is approximately given by:

$$V_g = L_g \left(G_g + \epsilon \frac{K_b}{R^2} + \frac{M_g^2}{2K_t} \right) + \frac{E_s R a}{d^2}, \quad (3.19)$$

where L_g is the DNA length in the condensed-globular state. The first term in parenthesis $G_g < 0$ is the energy of intersegment interaction per unit length and is negative due to attractive nature of the interactions due to multivalent ions in condensed DNA [13, 108, 112]. As stated in [13] there are ‘nominal configurational’ differences between globular DNA condensates and parallel DNA strands under osmotic stress, but in both cases the intermolecular interactions between DNA filaments are the predominant effect. Minimization of the surface and bending energy in the toroidal condensates would theoretically lead to the same local structure as parallel arrays [13]. Therefore $G_g = U(D_{eq})$ is the interaction energy given by Eq. (3.15) at the equilibrium distance [13]. In the Appendix B we describe a different model suggested by Battle et al. [98] to account for the torus internal energy interactions which leads to similar results.

The second term in Eq. (3.19) is the bending energy and the entropic cost of confinement. An estimate of the confinement free energy per unit length of the toroid can be computed from the configurational cost of a long persistent chain confined into a sphere. It is of the order of the bending energy of the toroid $\sim 2K_B R^{-2}$ [116]. If the configurational entropy is neglected then $\epsilon = 1/2$ from the bending contribution. The third term in Eq. (3.19) is the twisting energy per unit length, where M_g is the induced moment in globular state due to end rotations. The parameter E_s in Eq. (3.19)

represents the surface tension energy of a toroid of volume $\propto Ra^2$. The effective diameter of the DNA chain in the toroidal state is $d = D_{eq}$ and its value is regulated by the internal interactions of the loops making up the toroid. The free energy V_g has to be minimized subject to two constraints. The first constraint is given by the requirement that the volume of the globule (torus) in the condensed state has to be approximately equal to the volume of DNA chain [111]:

$$L_g d^2 \approx c R a^2, \quad (3.20)$$

where we have introduced the constant $c \in [1, 8\pi]$ to account for the prefactors in the volume of a distorted torus. Further discussion about the constant c can be found in the Appendix B. The chain forms approximately $N_g \approx L_g/(2\pi R) = (ca^2)/(2\pi d^2)$ loops [117], also known as the torus winding number [98]. The second constraint says that the number of turns n imposed on the DNA chain has to be equal to the linking number Lk_g (a topological invariant) of the globule:

$$Lk_g = \frac{M_g}{2\pi K_t} L_g + Wr, \quad (3.21)$$

where the first term is the twist of the chain Tw and Wr is writhe. The total writhe present in the globule is $Wr = Wr^+ - Wr^-$, where Wr^+ are the positive turns Wr^- are the negative turns under a given sign convention. However, note that the number of loops in the globule is equal to $N_g = Wr^+ + Wr^-$, such that $Wr = N_g - 2Wr^-$ and Wr^- is an unknown. Next we introduce the constraints Eq. (3.20) and Eq. (3.21) via the Lagrange multipliers μ and λ :

$$\hat{V}_g = V_g + \mu \left(L_g - cR \frac{a^2}{d^2} \right) + \lambda \left(Lk_g - \frac{M_g L_g}{2\pi K_t} - \frac{ca^2}{2\pi d^2} + 2Wr^- \right). \quad (3.22)$$

Then energy \hat{V}_g has to be minimized with respect to the unknown variables M_g, a, R and Wr^- . Performing $\partial \hat{V}_g / \partial M_g$ yields $\lambda = 2\pi M_g$. The second partial derivative $\partial \hat{V}_g / \partial a = 0$ gives:

$$\mu c \frac{aA}{k_B T d} = \frac{\gamma}{2} - M_g \frac{caA}{k_B T R d} \quad (3.23)$$

where $\gamma = AE_s/(k_b T d)$. The third partial derivative $\partial \hat{V}_g / \partial R = 0$ yields:

$$\frac{2\epsilon L_g}{R^3} = \frac{a}{A^2 d} \left[\gamma - \mu c \frac{aA}{k_B T d} \right], \quad (3.24)$$

and combining Eq. (3.23) and Eq. (3.24) we obtain a relationship between R and a as a function of M_g :

$$\frac{2\epsilon L_g}{R^3} = \frac{a}{A^2 d} \left[\frac{\gamma}{2} + M_g \frac{cAa}{k_B T R d} \right]. \quad (3.25)$$

Performing $\partial \hat{V}_g / \partial Wr^- = 0$ results in the condition $M_g = 0$, such that $Lk_g = Wr = N_g - 2Wr^-$ and there is no twist present in the globule. Setting $M_g = 0$ in Eq. (3.25) and using the constraint Eq. (3.20) we recover the solution for the size of the DNA condensate in the absence of twist [111, 112]:

$$R^o \approx \frac{A}{c\alpha^{2/5}} \left(\frac{L_g}{A} \right)^{1/5}, \quad a^o \approx d\alpha^{1/5} \left(\frac{L_g}{A} \right)^{2/5}, \quad (3.26)$$

where $\alpha = \gamma/(4\epsilon c^3)$ is a constant. The globule's winding number is given by:

$$N_g^o \approx \frac{ca^2}{2\pi d^2} = \alpha^{2/5} \frac{c}{2\pi} \left(\frac{L_g}{A} \right)^{4/5}, \quad (3.27)$$

and $Wr^- = (N_g^o - n)/2$. From the energy minimization procedure the obtained values of R^o and a^o yield a surface tension energy that is four times the bending energy of the system. In the case where $M_g = 0$, since the relations given in Eq. (3.26) are independent of the number of turns n and force F , the equilibrium energy of the globular system given by Eq. (3.19) is a constant independent of the controlled variables F and n .

The results Eq. (3.26) and Eq. (3.27) given by the condition $M_g = 0$ are independent of the number of turns n and reflect that the DNA chain can accommodate all the external applied link into the loops forming the torus. It should be noted that $Wr^- > 0$ is an independent variable as long as the number of applied turns n is smaller than the number of loops N_g^o that the globule can accommodate. When n is larger than N_g^o given by Eq. (3.27), the extra link that the torus can not accommodate as loops has to be stored as twist. If $n > N_g^o$, then $Wr^- = 0$, $M_g \neq 0$ and the solution to the minimization of \hat{V}_g is given by Eq. (3.25) together with the constraints Eq. (3.20) and Eq. (3.21). Note that the system will adopt the same solution with $M_g \neq 0$ in Eq. (3.25) for all values of n if due to kinetic or other constraints the globule is forced to store single signed loops (eg. $Wr^- = 0$). As expected for $M_g \neq 0$, the energy of the toroidal system is no longer independent of the number of turns n . Therefore we see that if $n \leq N_g^o$, then the theoretical scenario with $M_g = 0$ is the limiting case where the system is always able to reach thermodynamic equilibrium, while if $n \leq N_g^o$ then the solution with $M_g \neq 0$ in Eq. (3.25) is similar to taking into account some constraints that might force single signed loops in the toroid. Finally, we point out that the results obtained following the methods described in this section are valid as long as the results yield $R > a$, for which the DNA has a toroidal structure [111].

3.2.1 DNA Condensation: coil-globule transition

Mamasakhlisov et al. [112] recently developed a statistical mechanical model, based on a mean field approach and the Zimm-Bragg model, for the condensation of an insoluble flexible polymer under tension and no twist. According to their theory, for an infinite length chain the transition is first order, while for finite length polymer the transition is sharp and progresses over a small interval of tension.

To consider the straight to globule transition under tension and imposed end rotations we simplify the treatment given by Mamasakhlisov et al. [112] and use a similar idea to the transition model used in [108]. We assume the transition to be sharp (highly cooperative transition and small Zimm-Bragg cooperativity value) such as in the case of the homogeneous persistent chain [112]. This assumption is based on recent DNA single molecule experiments in multivalent salts done by van den Broek et al. [37], where a single nucleation cite for the toroid was observed, which leads to the idea that the surface energy term in the free energy of a toroidal condensate is indeed large. We can now compare the energy of the straight configuration (right before the transition) with the energy of the condensed-toroid (right after the transition). The energy of the condensed-torus is described in section 3.2 and is given by Eq. (3.19) where $L_g = L$, such that the entire DNA has collapsed and is in the toroidal state. We can estimate the value of the force F_{c-o} (as a function of the number of turns n) for which the transition (straight \rightarrow toroid) occurs as done previously by noting (a) that at the transition the energy of the straight configuration and condensed-globular configuration are equal, and (b) that the linking number $Lk_g = n_{c-s} = n_{sg}$ is a topological invariant that must be continuous at the transition. First we consider the case where $n \leq Wr_g^o$ together with $M_g = 0$. Comparison of the energies given by Eq. (3.1) and Eq. (3.19) together with the requirement of no twist in the globular region yields the relationship between the transition force F_{c-o} and the number of turns n_{sg} at the transition:

$$F_{c-o} = \frac{M^2}{2K_t} + G_{flu}^* + \frac{M^2}{4K_b K} - G_g - 5\epsilon K_B \left(\frac{c^5 \alpha^2}{A^4 L_g} \right)^{2/5} \quad (3.28)$$

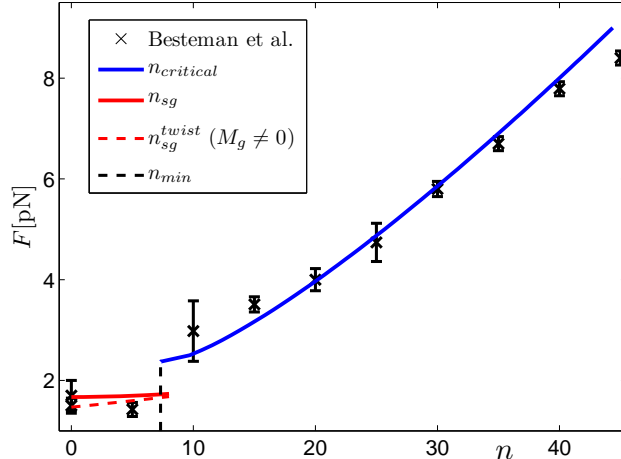


Figure 3.2: The curve of $n_{critical}$ shows the transition from straight configuration to supercoiled + straight coexistence. The curve n_{min} shows the minimum number of turns at which supercoiled + straight can coexist. The curve n_{sg}^{twist} (n_{sg} for toroids with no twist) shows the transition from straight to toroidal configurations. Spermine is the condensing agent, and we have used $C_A = 690\text{pN/nm}^2$, $K_b = 50k_B T$ and $\alpha = 0.013$ ($\beta \approx 2.6$ and $\Gamma d^2/(k_B T) \approx 0.6$).

where a^o and R^o have been replaced with the solution given in Eq. (3.26). The value of the moment M in the straight configuration is prescribed as a function of the number of turns at the transition n_{sg} given by Eq. (3.4).

In the case where $n \geq W r_g^o$ or in the case when the torus can only store one signed turns, comparison of the energies between the two states yields the critical force as a function of the number of turns n_{sg}^{twist} :

$$F_{c-o}^{twist} = \frac{M^2 - M_g^2}{2K_t} + G_{flu}^* + \frac{M^2}{4K_b K} - G_g - \epsilon \frac{K_b}{R^2} - \frac{E_s R a}{L d^2} \quad (3.29)$$

where M_g , a and R are given by solving the system of equations Eq. (3.20), Eq. (3.21) and Eq. (3.25) and M is given by Eq. (3.4) as a function of the number of turns at the transition in the presence of twist n_{sg}^{twist} .

We close this section by pointing out that although the step size in single molecule experiments may be dictated by a kinetic process we are interested in the final loop sizes governed by the thermodynamic parameters. We assume that for sufficiently long waiting times DNA will reach thermodynamic equilibrium. Our assumption is supported by the experimental studies [104–106] in the bulk under the introduction of static loops and the computer simulations in ref. [118]. For a brief discussion of the final size of the toroid we refer the reader to the Appendix B.

3.3 Results: Comparison with experiment

The experiments by Besteman et al. [15, 16] were performed for a DNA sample of length $L \approx 2720\text{nm}$ and for three different condensing agents: spermine, cosep and protamine. To see if we can fit their data we started by solving the transition from straight DNA to supercoiled-straight configuration. For the twisting modulus we have used $K_t = 86k_B T$ as done by Besteman et al. [15], while for the bending modulus we have used $K_b \approx 50k_B T$ for spermine (as done in [15, 37]), and $K_b \approx 25k_B T$ both for cosep and protamine. There is evidence that there is a substantial decrease in the bending modulus in the presence of condensing agents, where cations with higher charge density have a stronger effect [119]. The bending moduli have been measured to decrease up to $K_b \sim 15k_B T$ for

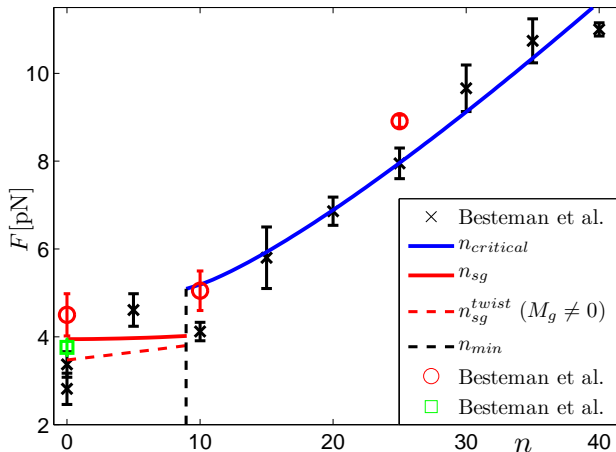


Figure 3.3: Same predictions as in Figure 3.2, but using cosep as the condensing agent. We have used $C_A = 1150\text{pN}/\text{nm}^2$, $K_b = 25k_B T$ and $\alpha = 0.013$ ($\beta \approx 7.5$ and $\Gamma d^2/(k_B T) \approx 1.1$). The black crosses, green squares and red circles correspond experimental data in Besteman et al. [15] for different molecules under same experimental conditions. The inset show the comparison to the experimental data using protamine as the condensing agent [16]. For protamine we have used $\beta = 1$ and the same values of C_A , D_{eq} and K_b as the ones used for cosep, since both condensing agent have a high charge density. Using K_b as low as $15k_B T$ still gives very good agreement with experiment for both protamine and cosep.

strong condensing agents such as cohex [108]. For the intersegment equilibrium distance $d = D_{eq}$ we have used $D_{eq} = 2.815\text{nm}$ for spermine as measured by Todd et al. [13], $D_{eq} = 2.7\text{nm}$ for protamine as done in [16] and approximated $D_{eq} \approx 2.7\text{nm}$ for cosep based on the measured value for other condensing agents [13]. The phenomenological parameters C_A and C_R used for the condensing agents were obtained as estimates from the measurements performed for different condensing ions in [13]. Finally, for the decay length λ we have used for spermine and cosep $\lambda = 0.46\text{nm}$ as done in [13] and for protamine we have used $\lambda \approx 0.485\text{nm}$ which is consistent with the values used by DeRouchey and Rau [103]. The blue solid line in Fig. 3.2 and Fig. 3.3 shows the transition from straight \rightarrow straight-supercoiled as a function of the controlled variables F and n . Our theoretical predictions accurately match experiment for moderately large number of turns. The black dashed line shows the minimum number of turns n_{min} for which supercoiled structures coexist with straight DNA portions (plectonemes with tails). For $n < n_{min}$ there is no solution to the transition problem straight \rightarrow straight-supercoiled. Interestingly, the theoretical value of n_{min} appears to agree with a *discontinuous* jump in Besteman et al. data for the condensing agents, where there seems to be a region of experimental points that is approximately independent of n (zero slope) and region where the points increase as a function of n .

Next we consider the possibility that DNA molecules can collapse into toroidal structures for both spermine and cosep. We use the same electrostatic parameters (C_A , C_R , D_{eq}) and bending modulus K_b that we used for the straight to supercoil transition. We neglect here the confinement cost in toroidal structures and set $\epsilon = 1/2$ (the effect of ϵ and other parameters on the calculations is described in the Appendix B).

The only parameter left is the surface tension term α ($E_s \propto \alpha$) in the free energy of the torus. The surface energy contribution for DNA condensates has not been completely worked out in the literature and we could not find specific values for any condensing agent. In one of their papers Ubbink and Odjik [120] suggest that the surface tension energy $E_{surface} \approx Ak_B T V^{1/3} d^{-2} \beta$ becomes important when the dimensionless parameter $\beta \sim O(1)$, where $V = \pi L(d/2)^2$ is the volume of the torus. Using our notation $\beta = 4\epsilon(cd)^2 V^{-1/3} (\alpha^4 L^3 A^{-8})^{1/5}$. In another paper Ubbink and Odjik [121] suggest that for $E_{surface} \approx \Gamma S$, where Γ is the surface tension parameter and $S = (2\pi)^2 R a$ is the surface

area of the torus, $\Gamma d^2/(k_B T) \sim O(1)$. In our notation $\Gamma S/(k_B T) \approx 4\epsilon c^2(\alpha^4 L^3 A^{-3})^{1/5}$. In their work on condensation Park et al. [117] used estimates of the cohesive energy where $E_{surface} = \Gamma S$ and $\Gamma d^2/(k_B T) \approx 0.06 \sim O(1/100)$. The values used by Park et al. and those suggested by Ubbink and Odjik differ greatly in magnitude. Finally, in ref. [116] the surface tension parameter is suggested to be $|E_o| \in [0.1 - 0.5] \text{ nm}^{-1}$, where our parameter $\alpha \propto |E_o| A c^{-5/2}$. Due to the wide range of suggested values for the surface tension parameter we take α as a fit to the point $n = 0$ in Besteman et al. [15] experiments. To validate the fitted value of α we compare it to the previously suggested values of the surface tension in the literature [116, 117, 120, 121], using them as upper and lower bounds. For both condensing agents in Figs. 3.2 and 3.3 the fitted value is $\alpha = 0.013$. This means $\beta \approx 2.6$ and $\Gamma d^2/(k_B T) \approx 0.6$ for spermine, while $\beta \approx 7.5$ and $\Gamma d^2/(k_B T) \approx 1.1$ for cosep, since $d = D_{eq}$ and A depend on the condensing agent. In both cases the surface tension parameter is of the order suggested by Ubbink and Odjik [120, 121].

The red solid line in Figs. 3.2 & 3.3 shows the transition from straight \rightarrow toroids (with no twist) as a function of the controlled variables F and n . The red dashed line shows the transition for toroids with twist. Our theoretical predictions show good quantitative agreement with experiment for $0 < n < n_{min}$. The solution to the problem in thermodynamic equilibrium shows that the supercoiled-straight configurations are energetically more favorable than the toroidal condensates for $n \geq n_{min}$ and $F > 0$.

The solution to the problem in thermodynamic equilibrium where supercoiled structures compete with toroidal condensates is given by a comparison of the energies Eq. (3.9) and Eq. (3.19). In the Appendix B we address how this competition of DNA configurations plays out as a function of persistence length A , the toroid configurational entropy parameter ϵ , and the toroid surface tension parameter α .

In the next section we complete a phase diagram based on the DNA configurations we studied.

3.3.1 Phase Diagram: DNA configurations

Fig. 3.4(a) shows a schematic of a phase diagram as a function of the controlled variables (F, n) . We have used the theoretical solution for the description of Besteman et al. experiments in spermine (see sub-section 3.3). For the labeling of each region see the caption of Fig. 3.4. The light blue region labeled U is an unknown region in the phase diagram. The minimization of energy over the studied configurations does not yield a favorable state in the unknown region. In this region $F < F_T$ such that there are no supercoiled-straight configurations and the number of turns n is not sufficiently large to consider a pure plectonemic state where the energy is approximately independent of F . The unknown region is below the yellow region (which correspond to a collapsed DNA structure), yet based on our theoretical analysis of constrained minimization, in parts of the unknown region the straight configurations are preferred to toroidal structures. This would mean that as F is decreased the DNA can collapse into supercoiled structures and as F is further reduced then it will go back to a straight configuration. This seems counterintuitive, so we identified this region as unknown. We suspect that in the unknown region, toroids and plectonemes compete and that the configurations in this region are more complex forms of DNA compaction which can be clarified by further experiments or simulations.

The solution for straight \rightarrow straight-supercoiled configurations for $F \sim F_T$ yields helices with large pitch $2\pi r \cot \theta$ ($\theta \rightarrow 0$) meaning that the DNA is approaching the configuration of two parallel rods. The value of the supercoiling angle θ is nearly constant for $F > F_T$, but as $F \rightarrow F_T$ the value of the supercoiling angle θ decreases sharply as depicted in Fig. B.1. As the pitch increases the two DNA filaments ‘see’ each other less and it is expected that the magnitude of the internal interactions will decrease. This is the case for DNA in monovalent salt solutions as evidenced by the energy expressions provided by Ubbink and Odjik [20]. Therefore even though for $F > F_T$ the phenomenological model $U_{es}(r)$ given in Eq. (3.15) for the helical region provides good estimates when compared to experiment, this approximation should become less accurate for F close to F_T . For

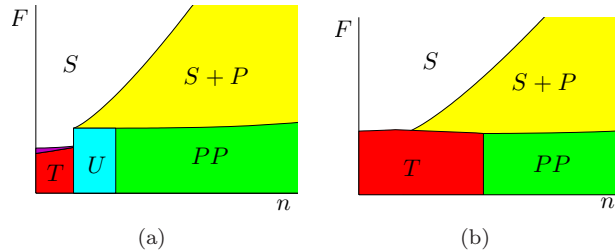


Figure 3.4: Phase diagrams constructed using the theoretical solution for the description of Besteman et al. experiments in spermine. In (a) we used $\alpha = 0.013$ as in Fig. 3.2 and in (b) $\alpha = 0.0003$. The straight configuration is the most favorable one in the white region and it is marked with S . The straight-supercoiled configuration is marked with $S+P$ and colored yellow, while the entirely supercoiled structure (pure plectonemes) is marked with PP and colored green. T stands for the region where we expect to see toroidal structures. The red color shows the region where we expect to see toroids (with or without twist) and the purple region where we expect to see toroids only without twist. The light blue region labeled U in (a) corresponds to an unknown region where there is no clear favorable configuration.

$F \sim F_T$ the internal energy becomes less negative and we expect the actual value of the transition force F to decrease when compared to the theoretical case used to construct the phase diagram in Fig. 3.4(a). Considering this effect would modify the $S \rightarrow S+P$ transition curve close to the region $F \sim F_T$ and $n \sim n_{min}$ in Fig.3.4(a). Consequently, we expect our theoretical solution for the transition close to F_T to provide some qualitative insight to the transition problem rather than an accurate prediction. In [122, 123] the authors provided a thorough analysis of a collapse of single, stiff polymers in poor solvent in the absence of twist and tension. It was suggested that the condensation proceeds via a cascade through metastable intermediates (‘racquets’) toward the equilibrium configuration which is the torus. We speculate that the high pitch solutions (that approach two parallel filaments) resemble the ‘racquets’ studied in [122, 123].

In Fig. 3.4(b) we show a scenario where we have decreased the surface tension parameter $\alpha = 0.0003$ (approximately 40 times smaller than the one used to fit Besteman et al experiments). The solution yields toroids of mean radius $R \sim 120$ nm. We can see that there is a region in (F, n) space, where DNA makes a transition straight-supercoiled \rightarrow toroid and that there is no unknown region. The unknown region tends to disappear for $\alpha \lesssim 0.001$ ($\Gamma d^2 \sim 0.1 k_B T$) which is of the order of the value used by Park et al. [117] and an order of magnitude lower than the ones used to fit Besteman et al. data in Fig. 3.4(a)

3.4 Conclusions

We have studied the mechanics of DNA in the presence of condensing agents. The result of our analysis is a phase diagram showing the most favorable DNA configuration as a function of the controlled variables – force F and number of turns n . Since the possible physical mechanisms of counterion-induced DNA condensation are still debated in the literature [13, 81], our theoretical-mechanical model can be used as a framework to test DNA-DNA interaction theories. Our model allows to easily see the effects of parameters such as K_b and D_{eq} (which are specific to the condensing agent and its concentration [13, 119]) on the final equilibrium structure. We have constrained our analysis to the observed DNA configurations – straight, supercoiled and toroidal. Here we give a short summary of all the results we have obtained.

A key result of our analysis is that there are no supercoiled solutions to the equations of equilibrium below a critical force F_T in the presence of condensing agents (see Fig.B.1). For moderately large values of n and F (not too close to F_T) the supercoiled structures are the most favorable state, but as the number of turns increases (or tensile force decreases for moderately large n) there is a

transition into pure plectoneme. For low values of n toroidal structures are favored when $F < F_T$. Our theoretical estimates of the transition force and number of turns n describe Besteman et al. experiments very accurately away from the point $F \sim F_T$ and $n \sim n_{min}$. The predicted DNA configurations from our phase diagram can be tested experimentally using fluorescence imaging methods to visualize the formation of the collapsed structure as done in [37]. Also, if there is formation of supercoiled structures, this process should be reversible and therefore there would be no hysteresis in the loading and unloading curves as long as the supercoils persist. This is the case in the formation of plectonemic structures in monovalent salt, [10–12] when there are no ‘structural’ changes in the DNA molecule [41, 44, 57]. Furthermore, as noted in Fig.B.1 in the Appendix B, our model predicts that for moderately large F the slopes of the hat curves (extension vs. rotation) are relatively constant as a function of n , while the value of the slope sharply increases as F decreases approaching F_T . The same methods used to measure the torque and slope in single molecule experiments using monovalent salts can be used in the presence of condensing agents and therefore test our theoretical predictions.

Our treatment of the toroidal structures assumes a large continuum limit where $R > a \gg D_{eq}$ with no geometrical defects. The final toroidal state depends on the value of the surface tension parameter α whose value remains unresolved. In Fig 3.2 the red solid line showing the transition from straight \rightarrow toroids, matches quantitatively the experimental data. But the match between theory and experiment is not so obvious in Fig. 3.3, where experiment shows a large range of force values for $n = 0$ and $n = n_{min}$ at the transition from straight \rightarrow toroids. Changing the value of α (while still inside the range used previously in the literature [117, 120, 121]), modifies the phase diagram (see Fig 3.4) for low values of F and n . Using different values of α can reproduce the entire range of values of Besteman et al. experiments in Fig. 3.3. Possible improvements to our theoretical framework include modeling thicker toroids when $R \sim a$ [124], deformed toroids [120] and toroidal structures with topological defects [117]. Also, we have performed only a constrained minimization of the free energy over a class of equilibrium structures which leads to a region in the phase diagram where the preferred configuration is unknown. More sophisticated calculations and experiments will be needed to determine what the correct equilibrium shape of the DNA is in that region.

Chapter 4

DNA superhelical Structures with non-constant helical pitch

Twisting leads to the formation of supercoils (also called plectonemes) and braids, both of which typically involve two interwound DNA filaments. In the literature *supercoil* refers to a superhelical structure with constrained linking number in each DNA filament. The term *braid* is reserved for superhelical structures where twist is not present and the combined linking number of the two filaments is controlled [125]. Since braids are supercoils with no twist, we will refer to both interwound structures as ply structures.

Supercoiling (plectonemes) takes place in experiments where a single filament of DNA is subjected to tension while turns are added to it as depicted in figure 4.2. This type of experiments are usually called rotation-extension experiments [10–12, 15, 27, 63, 64]. In rotation-extension experiments the vertical extension of the DNA and the external moment are recorded as functions of the number of turns. A well-known feature of the experimental data is a regime, corresponding to the formation of plectonemes, where there is almost a linear relationship between the DNA extension and the applied number of turns. Also, as shown in recent experiments of Forth et al. [12], Lipfert et al. [27] and Mosconi et al. [11], the external moment is approximately constant in this regime. Marko [55] and Marko and Siggia [52] show that the linear relationship between extension and number of turns can be captured using a two-phase model (phase 1 is straight DNA and phase 2 is superhelical DNA) in which the torque plateaus during the transition. Braiding can be seen in experimental set ups as the one depicted in figure 4.1 where two pieces of nicked DNA (that cannot carry twist) are interwound as shown in Charvin et al. [17]. A theoretical model for braided structures subjected to high and low tensions including the effects of thermal fluctuations has been proposed in Marko [125]. As mentioned in Chapter 2, supercoiling has been widely studied in the mechanics literature. In Fraser and Stump [65] study the effects of a loaded ply of constant angle while Coleman and Swigon [66] describe the mechanics of a variable balanced ply with no end loads, where the helical angle varies. Thompson et al. [67] and van der Heijden et al. [70] extend the variable ply formulation of Coleman and Swigon by considering the effects of end loads. Both Thompson et al. [67] and Coleman and Swigon [66] focus on variable plies formed from closed rods with two end loops (see figure 2 in Thompson et al. [67]). In more recent work Coleman and Swigon [126] extend some of their earlier results to variable plies formed from open rods with one end loop like the plectoneme depicted in figure 4.2. In the past few years Monte Carlo simulations have also been used to study DNA braiding [17] and supercoiling [5].

An objective of the present work is to develop an elastic-isotropic rod model for twisted DNA in the plectonemic regime where the helical angle is varying as a function of the arc-length of the rod. Our approach will follow Argudo and Purohit [21], where a variational formulation is used to

solve for the geometry of the plectoneme. Once the assumption of a constant curvature in the helical region is relaxed, we show why single molecule experiments as the one depicted in figure 4.2 can be accurately described by a uniform helical angle solution [4, 21, 45, 73]. .

Another objective of this work is to provide variable pitch solutions that are applicable in certain types of single molecule experiments [18]. Solutions for the variable pitch ply using essential (rigid/displacement loading) boundary conditions and natural (dead/force loading) boundary conditions where the length of the ply is fixed are available in the literature [70]. Here we relax the assumption of fixed length of the ply and solve the problem making use of the theory of optimal variable end-point in variational problems [127, 128]. We see potential applications of variable pitch helices in single molecule experiments where the external moment is not constant as a function of the number of applied turns, or in dual DNA manipulation single molecule experiments [18] where a force and number of turns can be applied to DNA molecules with no tails. Variable pitch solutions give rise to the possibility of lock-up in the helix, which can lead to formation of more complex DNA structures. Here we provide a complete set of analytical solutions for variable pitch helices for the case when the interaction energy in plectonemic DNA is approximately independent of the helical angle. This is a good approximation in the presence of high concentration of multivalent ions [22].

To present the theory in an organized manner we have divided the chapter into two sections. The first section describes a model for extended ply structures (including braids), as the one depicted in figure 4.1, where the force is parallel to the axis of the helical structure. The second section describes a model for plectonemes, as the one depicted in figure 4.2, where the force is perpendicular to the axis of the helical structure.

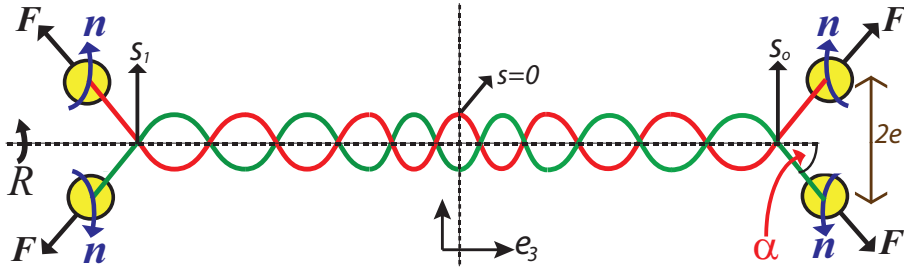


Figure 4.1: Sketch representing single molecule experiments where supercoils are formed. Two DNA molecules are fixed at one end, while the other end is connected to a magnetic or optical bead. The beads apply a pulling force F and twist the DNA strands by a given number of turns n . Moreover the beads can be moved in space to apply a given number of rotations R about the helical axis e_3 . If the DNA strands are nicked then the beads can not fix n and the molecules cannot carry twist. This leads to the formation of braids as in the case presented by Charvin et al. [17]. Similarly, in the dual DNA manipulation set up of Noom et al. [18] the beads are free to rotate and can not impose n . So there is no twist in the molecules. But, in general, an experimental set-up using optical or magnetic traps including twist can be realized [11, 12].

4.1 General Description of Model 1: extended ply structures

We model DNA as an elastic rod with bending moduli K_b and twisting modulus K_t . This model is based on a coarse grained representation [129] where base-pair details are neglected. Figure 4.1 gives a sketch of a dual angular optical trap experiment (also applicable to magnetic tweezer experiments) as the ones described in Noom et al. [18], but we have added the theoretical possibility where the molecules carry twist. We consider two DNA molecules, each of total contour length $2l$ with $-l \leq s \leq l$, where s is the arc-length of the space curve describing each DNA filament. A pulling force F is applied coaxially to the center-line of the space curve describing the tails (straight portions) of the DNA filament. Since the trapping beads at the ends of each filament can be moved independently in space, the DNA molecules can be inter-wound by a number of rotations R around

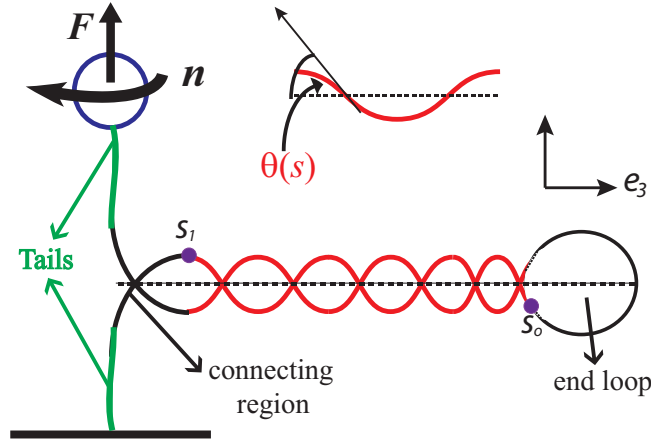


Figure 4.2: Sketch representing single molecule experiments where plectonemes are formed. A DNA molecule is fixed at one end, while the other end is subjected to a vertical pulling force F and twisted by a given number of turns n .

the horizontal (helical) axis \mathbf{e}_3 . Also each molecule can be rotated n turns by a twisting moment M_{ext} about its axis. Under these conditions DNA forms ply structures (braids in the case of nicked DNA molecules) once $R > 1/2$ [17].

It is clear from figure 4.1 that the straight tails are separated by a superhelical region composed of two identical helical curves. We will refer to these helical curves as general helices, with no need to have constant curvature and torsion. Considering variable curvature solutions brings up the possibility of finding more complex DNA structures because for two (or more) interwound helices there is a geometrical lock-up helical angle [19] – for angles larger than this lock-up value there are no helical solutions of the type depicted in figure 4.1. In fact, more complex structures have been observed in the Monte Carlo simulations of Charvin et al. [17] after lock up is reached.

Charvin et al. [17] used constant pitch helices reproduce the values of the extension as function of the number of rotations R from both experiments and MC simulations, but there is still a significant discrepancy in the values of the external moment M about the helical axis \mathbf{e}_3 and the critical number of rotations R at which the braided configuration makes a transition into a more complex DNA structure [17]. The values of the total initial extension z_o of the molecules and the radius r of the helices used in the simple geometrical model in Charvin et al. [17] are fitted to describe simulation and experiment. While the fitted values do reproduce qualitatively the braiding phenomena we note that the fitted theoretical radius in Charvin et al. [17] differs from the ones used in the hardcore interaction potential in their MC simulations. Consequently, we decided to use a variable pitch helix where the objective is to study how the external moment varies as a function of the applied rotation R and number of turns n . Note that our analysis can be easily simplified to account for the braids studied in Charvin et al. [17] by dropping the constraint on the linking number n such that the molecules carry no twist. We found that implementing a variable pitch helix, including a simple treatment of thermal fluctuation effects and configurational entropy, describes the MC simulations and experimental data very accurately while still using the same hard core potential radius r as in the MC simulations in Charvin et al. [17].

4.2 The Variational Problem

We follow a model similar to the one described in [21], where the DNA filament is considered to be an elastic rod. The details of the elastic rod model can be found in C.2. In our extended ply model the rod is made up of two regions:

- The tails are straight and aligned with the pulling force F . Due to thermal fluctuations the tails follow a writhed path. In the extended ply model we account for the effect of thermal fluctuations in a simple way by using l as an effective length $l = \epsilon L$, where L corresponds to the length with no thermal fluctuation and ϵ is the end to end distance of a worm like chain subjected to a pulling force F as given by [48]:

$$\epsilon \approx 1 - \frac{1}{2} \frac{k_B T}{\sqrt{F K_b}}, \quad (4.1)$$

where k_B the Boltzmann constant and T is the absolute temperature. A more accurate formula for the end to end distance of short-length filaments undergoing fluctuations is provided in [130].

- In the supercoiled region the position vector $\mathbf{r}_p(s)$ and the tangent vector $\mathbf{r}'_p(s)$ describe the superhelix. Note that each helix is itself a piece of double-stranded DNA. Using a cylindrical coordinate system and director frame Eq. (C.14) described in C.2, we have :

$$\mathbf{r}_p(s) = \chi r \mathbf{e}_r + z \mathbf{e}_3, \quad (4.2)$$

$$\mathbf{r}'_p(s) = \mathbf{d}_3 = \sin \theta \mathbf{e}_\psi + \cos \theta \mathbf{e}_3, \quad (4.3)$$

$$\psi' = \chi \frac{\sin \theta}{r}, \quad z' = \cos \theta, \quad 0 < \theta < \frac{\pi}{2},$$

where the chirality $\chi = \pm 1$ stands for the handedness of the helix; $\chi = 1$ for a right-handed helix and $\chi = -1$ for left handed helix [45]. The other filament of the supercoil is obtained by a rotation of π about the helical axis \mathbf{e}_3 . The supercoiled region is characterized by the helical radius r , which is assumed to be uniform, and the variable helical angle $\theta(s)$. Both r and θ may depend on the loading. The curvature $\kappa(s)$ is given by [76]:

$$\kappa^2 = (\theta')^2 + \frac{\sin^4 \theta}{r^2}. \quad (4.4)$$

By the arguments of conservation of torque about the body axis of an isotropic rod (see C.2), $M_3 = M_{ext}$, implying that the twist u_3 is constant in the helical region. Since we assume the tails to be straight the twist is equal in the tails and supercoiled region. The beads can be inter-wound R turns about the horizontal axis of the helices. This results in a conjugate moment M about the helical axis \mathbf{e}_3 .

Due to the vertical and horizontal lines of symmetry in the problem (see figure 4.1), we can simplify the analysis to the region of the curve contained in $0 \leq s \leq l$ and consider only one half of a DNA filament. The molecule contour length spent per tail is denoted by l_s and the contour length in the helical region is denoted by l_b . The sum of the length of the two regions is given by $l = l_b + l_s$. Our final objective is to compute the horizontal end to end extension $z(n, R)$ given by (see figure 4.1):

$$z(n, R) = \int_0^{l_b} (\cos \theta) ds + (\cos \alpha) l_s, \quad (4.5)$$

where α, θ, l_b and l_s are functions of n and R . The equilibrium configuration of the DNA is fully specified by the center-line, through the variables $r, \theta(s), M$ and M_3 . In what follows we compute

these parameters as a function of the loading and the arc-length s by minimizing the total energy of the system.

The experiments are performed under imposed end rotations R about the helical axis \mathbf{e}_3 and imposed number of turns n about the filament axis $\mathbf{r}'_p = \mathbf{d}_3$. Therefore, the energy minimization will be performed under the following constraints:

- The end rotation is given by (see C.1):

$$2\pi \left(R - \frac{1}{2} \right) = \chi \int_0^{l_b} \frac{\sin \theta}{r} ds, \quad (4.6)$$

where the $1/2$ appears because the inter-wound ($l_b > 0$) region starts for $R > 1/2$.

- The number of turns n imposed on the bead is equal to the excess link Lk_b of the DNA molecule in the helical region and the excess link Lk_s in the tails:

$$n = Lk_s + Lk_b \quad (4.7)$$

where the link Lk_b in the helical region corresponds to the classical partition into twist Tw and writhe Wr (see C.1):

$$Lk_b = \frac{M_3 l_b}{2\pi K_t} - \chi \int_0^{l_b} \frac{\sin 2\theta}{4\pi r} ds. \quad (4.8)$$

We assume that the force F is very large so that the link in the tails is entirely comprised of twist:

$$Lk_s = \frac{M_3 (l - l_s)}{2\pi K_t}, \quad (4.9)$$

In the experiments of Charvin et al. [17] the distance e is fixed. These facts together with the assumption that the tails are straight and carry only twist give the following geometrical constraint:

$$e = (l - l_b) \sin \alpha. \quad (4.10)$$

Note that the force enforcing the above constraint is $F \sin \alpha$ where α will be determined as part of the solution. This is consistent with experiments because the optical tweezers like those used by Charvin et al. [17] do not constrain the angle α .

4.3 Potential Energy of the System

It is convenient to express the total potential energy of the DNA as:

$$H = 2E_{coils} + 2E_{tails} - f(l - z(n, R)) \quad (4.11)$$

where $f = 2F \cos \alpha$ is the horizontal projection of the applied force, E_{tails} is the mechanical energy present in tail/straight region:

$$E_{tails} = \frac{M_3^2}{2K_t} l_s. \quad (4.12)$$

E_{coils} is the mechanical and internal energy present in the supercoiled region:

$$E_{coils} = \int_0^{l_b} \left(\frac{K_b}{2} \kappa^2 + \frac{M_3^2}{2K_t} + U^b(r, \theta) \right) ds, \quad (4.13)$$

where κ is given by Eq. (4.4) and $U^b(r, \theta)$ represents the free energy of entropic confinement per unit length of the strand in the supercoil and it may be written approximately as [20, 52, 84]:

$$U^b(r, \theta) = U_{conf}(r, \theta) = \frac{k_B T}{A^{1/3}} \left[\frac{c_p}{(p\pi)^{2/3}} + \frac{c_r}{r^{2/3}} \right], \quad (4.14)$$

where c_r and c_p are unknown constants fitted to experiment [84]. The term $2\pi p$ is the pitch of the helix and is given by $p = r \cot \theta$ and $A = K_b/(k_B T)$ is the persistence length of the fluctuating rod. The term $U^b(r, \theta)$ could also contain expressions describing the DNA-DNA electrostatic interactions but we choose to account for the electrostatic behavior with a hardcore radius as done in Charvin et al. [17], since we will compare our model with their experiments and MC simulations. So, the energy subject to constraints is:

$$\begin{aligned} H = & 2E_{braids} + 2E_{tails} - f(l - z(n)) + 2\pi M \left[R - \frac{1}{2} - \chi \int_0^{l_b} \frac{\sin(\theta)}{2\pi r} ds \right] \\ & + 2\pi \Lambda \left[2n - \frac{M_3}{\pi K_t} l + \chi \int_0^{l_b} \frac{\sin(2\theta)}{2\pi r} ds \right], \end{aligned} \quad (4.15)$$

where Λ is the Lagrange multiplier enforcing the linking number constraint Eq. (4.6) and M is the Lagrange multiplier (in fact, this is the external moment about the \mathbf{e}_3 axis) enforcing the constraint Eq. (4.7). For notation purposes we let:

$$H = \int_0^{l_b} I(\theta, \theta', s) ds + \hat{C}, \quad (4.16)$$

where

$$I(\theta, \theta', s) = K_b \left(\theta' + \frac{\sin^4 \theta}{r^2} \right) + 2U(r, \theta) - \chi M \frac{\sin \theta}{r} + \chi \Lambda \frac{\sin 2\theta}{r} + f(\cos \alpha - \cos \theta) \quad (4.17)$$

and \hat{C} is a constant independent of the arc-length s :

$$\hat{C} = \left[\frac{M_3^2}{K_t} + f(\cos \alpha - 1) \right] l + 2\pi M(R - 1/2) + 4\pi \Lambda n - 2 \frac{\Lambda M_3 l}{K_t} \quad (4.18)$$

4.3.1 Variational Formulation: Fixed F , R and n

We minimize the energy with respect to the independent variables M_3 and θ . The minimization with respect to twisting moment M_3 readily yields, as expected:

$$\Lambda = M_3. \quad (4.19)$$

Next we minimize with respect to θ :

$$\begin{aligned} \frac{\delta H}{\delta \theta} = \frac{\partial I}{\partial s} \frac{\partial I}{\partial \theta'} - \frac{\partial I}{\partial \theta} = 0 \\ 2K_b \left(\theta'' - 2 \frac{\sin^3 \theta \cos \theta}{r^2} \right) - 2 \frac{dU^b}{d\theta} - \chi M_3 \frac{\cos 2\theta}{r} + \chi M \frac{\cos \theta}{r} - f \sin \theta = 0 \end{aligned} \quad (4.20)$$

Note that:

$$\frac{dU^b(r, \theta)}{ds} = \frac{\partial U^b(r, \theta)}{\partial \theta} \theta'.$$

Therefore, we obtain from Eq. (4.20):

$$K_b \left(\theta'^2 - \frac{\sin^4 \theta}{r^2} \right) - 2U(r, \theta) + \chi M \frac{\sin \theta}{r} - \chi M_3 \frac{\sin 2\theta}{r} + f \cos \theta = D_1, \quad (4.21)$$

where D_1 is an integration constant. The equation given above for the variable pitch ply has been previously presented [70], but that study was restricted to fixed length plies with essential or natural boundary conditions. We approach the problem in a different manner by relaxing the fixed length constraint of the ply and solving it by making use of the theory of optimal variable end-points [127, 128].

4.3.2 Boundary Conditions for the Variational Problem

The first boundary condition is given by the symmetry requirement along the vertical axis at the center of the ply $s = 0$. The profile of θ has to be symmetric hence:

$$\theta'|_{s=0} = \gamma' = 0, \quad (4.22)$$

where $\gamma(s)$ is the curve where θ lies at $s = 0$.

On the other boundary $P_1 = (\alpha, s_1)$ is a variable end point [127]. Imagine a particle at the end of the helical region at $s = s_1$. For a prescribed force F , as rotations R are added to the system, s_1 and α change because the ply gets longer. Let $\alpha(s)$ be the curve where P_1 lies. The transversality condition for variable end points is [127]:

$$\alpha' \frac{\partial I}{\partial \theta'} - \left(\theta' \frac{\partial I}{\partial \theta'} - I \right) \Big|_{s=l_b} = 2K_b \alpha' \theta' - (D_1 - f \cos \alpha) \Big|_{s=l_b} = 0, \quad (4.23)$$

where we have made use of Eq. (4.17) and Eq. (4.21). The subscripts in the above expression denote partial differentiation with respect to that variable. $\alpha(s)$ is the curve that describes the Euler angle θ in the connecting region between the tails and the helices for $s \geq s_1 = l_b$. We enforce the continuity of the tangent vector $\mathbf{d}_3 = \mathbf{r}'(s)$ and the moment vector \mathbf{m} (equivalently the curvature κ^2) at the intersection point $s_1 = l_b$. Continuity of $\mathbf{d}_3 = \mathbf{r}'(s)$ implies $\theta(s)$ is continuous at $s = s_1$, while continuity of the moment implies that $\mathbf{m} \cdot \mathbf{e}_r = m_r = -K_b \theta'(s) = -K_b \alpha'(s)$ is continuous at $s = s_1$. Therefore, the second boundary condition in equation Eq. (4.23) is equivalent to:

$$2K_b \alpha'^2 - (D_1 - f \cos \alpha) = I \Big|_{s=l_b} = 0 \quad (4.24)$$

4.3.3 Solution

In this section we outline an approach to solve for the parameters of the braid under the condition in Charvin et al. [17]. We introduce a hardcore radius to account for the electrostatics and excluded volume effects of the braided DNA. This not only simplifies the calculations but allows comparison with the Monte Carlo simulations in Charvin et al. [17] in which a hardcore potential is used.

For the experimental set up in figure 4.1 we drop the constraint on n so that $M_3 = 0$ in the problem. Moreover, if the separation distance $2e$ is fixed, then for a tensile force F , as R increases the length of the straight portion l_s decreases. We also keep in mind that in the two bead set up of Charvin et al. [17] the horizontal component f is the controlled parameter rather than F . For fixed $2e$, it is clear that α monotonically increases as a function of R and that this relation can be inverted. So we solve the problem by letting α be the independent variable and obtain $R(\alpha)$ from Eq. (4.6).

For a given value of α , the iterative procedure to find the solution is:

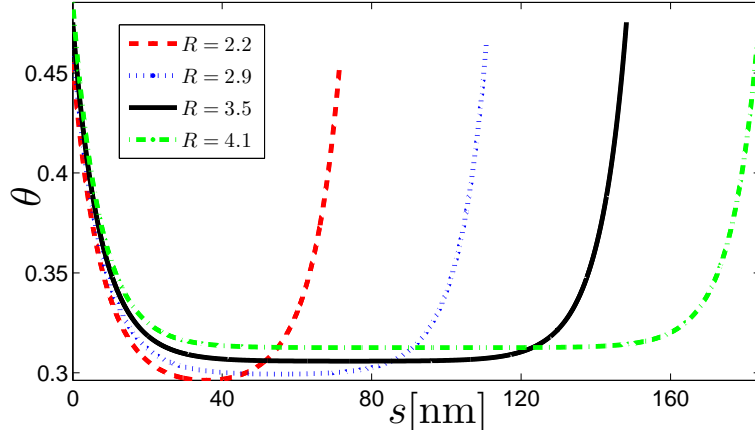


Figure 4.3: Profile of $\theta(s)$ as a function of s for nicked DNA (no twist) in a setup as shown in figure 4.1. In the calculations we have used $L = 1200\text{nm}$, $K_b = 180\text{pNm}$, $f = 2\text{pN}$, $2e = 0.36L$ and the hardcore radius $2r = 6\text{nm}$, as done in a set of MC simulations in [17]. $\theta(s)$ approaches the mid-section angle γ but diverges fast close to the end points s_1 and s_o . The plateau value of θ is a function of the prescribed R . We also computed \hat{R} using the mid-section approximation where we assume that the helical region has a constant $\theta = \gamma$ in the interval $-l_b < s < l_b$, while $\theta = \alpha$ only at the end points s_1 and s_o . The percentage error between R and \hat{R} decreases for longer braids. For $R = 4.1$, the percentage error is approximately 3%.

- Using Eq. (4.21) evaluated at $s = s_1 = l_b$ and Eq. (4.24) we can get an expression of $D_1 = D_1(\alpha, r, M)$ as function of α, r and M , where r and α are known and M is unknown.
- Guess an initial value M^i . Then, solve the differential equation Eq. (4.20).
- We know the limit point at $s = l_b$ to be $\theta = \alpha$ and using the two boundary conditions Eq. (4.22) and Eq. (4.24) we can find the value of the other limit point $\theta|_{s=0} = \gamma$ as a function of M^i :

$$\frac{\sin^4 \gamma}{r^2} + 2U(r, \gamma) - \chi M^i \frac{\sin \gamma}{r} - F \cos \gamma + D_1(\alpha, r, M^i) = 0,$$

- With the two boundary points, the length of the braided region l_b^i can be obtained by solving the differential equation. The iterative process repeats until the values of $l = l_b^i$ and $M = M^i$ satisfy constraint Eq. (4.10).
- Finally, R is obtained from Eq. (4.6).

We find that for l_b larger than a few nanometers ($\sim 20\text{nm}$) the midsection and most of the supercoiled region approaches the helical angle $\theta = \gamma$. The solution diverges from γ only near the end points where $\theta \rightarrow \alpha$. The problem can be approximated by simultaneously solving Eq. (4.20) and Eq. (4.21) at the midsection point $s = 0$ (where $\gamma' = 0$ and $\gamma'' = 0$), which yields $\theta \approx \gamma$ for $-l_b < s < l_b$ and $\theta = \alpha$ at the end points of the braided region. We call this approximation the mid-section γ solution. Therefore $M \approx M_\gamma$ where the subscript γ denotes the value corresponding to mid-section solution at $s = 0$. We found that using the approximation $\hat{R} \approx l_b \sin \gamma / (2\pi r)$ with $l_b \sim 50\text{nm}$ yields less than 5% error when compared to the full solution using the variable angle. The error decreases for longer braids reaching approximately 3% error for $l_b \sim 100\text{nm}$. This approximation does not preclude the idea that $\theta = \alpha > \gamma$ at $s = l_b$ and consequently the braid can reach lock-up well before γ is close to $\pi/4$. Note that the midsection γ solution is not the same as a constant helix solution where $\theta' = \theta'' = 0$ for all values of s . The constant solution would correspond to replacing D_1 by $f \cos \alpha$ in equation Eq. (4.21), but this substitution leads to a different solution for $\theta \neq \gamma$. For the parameters used in [17] there is no solution for θ if we set $\theta' = \theta'' = 0$.

Using the mid-section γ solution we compare our model to simulations and experiments in Charvin et al. [17]. The experiments were performed using 11kb \sim 3600nm double stranded DNA filaments at $T = 298\text{K}$ for different salt concentrations. The DNA was subjected to a pulling force $f = 2\text{pN}$, and the authors measured the persistence length to be $A = 44 \pm 4 \text{ nm}$. The spacing $2e$ between beads in the experiment was set to be $0.36L$. To see the effect of the bead spacing $2e$ and the electrostatics (hardcore diameter $2r$) the authors also performed MC simulations. In the simulations they used $L = 1200\text{nm}$ with $2e = 0.36L$ and $2e = 0.02L$ and performed calculations for diameters ranging from 4.8 to 6nm. The authors also provided a simple geometrical model with a constant helical angle θ that captures the behavior of DNA braiding, but the analytical model required much larger diameters $\sim 8.8\text{nm}$ to provide good agreement with the experimental and simulation data. As noted in Charvin et al. [17], the MC simulations using $2r = 6\text{nm}$ accurately match the experimental data using a 100 mM monovalent salt solution, while the simulations using $2r = 4.8\text{nm}$ match experiment using a 100 mM PB + 5mM Mg^+ salt solution. Since we are using a hardcore potential radius to describe electrostatics we compare our model to their MC simulations. This also allows us to compare our model with their experiments.

Figures 4.4-4.6 show the comparison between our analytical model and the MC simulations in Charvin et al. [17]. There is quantitative agreement in the end to end extension z/L as a function of the catenation number $Ca \approx 3.4\hat{R}/L$ (see figure 4.4). The MC simulations show that braids are formed up to a critical catenation number $Ca_c \sim 0.04 - 0.05$ (for $2e = 0.036L$) and $Ca_c \sim 0.075$ (for $2e = 0.02L$). Our calculations are valid up to the point where constraint Eq. (4.10) holds for $\alpha \leq \pi/4$ (smaller than the lock-up angle). At this point we expect that there would be a transition to more complex DNA structures including supercoils of inter-wound braids as seen in the MC simulations [17]. Ca_c ranges in our calculation from $Ca_c = 0.031$ ($2r = 6\text{nm}$) up to $Ca_c = 0.045$ ($2r = 4.2\text{nm}$) when $2e = 0.36L$, and $Ca_c = 0.072$ when $2e = 0.02L$. For both spacings $2e$, our results show good agreement with the simulations. Note in figure 4.4 that in the case $2r = 4.2\text{nm}$ our $Ca_c = 0.045$ prediction is the same as the transition point obtained from the MC simulation (see figure 6 in Charvin et al. [17]). Figure 4.5 shows the evolution of M as a function of Ca . The predictions of our model are in good agreement with the simulations up to our predicted Ca_c point. But, our model shows a non-linear dependence of M on Ca (specially for the case $2e = 0.02$, see inset in figure 4.5), while the dependence is linear in Charvin et al. [17]. We found that as we increased f , the dependence of M on Ca approached linear behavior. Figure 4.6 shows the behavior of the helical angle as a function of Ca . We recall that $\theta \approx \gamma$ is the mid-section angle of the braid, while α is the opening helical angle at the end points. Charvin et al. [17] point out that the value of θ varies little with number of rotations in their simulations. But, our model shows that both θ and α vary as functions of Ca . Both these parameters affect the energetics of the system and could explain the discrepancy in the predicted M moments.

Experiments as the one depicted in figure 4.1 including twist are not available in the literature at the moment but could be realized in future. Therefore, we use our model to make some predictions that can be tested. Using the mid-section γ solution, in figure 4.7 we show the results of DNA subjected to the linking number constraint Eq. (4.7) when the filaments carry twist. Introducing the linking number constraint through n does not affect the helical angle θ and the end to end distance z/L significantly (not shown), but it does change the values of the external moment M and twisting moment M_3 .

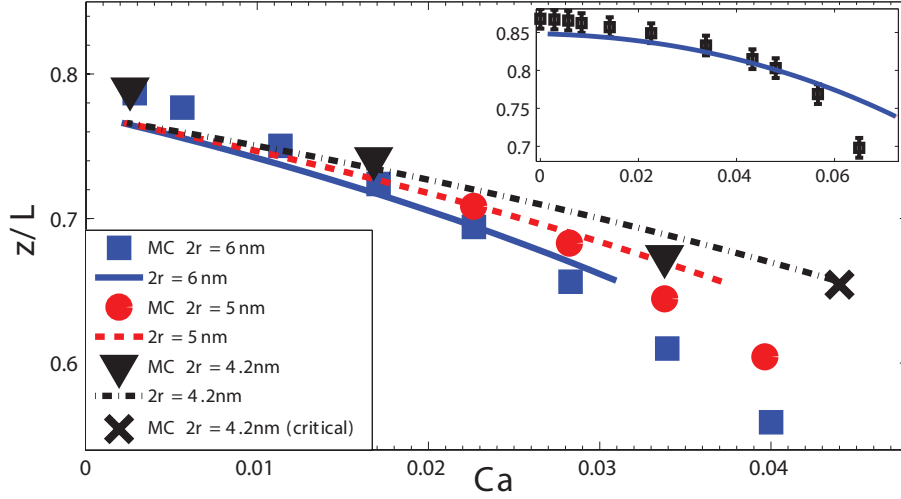


Figure 4.4: Normalized end to end extension z/L as a function of catenation number $Ca \approx 3.4\hat{R}/L$ for nicked DNA (no twist). In the calculations we have used the same parameters as in figure 4.3. The lines show the solution up to the point where constraint Eq. (4.10) holds for $\alpha \leq \pi/4$. The opening angle α in the helical region has to be smaller than or equal to the lock-up angle [19]. The end of each line depicts the critical catenation number Ca_c at which the DNA is believed to make a transition into more complex structures (supercoiling of braids) [17]. The X marker shows $Ca_c \approx 0.045$ from MC simulations using $2r = 4.2\text{nm}$. The inset shows z/L as a function of Ca using $2e = 0.02L$ and $r = 6\text{nm}$ and the markers correspond to MC data [17].

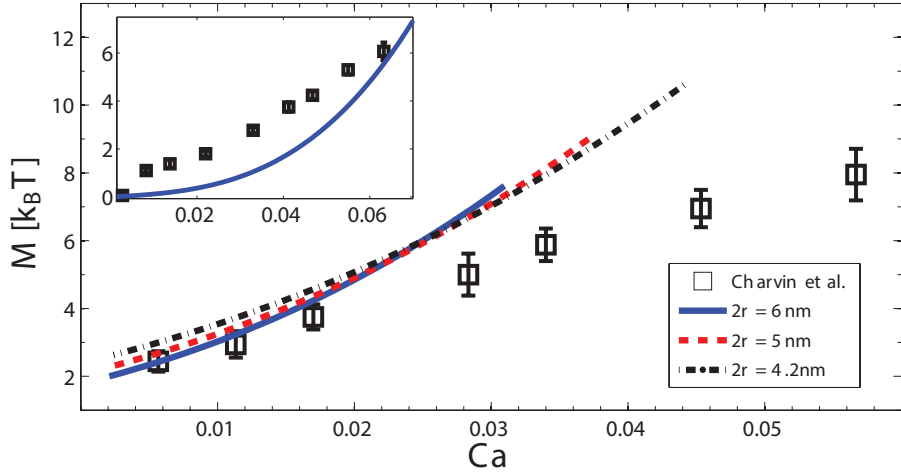


Figure 4.5: External moment M as a function of catenation number Ca for nicked DNA (no twist). In these calculations we have used the same parameters as in figure 4.4. The inset shows results using $2e = 0.02L$ and $r = 6\text{nm}$ and the markers correspond to MC data [17].

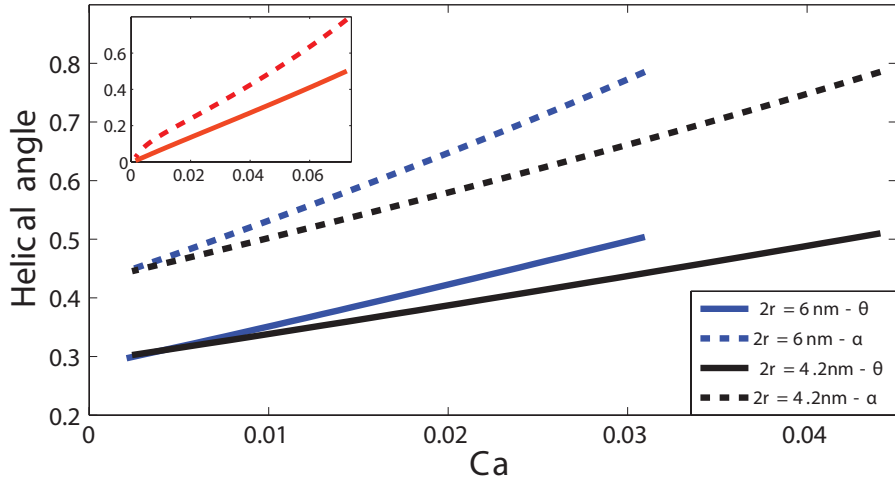


Figure 4.6: Helical angle as a function of catenation number Ca for nicked DNA (no twist). In these calculations we have used the same parameters as in figure 4.4. Here $\theta \approx \gamma$, the mid-section angle of the braid, while α is the opening helical angle at the end points. The inset shows results using $2e = 0.02L$ and $r = 6\text{nm}$.

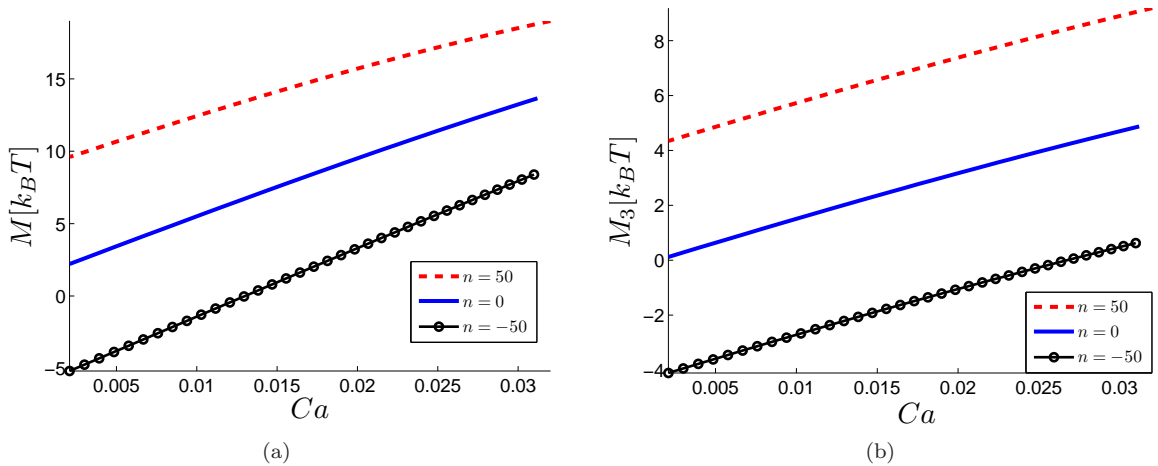


Figure 4.7: (a) External moment M as a function of catenation number Ca for twisted DNA. In these calculations we have used the same parameters as in figure 4.4. (b) Twisting moment M_3 as a function of catenation number Ca for twisted DNA. For a given n adding external rotation R to the system is equivalent to allocating part of the link into writhe. For positive R the induced writhe is negative as noted from Eq. (4.6) and Eq. (4.7). Hence, for fixed n , the twist increases to counteract the effect on the writhe effect, which can be seen as an increase in the value of M_3 as a function of Ca .

4.4 General Description of Model 2 : Plectonemes

As in section 4.1, we continue with our model of DNA as an elastic rod under tension and torsion, but this time in the configuration depicted in figure 4.2. We consider a DNA molecule of total contour length $2l$ with $-l < s < l$ where s is the arc-length of the space curve describing the DNA filament. The molecule is fixed at $s = -l$, while the other end $s = l$ is attached to an angular trapping instrument (optical or magnetic bead). A vertical pulling force F is applied coaxially to the center-line of the space curve describing the tails of the DNA filament and the DNA molecule is rotated n turns by a twisting moment M_{ext} . Under these conditions DNA forms plectonemes

once $n > n_{critical}(F)$. It is clear from figure 4.2 that two twisted, straight tails are separated by a plectonemic region composed of two identical helical curves. Intuitively, we expect the curvature to be greater in the region closer to the end loop than the region closer to the tails since the bending moment is larger at the end loop. This is part of the motivation for considering helical solutions with variable pitch. Previous studies using a constant helix approximation reproduce the values of the slope in the rotation vs. extension graphs from both experiments and MC simulations very accurately, but there is a significant discrepancy in the values of the external torque obtained from direct experimental methods [12], indirect experimental methods [11], MC simulations [5] and theory [4, 21, 45, 73]. So here we generalize to a variable pitch helix with the objective of studying how the external moment values influence the helical solution. We also present an analysis that shows that if the external torque is assumed to be independent of the number of turns n then the system must yield a constant angle solution.

4.4.1 The Variational Problem

In this section we follow more closely the model described in [21], where the DNA filament is considered to be an elastic rod undergoing thermal fluctuations. The DNA in the experiments is made up of three regions:

- In the linear region the tails are nearly straight and aligned with the vertical axis. The tails are not completely straight and the center line follows a writhed path due to thermal fluctuations of the DNA molecule. Here we follow the analysis of fluctuating polymers subjected to tension and twist in the straight regime which was carried out in detail by Moroz and Nelson [40, 56] who give expressions for the twist and writhe.
- The plectonemic region is exactly the same as described in the extended ply model in section 4.2. The external moment M_{ext} applied in the upper tail of the DNA molecule is equivalent to a total moment M_3 about $\mathbf{r}'_p = \mathbf{d}_3$ at the beginning of the plectonemic region. By the arguments of conservation of torque about the body axis of an isotropic rod in C.2, $M_3 = M_{ext}$, implying that the twist u_3 is constant in the helical region ¹. One consequence of the use of the expressions given by Moroz and Nelson [40, 56] is that the twist u_3 in the tails is different from that in the plectoneme even though the twisting moment is the same.
- In the end of the plectonemic region there is a loop. The end loop is formed in a transition from the straight configuration into the plectonemic configuration. In order to model the loop we propose an approximation based on the localizing solution of the rod [78, 79, 131].

The molecule contour length spent per tail is denoted by l_t , the contour length in the loop is denoted by $2l_o$ and the contour length per helix is denoted by l_p . The sum of the length of all regions is given by $2l = 2l_p + 2l_t + 2l_o$. The equilibrium configuration of the rod is fully specified by the center-line through the variables r , $\theta(s)$ and M_3 . In what follows we compute these parameters as a function of the loading (pulling force, F , and the number of turns, n) and the arc-length s by minimizing the total energy of the system.

Due to the symmetry of the problem we can simplify the analysis to the region of the curve contained in $0 \leq s \leq l$. The experiments are performed under imposed end rotations, so the energy minimization will be performed under the constraint that the number of turns n imposed on the bead is equal to the excess link Lk_p of the DNA molecule in the helical region, the excess link Lk_t

¹At the transition point between the straight state and the plectonemic state there is a jump in the external torque. We define $M_{critical}$ as the twisting moment in the straight configuration right before the transition, M_3 as the twisting moment when plectonemes (helices) are present, and $\delta M = M_{critical} - M_3$ as the jump in the twisting moment at the transition. We use the notation M_{ext} in section 4.4.1 for the external torque. The equations describing the DNA tails can be used for the plectonemic state and the straight state by replacing M_{ext} with either M_3 or $M_{critical}$ respectively.

in the tails and the excess link Lk_o in the loop ²:

$$n = Lk_p + Lk_t + Lk_o, \quad (4.25)$$

where the link Lk_p in the helical region corresponds to the classical partition into twist Tw and writhe Wr (see C.1):

$$Lk_p = \frac{M_{ext}l_p}{2\pi K_t} - \chi \int_0^{l_p} \frac{\sin 2\theta}{4\pi r} ds. \quad (4.26)$$

At this point we note that clockwise rotations n about the \mathbf{e}_1 axis, corresponding to a positive external moment M_{ext} as shown in figure 4.2, generate a left-handed helix with $\chi = -1$, while a negative external moment generates a right-handed helix with $\chi = 1$. We also note that in the presence of thermal fluctuations there is a writhe contribution from the tails which can be accounted for by using the results of [56]:

$$Lk_t = \frac{M_{ext}(l - l_p - l_o)}{2\pi} \left(\frac{1}{K_t} + \frac{1}{4K_b K} \right) + O(K^{-3}), \quad (4.27)$$

where,

$$K = \frac{\sqrt{K_b F - M_{ext}^2/4}}{k_B T}. \quad (4.28)$$

The link in the end loop can be approximated as (see C.1):

$$Lk_o = \frac{M_{ext}(l_o)}{2\pi K_t} + Wr_o, \quad (4.29)$$

where $2Wr_o \approx 1$ is the writhe present in the loop.

4.4.2 Potential Energy of the System

Following the treatment in [21], it is convenient to express the total potential energy of the DNA filament as:

$$V = \int_0^l L(s, \theta, \theta', q_i) ds = V_{tails} + V_{loop} + V_{helices}, \quad (4.30)$$

where q_i are variables like M_3, r, \dots independent of s . The free energy of the tail region under imposed end rotations Lk_t is given by [21]:

$$V_{tails} = 2\pi Lk_t M_{ext} + \left(-F - \frac{M_{ext}^2}{2K_t} + G_{flu}^* \right) l_t, \quad (4.31)$$

where the last term is a correction due to thermal fluctuations:

$$G_{flu}^* = \frac{(k_B T)^2}{K_b} K \left(1 - \frac{1}{4K} - \frac{1}{64K^2} \right) + O(K^{-3}), \quad (4.32)$$

²Since the analysis is simplified to $0 \leq s \leq l$ the given expressions for the energy and number of turns are half of the values associated with the complete problem.

where K is given by Eq. (4.28). The extension with thermal fluctuations taken into account is given by $-\frac{\partial V_{tails}}{\partial F} = \rho l_t$ ³, where

$$\rho = 1 - \frac{1}{2} \frac{1}{\sqrt{\frac{K_b F}{k_B^2 T^2} - \frac{M_{ext}^2}{4k_B^2 T^2} - \frac{1}{32}}} + \frac{K_b k_B T}{l_t (K_b F - \frac{M_{ext}^2}{4})}. \quad (4.33)$$

The slope of the rotation-extension curve after the formation of plectonemes can be obtained from constraint Eq. (4.25) together with the expression Eq. (4.33). The extension of the filament is given as $\Delta z = \rho(l - l_o - l_p)$ where we note that l is constant, l_o is approximately constant and ρ is independent of n .

The free energy of the loop will be approximated under the assumption that the bending and twisting energy decouple (see [21]). The twisting moment M_{ext} is a constant along the molecule and therefore the twist energy per unit length is a constant too. The expression for the bending energy per unit length E_{o-bend} and the length of the loop l_o are obtained from to the expressions given by [79] in the absence of twist.

$$V_{loop} = \left(\frac{M_{ext}^2}{2K_t} + E_{o-bend} \right) l_o, \quad (4.34)$$

where

$$2l_o = 4\sqrt{\frac{K_b}{F}}, \quad E_{o-bend} = F.$$

The free energy of the plectonemic region can be divided into elastic energy $V_{el}^{helices}$ and the energy due to internal interaction $V_{int}^{helices}$. The elastic energy is given by:

$$V_{el}^{helices} = \int_0^{l_p} \left(\frac{K_b}{2} \kappa^2(s) + \frac{M_{ext}^2}{2K_t} \right) ds. \quad (4.35)$$

Equation Eq. (4.35) captures the elastic behavior of the DNA in response to the applied loadings. The electrostatic and entropic interactions present in the plectonemic region (only along l_p) will be described in more detail later. Here we just introduce a general expression to account for both of them:

$$V_{int}^{helices} = \int_0^{l_p} U(r, \theta, x_i) ds, \quad (4.36)$$

where x_i represents any auxiliary parameters or internal variables that may appear in the free energy of the system depending on the model picked to described the electrostatic and entropic parts of the energy.

The potential energy can be written by separating the terms that contribute along l , and the ones that only contribute along l_p and l_o . We introduce a Lagrange multiplier λ to account for the constraint Eq. (4.25) in the potential energy:

$$\begin{aligned} V &= \int_0^{l_p} \left(\frac{K_b}{2} \kappa^2(s) + F + U(r, \theta) - G_{flu}^* - \frac{M_{ext}^2}{4K_b K} \right) ds \\ &+ \left(\frac{M_{ext}^2}{2K_t} - F + G_{flu}^* + \frac{M_{ext}^2}{4K_b K} \right) l \\ &+ \left(E_{o-bend} + F - G_{flu}^* - \frac{M_{ext}^2}{4K_b K} \right) l_o + \lambda (n - (Lk_t + Lk_o + Lk_p)), \end{aligned} \quad (4.37)$$

³The given formula for ρ includes corrections as detailed in [40].

where $\kappa(s)$ is given from Eq. (4.4). Let us define:

$$\begin{aligned}
I(s, \theta, \theta', M_3, r) &= \frac{K_b}{2} \left(\theta'^2 + \frac{\sin^4 \theta}{r^2} \right) + F + U(r, \theta) - G_{flu}^* \\
&- \frac{M_{ext}^2}{4K_b K} + \frac{\lambda}{2\pi} \left(\frac{M_{ext}}{4K_b K} + \chi \frac{\sin 2\theta}{2r} \right).
\end{aligned} \tag{4.38}$$

Then the final expression for the potential energy of the system subject to the constraint Eq. (4.25) is:

$$\begin{aligned}
V &= \int_0^{l_p} I(s, \theta, \theta', r, M_3) ds + \left(\frac{M_{ext}^2}{2K_t} - F + G_{flu}^* + \frac{M_{ext}^2}{4K_b K} \right) l \\
&+ \left(E_{o-bend} + F - G_{flu}^* - \frac{M_{ext}^2}{4K_b K} \right) l_o \\
&+ \lambda \left[n - \frac{M_{ext} l}{2\pi} \left(\frac{1}{K_t} + \frac{1}{4K_b K} \right) + \frac{M_{ext} l_o}{8\pi K_b K} - W r_o \right].
\end{aligned} \tag{4.39}$$

4.4.3 Internal Energy: Entropy and Electrostatics

In section 4.4.2 we introduced the term $U(r, \theta, x_i)$ as a general expression to account for the internal interactions and configurational entropy cost in the plectonemic region. We split the internal energy of the plectoneme $U(r, \theta, x_i)$ into the configurational entropy cost contribution $U_{conf}(r, \theta, x_i)$ and the purely electrostatic interactions between the charged helices in ionic solution $U_{el}(r, \theta, x_i)$, such that $U = U_{conf} + U_{el}$. The free energy of entropic confinement per unit length of the strand in the plectonemic supercoil is given by expression Eq. (4.14). As mentioned before, the aim of this work is not to validate the theoretical model for plectonemes with experimental data, but rather study the more general case of variable pitch helices. DNA-DNA interactions and electrostatics are still not well understood and there are several models in the literature that have been used to capture the behavior of supercoiled DNA [4, 5, 21, 22]. The specific electrostatic model should be picked according to the type of solution which may contain monovalent or multivalent ions. When appropriate, we will point out the electrostatic model that we choose to make the calculations.

4.4.4 Variational Formulation: Fixed F and n

Once the DNA has transitioned from the straight configuration into the plectonemic configuration, we will define $M_{ext} = M_3$ as the external moment present in the molecule. We minimize the energy with respect to the independent variables r , M_3 and θ :

$$\left\{ \frac{\delta V}{\delta M_3}, \frac{\delta V}{\delta r}, \frac{\delta V}{\delta \theta} \right\} = 0$$

Minimizing with respect to M_3 trivially recovers the result:

$$\lambda = 2\pi M_3 + O(K^{-3}). \tag{4.40}$$

Minimization with respect to the radius r yields:

$$\int_0^{l_p} \left(\frac{-K_b \sin^4 \theta}{r^3} + \frac{\partial U(r, \theta)}{\partial r} - \chi M_3 \frac{\sin 2\theta}{2r^2} \right) ds = 0. \tag{4.41}$$

Finally, performing the variation with respect θ gives the following differential equation:

$$\frac{\delta V}{\delta \theta} = \frac{\partial I}{\partial s} \frac{\partial I}{\partial \theta'} - \frac{\partial I}{\partial \theta} = 0$$

$$K_b \left(\theta'' - \frac{2 \sin^3 \theta \cos \theta}{r^2} \right) - \frac{\partial U(r, \theta)}{\partial \theta} - \chi M_3 \left(\frac{\cos 2\theta}{r} \right) = 0. \quad (4.42)$$

Note that:

$$\frac{dU(r, \theta)}{ds} = \frac{\partial U(r, \theta)}{\partial \theta} \theta' + \frac{\partial U(r, \theta)}{\partial r} r'.$$

But $r' = 0$, so from Eq. (4.42) we get:

$$\frac{K_b}{2} \left(\theta'^2 - \frac{\sin^4 \theta}{r^2} \right) - U(r, \theta) - \chi \frac{M_3 \sin 2\theta}{2r} = C_1. \quad (4.43)$$

The results obtained in this section for the plectonemic state of the DNA molecule could be obtained also under the assumption that $l_o \ll l$, when the loop size can be neglected in comparison to the length of the tails and plectoneme.

4.4.5 Boundary Conditions for the Variational Problem

Using variational principles, as previously done for the extended ply model in section 4.3.2, we will consider the boundary conditions for equation Eq. (4.42) that arise naturally from the transversality conditions [127]. So far, for convenience in notation, we have been writing all the integrals in the form $\int_0^{l_p} (...) ds$. In a more rigorous and general approach, the lower bound of integration corresponds to point $P_o = (\theta_o, s_o)$ on the $\theta - s$ plane, previously associated with $s = 0$, and the upper bound corresponds to point $P_1 = (\theta_1, s_1)$, previously associated with l_p . Therefore, $l_p = s_1(s) - s_o(s)$, such that the length variations of the end loop are included in the total length of the plectoneme (distance between points P_o and P_1). In the process of forming a plectoneme in rotation-extension experiments, before a helix is formed the filament undergoes a dynamic transition and forms an end loop. In the light of these calculations, it is easy to see why point P_1 is a variable end point. Imagine a particle at the end of the plectonemic region at $s = s_1$. For a prescribed force F , as rotations are added to the system, s_1 changes because the plectoneme gets longer. Similarly, for a prescribed number of rotations n , as F increases the helical region will get shorter. On the other hand, as more turns are added to the system, qualitatively the curvature at point P_o becomes tighter and the end loop becomes smaller, so that s_o is also changing. So, we say that P_o and P_1 are, in general, variable end points. In practice, a model based on the assumption that the length variation of the end loop is negligible compared to the total variation of s_1 gives good agreement with experimental data on jumps in the extension at the transition [21]. Therefore, we are justified in making the approximation $l_p = s_1$, $s_o = 0$ and l_o is independent of the number of turns. Now, if l_o is fixed then point P_o has a natural boundary condition (only fixed position but θ is allowed to vary). Let $\Omega(s) = s - s_o = 0$ be the curve where P_o lies and let $\zeta(s)$ be the curve where P_1 lies. The transversality conditions for points P_o and P_1 are [127]:

$$\left. \frac{\partial I}{\partial \theta'} \right|_{s=0} = 0 \rightarrow \theta'|_{s=0} = \alpha' = 0, \quad (4.44)$$

$$\zeta' \left. \frac{\partial I}{\partial \theta'} - \left(\theta' \frac{\partial I}{\partial \theta'} - I \right) \right|_{s=l_p} = K_b \zeta' \theta' - (C_1 - F + G_{flu}^*) \Big|_{s=l_p} = 0, \quad (4.45)$$

where we have made use of Eq. (4.38) and Eq. (4.43). The subscripts in the above expression denote partial differentiation with respect to that variable.

We point out that the transversality conditions are in fact conditions on the curves where the end

points lie such that the variational problem is optimized [128]. The form of the second transversality condition Eq. (4.45) can be derived by solving the variational problem for a fixed end point $s_1 = l_p$, such that $\theta^*(s; l_p)$ is the solution for each value of l_p . A minimization with respect to l_p results in Eq. (4.45) as long as the integrand $I(s, \theta^*, \theta^{*'}, M_3, r)$ is not an explicit function of l_p ⁴. Therefore the solution θ^* to the variational problem satisfying Eq. (4.45) in fact satisfies:

$$\frac{\partial V}{\partial l_p} = \frac{\partial G}{\partial l_p} = 0 \quad (4.46)$$

where we use

$$G = \int_0^{l_p} I(\theta^*(s), \theta^{*'}(s), M_3, r; s) ds \quad (4.47)$$

to simplify notation in later equations.

$\zeta(s)$ corresponds to the curve that describes the Euler angle θ in the connecting region between the tails and the helices for $s \geq s_1$. It can be obtained following an analysis similar to [71, 72], where a non-planar localizing solution is patched to a helical solution to get a plectoneme. We enforce the continuity of the tangent vector $\mathbf{d}_3 = \mathbf{r}'(s)$ and the moment vector \mathbf{m} (equivalently the curvature κ^2) at the intersection point $s_1 = l_p$.⁵ Continuity of $\mathbf{d}_3 = \mathbf{r}'(s)$ implies $\theta(s)$ is continuous at $s = s_1$, while continuity of the moment implies that $\mathbf{m} \cdot \mathbf{e}_r = m_r = -K_b \theta'(s) = -K_b \zeta'(s)$ is continuous at $s = s_1 = l_p$. The latter means that the second boundary condition in equation Eq. (4.45) can be rewritten as⁶:

$$K_b(\zeta'(l_p))^2 - (C_1 - F + G_{flu}^*) = I|_{s=l_p} = 0. \quad (4.48)$$

4.4.6 Constant pitch solutions

Going back to the variational problem, we see that the solution to the profile of the helical angle described by the differential equation Eq. (4.42) is given by θ^* , M_3 , r and the boundary conditions Eq. (4.44) and Eq. (4.45). In general, the integrand $I(s, \theta^*, \theta^{*'}, M_3, r)$ can be a function of both n and F through $r = r(n, F)$ and $M_3 = M_3(n, F)$ which can both be functions of the controlled variables. So, the free energy is of the form $V(F, n) = V[F, n, l_p(F, n), r(F, n), M_3(F, n)]$:

$$\begin{aligned} V(F, n) &= \int_0^{l_p} I(s, \theta^*, \theta^{*'}, M_3, r) ds + 2\pi n M_3 + \left(-\frac{M_3^2}{2K_t} - F + G_{flu}^* \right) l \\ &+ (E_{o-bend} + F - G_{flu}^*) l_o - 2\pi M_3 W r_o. \end{aligned} \quad (4.49)$$

Thermodynamic derivatives yield:

$$\left. \frac{\partial V}{\partial n} \right|_F = \partial_n V + (\partial_r V) \left. \frac{\partial r}{\partial n} \right|_F + (\partial_{M_3} V) \left. \frac{\partial M_3}{\partial n} \right|_F + (\partial_{l_p} V) \left. \frac{\partial l_p}{\partial n} \right|_F = 2\pi M_3 \quad (4.50)$$

$$\left. \frac{\partial V}{\partial F} \right|_n = \partial_F V + (\partial_r V) \left. \frac{\partial r}{\partial F} \right|_n + (\partial_{M_3} V) \left. \frac{\partial M_3}{\partial F} \right|_n + (\partial_{l_p} V) \left. \frac{\partial l_p}{\partial F} \right|_n = -\rho l_t \quad (4.51)$$

where $|_x$ means holding x fixed and $\partial_n x$ represents explicit differentiation. Since $\partial_n G = 0$, $(\partial_r G) = (\partial_r V) = 0$ (from optimizing the energy with respect to r in Eq. (4.41)) and $(\partial_{l_p} G) = (\partial_{l_p} V) = 0$

⁴Note that M_3, r and l_p are independent variables.

⁵Our analysis is valid if there are no point forces and moments at $s = s_1$, which is justified under the assumption that there is no physical contact of the DNA strands due to the electrostatic interactions which lead to high repulsion forces at the length scale of the problem.

⁶If we neglect entirely the presence of the connecting region, which is a common assumption in the literature [21, 55, 73], then $m_r = -K_b \theta' = 0$ at $s_1 = l_p$ and $C_1 = F$. This assumption together with the assumption that l_o is fixed will lead to a uniform helix.

(from Eq. (4.46)), where G is given by Eq. (4.47), we obtain:

$$\left. \frac{\partial G}{\partial n} \right|_F = (\partial_{M_3} G) \left. \frac{\partial M_3}{\partial n} \right|_F \quad (4.52)$$

Now we will examine some connections of the above with the constant helical angle solutions and the trends in the experimental data. Experimental data and MC simulations [5, 10, 12, 27] show evidence that the external moment $M_{ext} = M_3$ becomes approximately a constant independent of the number of turns in the presence of plectonemes. So, let us assume that $M_3 = M_3(F)$ is independent of n . Now let $l_p^i(n_i)$ be the length of the supercoiled region at the beginning of the transition for $n_{critical} = n_i$ number of turns. Note l_p^i is not necessarily zero since there is a dynamic jump from the straight configuration to the supercoiled-straight coexistence where an end loop is formed [10, 12]. During this dynamic process there is a discontinuity in the measured external moment M_{ext} . Let $l_p^f(n_p) = l - l_o$ be the length of the supercoiled region when all the DNA is a plectoneme with no tails and $n = n_p$. For notation purposes we will let $\hat{n} \in [n_i, n_p]$ such that $l_p(\hat{n}) \in [l_p^i, l_p^f]$. The assumption that $M_3(F)$ is independent of the number of turns is the signature of a first order phase transition and implies from equation Eq. (4.52), $G(n_i) = G(n_p) = G(\hat{n})$, so that:

$$\begin{aligned} \int_0^{l_p^i(n_i)} I(s, \theta^*, \theta^{*'}, M_3, r(n_i); n_i) ds &= \int_0^{l_p^f(n_p)} I(s, \theta^*, \theta^{*'}, M_3, r(n_p); n_p) ds \\ &= \int_0^{l_p(\hat{n})} I(s, \theta^*, \theta^{*'}, M_3, r(\hat{n}); \hat{n}) ds. \end{aligned} \quad (4.53)$$

Making use of equation Eq. (4.38), Eq. (4.43) and Eq. (4.48) we can rewrite condition Eq. (4.53) as:

$$\int_0^{l_p^i(n_i)} [\theta^{*'}(s; n_i)]^2 - [\theta^{*'}(l_p^i; n_i)]^2 ds = \int_0^{l_p(\hat{n})} [\theta^{*'}(s; \hat{n})]^2 - [\theta^{*'}(l_p; \hat{n})]^2 ds. \quad (4.54)$$

Since $[\theta^{*'}(l_p; \hat{n})]^2 - [\theta^{*'}(s; \hat{n})]^2 \geq 0$ (see figure 4.9) and $l_p(\hat{n})$ is an increasing function, ensuring equality in condition Eq. (4.54) requires that the integrand decrease as \hat{n} increases. Hence, as $l_p \rightarrow \infty$, $[\theta^{*'}(s; \hat{n})]^2 \rightarrow [\theta^{*'}(l_p; \hat{n})]^2$ (for all values of s). The latter means that as the number of turns increases, the profile of $\theta^{*'}(s; \hat{n})$ flattens and becomes a constant value. Moreover, we required from the first boundary condition Eq. (4.44) that $[\theta^{*'}(s_o; n_i)] = \alpha' = 0$, so as \hat{n} increases, $|\theta^{*'}(l_p; \hat{n})|$ must decrease with a lower bound $\theta^{*'}(l_p; \hat{n}) \rightarrow \alpha' = 0$. At this point it is easy to see that for large to moderate values of \hat{n} (long DNA filaments in supercoiled state) $\theta^{*'} \sim 0$. We can take the analysis a step further. Notice from Eq. (4.48) that in order to satisfy condition Eq. (4.54) (for a non-constant helical angle), as \hat{n} decreases C_1 must increase. Now, if M_3 and r are both independent of n and $\theta'|_{s=0} = 0$, then the profile of θ^* given by equation Eq. (4.43) is only a function of C_1 . Now start from the case where l_p can be extended to infinity such that $\theta^{*'}(l_p; \hat{n}) = \theta^{*'}(s; \hat{n}) = 0$. From condition Eq. (4.48) we have that $C_1 = F - G_{flu}^*$ and in order to have this constant solution for a given \hat{n} , $\theta'' = \theta' = 0$ for all s , which means that equation Eq. (4.43) has only one solution for the helical angle, $\theta = \theta_c$. For shorter l_p (smaller \hat{n}) we must then have that $C_1 \geq F - G_{flu}^*$ to satisfy condition Eq. (4.54). If we let $C_1 > F - G_{flu}^*$ such that $\theta^{*'}(l_p; \hat{n}) \neq 0$ (variable pitch profile), then equation Eq. (4.43) has no value of θ for which $\theta' = 0$. This implies that if $C_1 > F - G_{flu}^*$ then the first boundary condition Eq. (4.44) can not be satisfied (See figure 4.10). Therefore, we have shown that for a system where M_3 and r are independent of n , we recover the uniform helix equations previously used in the literature [4, 21, 73] as the only possible solutions. The free energy per unit length is then given by:

$$u_{total} = \frac{l_p}{l} u_{helices} + \frac{l_t}{l} u_{tails} + \frac{l_o}{l} u_{loop}$$

which is equivalent to the double tangent construction proposed in the phase transition model of Marko [55] and Marko and Siggia [52]. Figure 4.8 shows a sample phase diagram based on the idea that the external moment M_3 is a constant independent of n . It illustrates the preferred DNA state as a function of the controlled variables F and n for given ionic conditions.

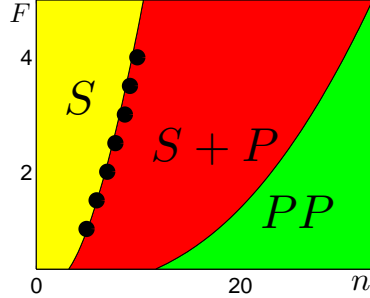


Figure 4.8: S stands for a completely straight configuration. $S + P$ stands for the configuration depicted in figure 4.2, where straight tails coexist with a helical region. PP represents the state when the entire DNA molecule has been converted into a superhelix. The scenario depicted represent a monovalent ionic solution of 0.1M NaCl. Black circles correspond to the 0.1M experimental data points from [5, 10]. The transition from $S \rightarrow S + P$ is accurately predicted by the model. We have used for the DNA-DNA electrostatic interaction the model provided by Ubbink and Odijk [20]. For details on the calculations we refer the reader to [21, 22].

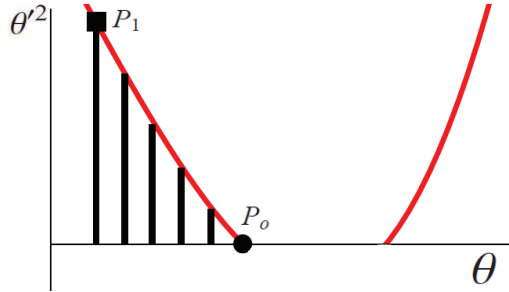


Figure 4.9: Profile of θ'^2 as a function of θ . The value of $[\theta'^*(l_p; \hat{n})]^2$ at P_1 is the maximum value of θ'^2 . The convexity of θ'^2 as a function θ is shown in C.3.

4.4.7 Variable pitch solutions: External moment $M_{ext} = M_3$ depends on n

So far in the description of the problem the end point condition at $s = 0$ is clearly specified, but the end point condition at $s = l_p$ is still not prescribed since we don't know the curve $\zeta'(s)$ ($\zeta(s)$ is the curve where the end point lies). Expression Eq. (4.48) is just a condition that the system must satisfy in order to minimize the energy. Since it is not feasible to find the exact end point condition at $s = l_p$ due to the presence of thermal fluctuations, we choose a different path to solve the problem by prescribing M_3 . The value of M_3 can be taken from experiment or simulations as a function of the applied number of turns n . Now the right hand side of Eq. (4.52) is not equal to zero and Eq. (4.53) no longer holds.

Next, we outline the procedure to solve the variational problem for a fixed force F . Given the prescribed external moment M_3 for a given $n = \hat{n}^*$ a solution $S[C_1^*]$ for each \hat{n}^* can be obtained as follows:

- Assume $C_1(n)$ (or equivalently $\zeta'(l_p)$) in Eq. (4.48) to be known and plug it into equation Eq. (4.43).

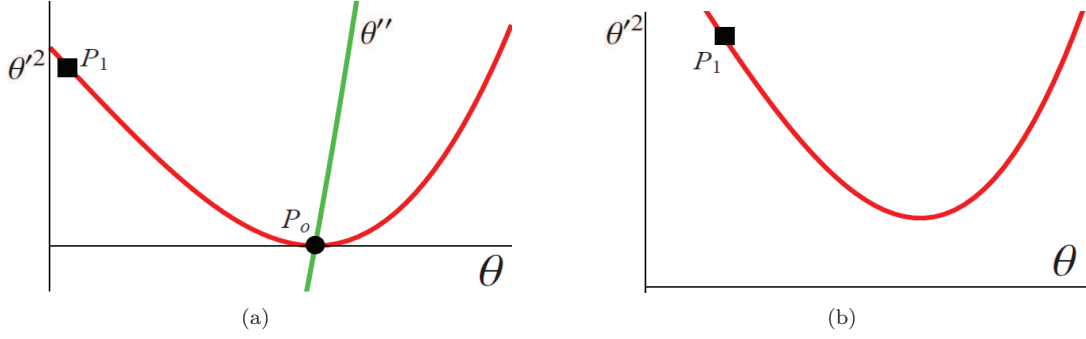


Figure 4.10: (a) The constant helix solution is at point P_o where $\theta' = \theta'' = 0$ and $C_1 = F - G_{flu}^*$. (b) If $C_1 > F - G_{flu}^*$ (with M_3 and r independent of n) in Eq. (4.43), the profile of θ'^2 shifts upward and there is no value of θ for which $\theta' = 0$. Consequently, the first boundary condition Eq. (4.44) can not be satisfied.

- Pick a value of $r^t(n)$ and solve the differential equation Eq. (4.42). The value of l_p is obtained from condition Eq. (4.44) at $s = 0$. Keep iterating on the value of r until $r = r^*$, where r^* satisfies equation Eq. (4.41).
- Once $S[C_1] = \{M_3, r^*, \theta^*(s), l_p\}$ is known, the associated value of n with $S[C_1]$ can be obtained from the link constraint Eq. (4.25). The value of $C_1 = C_1^*$ that recovers $n = \hat{n}^*$ determines the solution $S[C_1^*]$, therefore for each value of $\hat{n}^* \in [n_i, n_p]$ we have a unique solution given by $\{M_3, r^*, \theta^*(s), l_p^*, C_1^*\}$.

For the special case of multivalent ions $r = d_o/2$ can be assumed to be approximately constant [22], such that there is no need to satisfy the integral constraint Eq. (4.41) and the computational cost of obtaining the solution is significantly reduced.

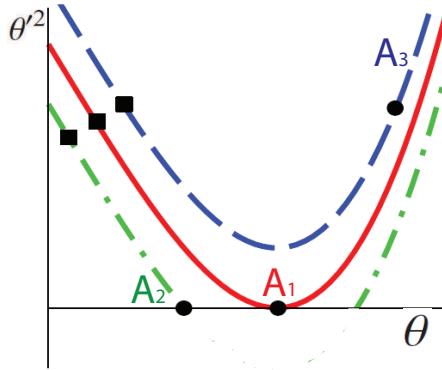


Figure 4.11: Profile of θ'^2 as a function of θ . In the red curve A_1 the point $\theta_{min} = 0$ corresponds to the helix solution that can be extended to infinity. The green curve A_2 shows the scenario corresponding to a fixed end loop length such that point P_o has a natural boundary condition. The blue curve A_3 describes the scenario where the end loop length is variable as described in C.4. Black squares represent point P_1 while black circles represent point P_o .

4.4.8 Variable Pitch – Analytical Solutions: *Three* cases

Here we assume that the values of r and M_3 are prescribed, and in order to obtain closed analytical solutions, we further assume that $U(r, \theta) = U(r)$. This can be done by setting $\theta = \pi/6$ (Clauvelin et al. [73] use $\theta = 0$, this does not affect the solution by much) in $U(r, \theta)$ when using expressions

corresponding to monovalent salts [21], or using electrostatic interactions in the presence of multivalent ions [22]. In any case, the differential equations Eq. (4.42) and Eq. (4.43) reduce to the form of equations (2.41) and (2.42) analyzed by Coleman and Swigon [66]. The differential equation for the helical angle θ is:

$$(\theta')^2 = \frac{4}{D^2} \sin^4 \theta + \frac{2wM_3}{D} \sin 2\theta - \Gamma, \quad (4.55)$$

where,

$$\Gamma = -2 \frac{U + C_1}{K_b}, \quad w = \frac{\chi}{K_b}, \quad D = 2r. \quad (4.56)$$

Let $\eta = \cot \theta$ as in [66], then we obtain the differential equation in the variable η :

$$P(\eta) = (\eta')^2 = \frac{4}{D^2} + \frac{4wM_3}{D} \eta(1 + \eta^2) - \Gamma(1 + \eta^2)^2. \quad (4.57)$$

The 4 roots of the polynomial $P(\eta)$ are functions of r, K_b , M_3 and Γ . Once we know how many roots of Eq. (4.57) are real numbers, the solution is given by an integral of the form [66, 132, 133]:

$$\Delta s = \int_{\eta_1}^{\eta} \frac{dx}{\sqrt{P(x)}}, \quad (4.58)$$

where $\eta_1 = \cot \alpha$ and $\alpha = \theta(s_1)$. In the rest of section 4.4.8 we will make use of the treatment given by Enneper and Muller [132] and Greenhill [133] on elliptic functions and integrals.

First Case: Curve A_2 Here we consider the situation where Eq. (4.57) has only two real roots, such that $(\theta')^2_{min} < 0$ (See curve A_2 in figure 4.11). In this case point P_o has a natural boundary condition where $\theta' = 0$ at $s = s_o$, while point P_1 is a variable end point that satisfies Eq. (4.48), where $\theta = \zeta$ at $s = s_1$ is smaller in magnitude than $\theta = \alpha$ at $s = s_o$ since the curvature is larger at the end loop. This case is also presented in Coleman and Swigon [66] with a different set of boundary conditions. Let Γ_i for $i = 1, 2$ be the value of Γ when $\theta' = 0$:

$$\Gamma_i = \frac{4}{D^2} \sin^4 \alpha_i + \frac{2wM_3}{D} \sin 2\alpha_i, \quad (4.59)$$

where α_1 and α_2 are the values of θ at one of the roots of $\theta' = 0$. By using $\eta_1 = \cot \alpha_1$ and $\eta_2 = \cot \alpha_2$, $P(\eta)$ has two real roots ($\eta_1 > \eta_2$) and hence there are numbers b and c for which Eq. (4.57) can be rewritten as:

$$P(\eta) = B(\eta - \eta_1)(\eta - \eta_2) ((\eta - b)^2 + c^2), \quad (4.60)$$

where

$$B = -4 \frac{1 + wM_3 D \eta_1 (1 + \eta_1^2)}{(1 + \eta_1)^2 D^2}. \quad (4.61)$$

For $\infty > \eta > \eta_1$ and $\eta_2 > \eta > -\infty$ [132, 133]:

$$\Delta s = \frac{1}{\sqrt{B}} \frac{1}{\sqrt{pq}} \text{cn}^{-1} \left\{ \frac{p(\eta - \eta_2) - q(\eta - \eta_1)}{p(\eta - \eta_2) + q(\eta - \eta_1)}, m \right\}, \quad (4.62)$$

where

$$p^2 = ((\eta_1 - b)^2 + c^2), \quad q^2 = ((\eta_2 - b)^2 + c^2). \quad (4.63)$$

and the elliptic modulus m is given by:

$$m^2 = \frac{1}{4} \left\{ \frac{(p+q)^2 - (\eta_1 - \eta_2)^2}{pq} \right\}. \quad (4.64)$$

Let $u = \sqrt{B}\Delta s\sqrt{pq}$ such that:

$$\text{cn}(u) = \cos \phi = \frac{p(\eta - \eta_2) - q(\eta - \eta_1)}{p(\eta - \eta_2) + q(\eta - \eta_1)}, \quad (4.65)$$

$$\phi = \text{am}(u) = 2 \arctan \sqrt{\frac{q(\eta - \eta_1)}{p(\eta - \eta_2)}}. \quad (4.66)$$

From Eq. (4.65) we can solve for θ ⁷:

$$\theta = \cot^{-1} \left\{ \frac{\eta_1 q [\text{cn}(u) + 1] + \eta_2 p [\text{cn}(u) - 1]}{q [\text{cn}(u) + 1] + p [\text{cn}(u) - 1]} \right\}. \quad (4.67)$$

Second Case: Curve A_1 Here we consider the situation where Eq. (4.57) has only one real root, such that $(\theta')_{min}^2 = 0$. In this case, again, point P_o has a natural boundary condition where $\theta' = 0$ at $s_o = 0$, while point P_1 is still a variable end point with less curvature than point P_o (See curve A_1 in figure 4.11). The polynomial $P(\eta)$ is given by:

$$P(\eta) = B(\eta - \eta_1)^2((\eta - b)^2 + c^2), \quad (4.68)$$

where B is the same as in the *First Case*. We are interested in the region $\alpha_1 < \theta < \frac{\pi}{4}$ such that η lies between η_1 and $-\infty$ we have:

$$\int_{-\infty}^{\eta} \frac{d\eta}{\sqrt{P(\eta)}} = \frac{1}{\sqrt{BN(\eta_1)}} \cosh^{-1} \left\{ \frac{N(\eta_1)N(\eta)}{c(\eta_1 - \eta)} \right\}, \quad (4.69)$$

where $N(x) = \sqrt{(x-b)^2 + c^2}$. Now consider the scenario where $\alpha_1 < \theta_2 < \theta_3 = \pi/4$, with $\eta_2 = \cot \theta_2$ and $\eta_3 = \cot \theta_3$ such that $\eta_3 < \eta_2 < \eta_1$. The change in the arc-length from η_3 up to η_2 is:

$$\Delta s = \int_{\eta_3}^{\eta_2} \frac{d\eta}{\sqrt{P(\eta)}} = \int_{-\infty}^{\eta_2} \frac{d\eta}{\sqrt{P(\eta)}} - \int_{-\infty}^{\eta_3} \frac{d\eta}{\sqrt{P(\eta)}}. \quad (4.70)$$

Note that in equation Eq. (4.69) if $\eta = \eta_1$, then the integral approaches infinity. The point η_1 corresponds to point $\theta'_{min} = 0$ in curve A_1 , and if $C_1 = F - G_{f;u}^*$ this represents the constant helix that can be extended to infinity⁸.

Third Case: Curve A_3 Here $(\theta')_{min}^2 > 0$ as in curve A_3 in figure 4.11, such that Eq. (4.57) has no real solution. One can envision this scenario when both points P_o and P_1 are variable and satisfy transversality conditions of the form of Eq. (4.45), as in the case where the end loop length might be changing as described in C.4. Here there is no real root of the quartic $P(\eta)$, such that we can

⁷The solution given in equation Eq. (4.67) can also be expressed as: $\theta = \cot^{-1} \left\{ \frac{q\eta_1 - p\eta_2 \tan^2(\text{am}(u)/2)}{q - p \tan^2(\text{am}(u)/2)} \right\}$ where $u = \sqrt{B}\Delta s\sqrt{pq}$ and the elliptic modulus of the am function is given by Eq. (4.64).

⁸The analytical solutions arising in the three scenarios were checked using numerical integration in MATLAB.

find two polynomials:

$$Q(\eta) = (\eta - m)^2 + e^2, \quad \hat{Q}(\eta) = (\eta - \hat{m})^2 + \hat{e}^2, \quad (4.71)$$

where $P(\eta) = BQ(\eta)\hat{Q}(\eta)$, and B is a constant. Following the treatment in [132], we define the constants \hat{h} and h :

$$\hat{h}^2 = (m - \hat{m})^2 + (e + \hat{e})^2, \quad h^2 = (m - \hat{m})^2 + (e - \hat{e})^2. \quad (4.72)$$

The solution to the problem is given by:

$$\int \frac{d\eta}{\sqrt{P}} = \frac{2}{\sqrt{B}(\hat{h} + h)} F(\phi, k), \quad (4.73)$$

where,

$$\omega = \tan^{-1} \left\{ \sqrt{\frac{4e^2 - (\hat{h} - h)^2}{(\hat{h} + h)^2 - 4e^2}} \right\}, \quad \phi = \omega + \tan^{-1} \left\{ \frac{\eta - m}{e} \right\}, \quad k^2 = \frac{4rh}{(\hat{h} + h)^2}, \quad (4.74)$$

and $F(\phi, k)$ is the incomplete elliptic integral of the first kind with modulus k and argument ϕ . The arc-length between η_1 and η_2 is given by:

$$\Delta s = \int_{\eta_1}^{\eta_2} \frac{d\eta}{\sqrt{P}}, \quad (4.75)$$

where $\eta_2 = \cot \theta_2$, $\eta_1 = \cot \theta_1$ and $0 \leq \theta_2 < \theta_1 \leq \frac{\pi}{4}$.

In figure 4.12 we show a non-constant profile of θ as a function of s for case 3 as described above. $\theta(s)$ is nearly constant close to the midsection but diverges fast near the end points. We have used different values of M_3 , where the average values of $\theta = \theta_{mid}$ increases as a function of M_3 . The minimum value of θ_{mid} corresponds to the constant solution value of the applied torque $M_3 = M_c$. The inset shows case 2 described in section 4.4.8. Note that the plateau value of θ_{mid} is only a function of the prescribed M_3 and it is the same for both case 2 as well as case 3.

4.5 Conclusions

We have used a variational approach to solve the energy minimization problem that corresponds to DNA single molecule experiments in two different set ups. Our focus has been on helical solutions with non-constant pitch. Here we give a short summary of the results.

In the experiments of Charvin et al. [17], Noom et al. [18] two DNA molecules are subjected to an external force F and rotations R about the helical axis. In this type of experiments extended ply structures are formed as depicted in figure 4.1. When the DNA molecules are nicked, so that there is no twist stored in the filament, the structures are known as braids [17]. We used variable curvature helices to model extended ply structures that can reach lock-up at a critical number of rotations Ca_c . At lock-up the helical angle α at the end points reaches the maximum physical value of $\pi/4$ [19], and as more rotations are added to the system more complex DNA structures are formed as shown by Charvin et al. [17]. In fact, our model is in excellent agreement with MC simulations and experiments reported in Charvin et al. [17]. Variable pitch plies using essential (rigid/displacement loading) boundary conditions and natural (dead/force loading) boundary conditions where the length of the ply is fixed have been studied [70] before but the novelty of our work is in relaxing the fixed length of the ply and solving the problem such that the ply's length minimizes energy [127, 128].

In the experiments of Brutzer et al. [10], Mosconi et al. [11], Forth et al. [12], Lipfert et al. [27]

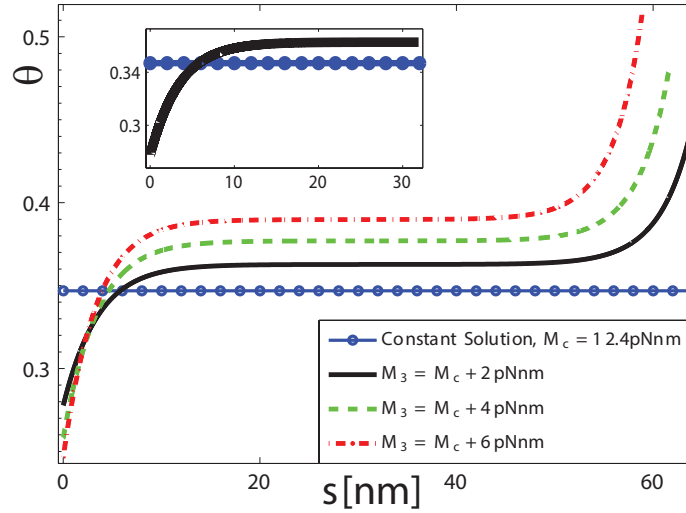


Figure 4.12: Profile of θ as a function of s for case 3 described in section 4.4.8. In the calculations we have used $F = 4\text{pN}$ and $r = 1.5\text{nm}$, together with the internal interactions described by multivalent ions in [22]. $\theta(s)$ is nearly constant close to the midsection but diverges fast near the end points. We have used different values of M_3 with average values of θ_{mid} increasing as a function of M_3 . The inset shows case 2 described in section 4.4.8. Note that the plateau value of θ is only a function of the prescribed M_3 and it is the same for case 2 and case 3.

a DNA molecule is attached to a substrate at one end while subjected to a tensile force and twisted by a specific number of turns n at the other end (see figure 4.2). Previous studies using constant pitch helices describe experiment quantitatively. Although experiments and simulations show that the external moment as a function of the applied number of turns n is nearly constant, there is still a significant discrepancy in the values of the external torque obtained from direct experimental methods [12], indirect experimental methods [11], MC simulations [5] and theory [4, 21, 45, 73]. Motivated by this we generalized to a variable pitch helix and presented a model where the influence of the external torque on the helical solution can be studied. In this general model M_3 is treated as prescribed quantity which is a function of the number of turns n that can be extracted from experiment or simulations. We found that the helical angle is approximately constant $\theta \sim \theta_{mid}$ along the mid-section of a plectoneme and it quickly diverges at the boundary points near the tails and end loop. The value of $\theta = \theta_{mid}$ increases as a function of the prescribed M_3 (as shown in figure 4.12) where the minimum value of θ_{mid} is obtained for the constant pitch solution. We have also shown, starting from the variable pitch solution, that the system can only yield a constant pitch helix when $M_{ext} = M_3$ and r are independent of n . Furthermore, even if r is not independent of n , the system can be approximated as having a constant angle for long DNA filaments and large number of turns n .

Chapter 5

A heterogeneous fluctuating rod under imposed tension and twist

Single molecule torsion experiments where the molecules (i.e. actin, DNA, RNA, microtubules, oligoprotein chains) are subjected to stretching and/or twisting using magnetic tweezers and optical traps have been usually interpreted using two models that were proposed in the late 1990s. The first model is that of Moroz and Nelson [40, 56] which assumes that the molecule is a homogeneous elastic rod under constant force and torque fluctuating around its minimum energy configuration. The second model is that of Marko [134] which is similar to that of Moroz and Nelson except that the number of turns applied by the experimental device is specified rather than the twisting moment. Both models are applicable when the force is sufficiently high and the molecules are long. In the last fifteen years torsion experiments on bare DNA and RNA have been successfully interpreted using these models. Motivated by the success of these models in interpreting data, more sophisticated torsional experiments are being performed where small molecules, including drugs, are bound to DNA while they are loaded in tension and torsion. Binding of small molecules changes the local mechanical properties of DNA so that one can no longer think of it as a homogeneous rod. Thus a more sophisticated version of these models is needed where variation in mechanical properties is accounted for.

Besides Moroz and Nelson [40, 56] and Marko [134] several other authors have provided theoretical contributions to the challenges of DNA mathematical modeling. For example, a continuous rod model was used in Manning et al. [135] to study DNA ring closure and cyclization. In order to account for the base-pair details and sequence dependence of the DNA structure various discrete rigid-base (basepair) models are available in the literature [136–139]. More recently, Moakher and Maddocks [140] developed a theoretical description of double-stranded filaments using a continuous birod. In Wolfe et al. [141] continuous and discrete versions of both rods and birods are reviewed together under a unified Lie group notation. The authors show that the birod model in fact converges to the single filament model for sufficiently long filaments. This is the regime where our (fluctuating rod) model of DNA is situated. For more detailed reviews on different mechanical and mathematical tools used in the study of DNA, we refer the reader to [42] and [142] and the references therein.

For the applications discussed in this chapter a base-pair level description of DNA would be too expensive. Instead, what is needed is a twistable worm-like chain model in which mechanical properties such as bending and twisting moduli can take on a few different values. Such a model for calculating the thermo-mechanical properties of fluctuating molecules was proposed by Su and Purohit [49]. Their model was based on Gaussian integrals (using similar techniques to the ones used in Gaussian-elastic networks [143]), and included the effects of bending but neglected twist. In the present work we build on their methods and those of Zhang and Crothers [144] to include the

effects of twist, stretch and the coupling of these deformation modes. Using our methods we can study both types of ensembles where number of turns is controlled or torque is controlled. Our model has the advantage that it is useful not only for very long molecules (length $\gg \mu m$), but also for shorter ones (a few hundred nanometers in length), where finite size effects arising from boundary conditions need to be accounted for. In this chapter we will confine ourselves to only one type of boundary condition which corresponds to a partially clamped set-up. This is appropriate for single molecule experiments in which one end of the molecule is fixed at the origin (clamped) while the other end is free to have transverse displacements (see figure 5.1). For a more complete discussion of various types of boundary conditions we refer the reader to Su and Purohit [49].

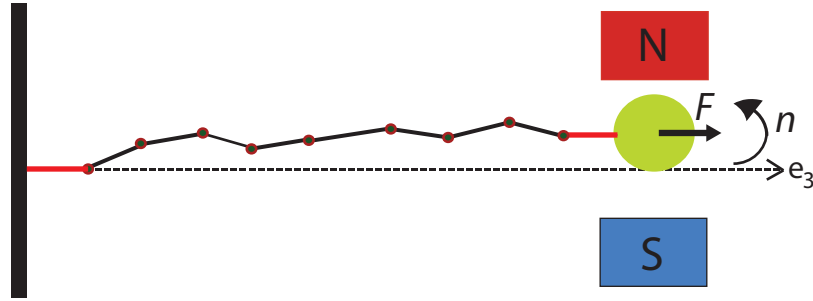


Figure 5.1: Partially clamped set-up realized in single molecule experiments using magnetic tweezers. The rod is subject to tension and twist. The rod is discretized so that we view it as a chain of segments of length l . The chain is free to fluctuate due to thermal effects, but the first and last segments (red) are parallel to the \mathbf{e}_3 axis. The first segment is clamped while the last one is free to have transverse displacements.

We use our model to interpret DNA torsion experiments with small molecule binding. We also use it to understand another set of experiments in which DNA exists as a mixture of metastable phases. Such phase co-existence has been observed and documented using extension-rotation curves (also called ‘hat’ curves by some authors) which are the standard output of DNA torsion experiments using magnetic or optical tweezers. It is well known that the extension-rotation curves are not always symmetric for positive and negative number of turns n (see fig. 5.1). To understand the origins of the asymmetry we first recall that when $n > 0$ DNA is twisted in the same sense as its intrinsic twist while when $n < 0$ it is twisted in the opposite sense. For small applied tensions F , when $n > 0$ there is a transition from straight to plectonemic DNA. DNA can be treated as a homogeneous elastic rod in this regime and with a careful treatment of electrostatics (and excluded volume interactions) the results of a large number of experiments and simulations can be systematically analyzed [4, 5, 21, 22, 68, 73]. The situation is more complicated for large F and/or $n < 0$. Depending on F , DNA sequence content or salt concentrations, DNA may change its structure under negative n instead of adopting a B-DNA plectonemic configuration. This results in the asymmetry observed in the extension rotation curves and has lead the field to study other possible DNA forms with left-handed internal structure (B-DNA is right handed). It was suggested that negative turns can induce DNA denaturation, where double stranded DNA separates into two single stranded DNA pieces Marko [55]. For a long time denatured DNA, also known as DNA bubbles, was believed to be the only stable DNA state under negative n , but recent single molecule studies show that twisted DNA at low tension coupled with high monovalent salt concentration may be in the Z-form [145]. Independently, Bryant et al. [57] suggested that DNA subjected to negative n may be making a transition into another possible DNA state L-DNA, which could itself be a mixture of two or more known states (B-DNA, Z-DNA, ssDNA, S-DNA, P-DNA). Furthermore, DNA structural transitions have been shown to be not only functions of the mechanical loads, temperatures and ion concentrations, but also sequence (GC or AG tracts). Since the experimental work leaves an open question regarding the DNA state under negative n , our fluctuating rod/chain model seems to be a plausible tool to investigate quantitatively the consequences of various hypotheses and make predictions that can be

verified by further experiments.

5.1 Model Description

Consider the (discretized) rod subject to tension and twist depicted in figure 5.1. Our problem is to evaluate the partition function and free energy of a such a rod assuming that the energy can be expressed as a quadratic function of the kinematic variables that characterize its configuration.

5.2 Energy of the system

A general expression for the stored energy density of a one-dimensional, isotropic, unsharable rod which is a quadratic function of the deformations is given in Chouaieb et al. [75] and Nelson [146]:

$$\Lambda(\kappa, u_3, \epsilon) = \frac{1}{2} [A\kappa^2 + Cu_3^2 + 2gu_3\epsilon + S\epsilon^2], \quad (5.1)$$

where u_3 is the physical twist (strain) of the rod, κ is the bending curvature and ϵ is the stretch. A, C, g and S are the bending, twist, twist-stretch and stretch modulus respectively. The bending modulus A and the twisting modulus C have units of pNnm² and can be related to a bending persistence length $A = k_B T A_p$ and a twisting persistence length $C = k_B T C_p$. So from here on we will be giving the bending and twisting modulus in terms of $k_B T$, keeping in mind that the units are in pNnm². Expression Eq. (5.1) it not the most general model for rod-like macromolecules, as there can be effects associated with anisotropic bending and shearing along the cross-sections of the filament (see for instance [147]). Furthermore, in the case of DNA, equation Eq. (5.1) does not account for base-pair details. But, we show in this chapter that the energy Eq. (5.1) is good enough to provide new insights regarding the mechanical response of twisted DNA in certain single molecule experiments.

Next, for simplicity we present our methods in the context of an inextensible chain subjected to only bending and twist, but the same procedure can be carried out adding contributions from stretch and twist-stretch coupling. Here we present our method for a fixed torque ensemble in order to compare our results with those of [56]. Results including stretch, twist-stretch coupling and the n -ensemble are presented in the D.3.

Given a kinematic variable x , we denote the fluctuations away from the static (zero temperature state) as $\delta x = x - x_{min}$. Then the energy of the states around the minimum energy configuration (static) can be expanded using a Taylor series up to second order:

$$E = E_{min} + \frac{\partial E}{\partial x} \delta x + \frac{1}{2} \frac{\partial^2 E}{\partial x^2} (\delta x)^2 = E_{min} + \frac{1}{2} \frac{\partial^2 E}{\partial x^2} (\delta x)^2, \quad (5.2)$$

since at equilibrium we require $\partial E / \partial x = 0$. The energy for our rod is given by

$$E = \int_0^L \frac{1}{2} [A\kappa^2 + Cu_3^2] ds - 2\pi\Delta L_k M_3 + F\Delta x. \quad (5.3)$$

where L is the length of the rod, s is the arc-length, $2\pi\Delta L_k M_3$ is the potential energy of the external torque M_3 and $F\Delta x$ is the potential energy of the force due to the shrinking ($\Delta x = L - x$) caused by thermal fluctuations. Now let $[\mathbf{e}_i] = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ be a standard lab frame. The rod has one end fixed at the origin and the other end subjected to the external force $\mathbf{F} = F\mathbf{e}_3$ and external moment $\mathbf{M} = M_3\mathbf{t}$, where \mathbf{t} is the unit tangent vector which forms an angle $\theta(s)$ with the \mathbf{e}_3 axis (note that

at the end where the force and torque are applied $\theta = 0$):

$$\mathbf{t}(s) = -\sin\theta \sin\psi \mathbf{e}_1 + \sin\theta \sin\psi \mathbf{e}_2 + \cos\theta \mathbf{e}_3. \quad (5.4)$$

In the above $\psi(s)$ is an Euler angle that measures the rotation of the filament about \mathbf{e}_3 (see D.3). Then the curvature of the rod is given by:

$$\frac{d\mathbf{t}}{ds} \cdot \frac{d\mathbf{t}}{ds} = \kappa^2 = \theta'^2 + \psi'^2 \sin^2\theta, \quad (5.5)$$

where $()'$ denotes differentiation with respect to the arc-length s . Also note that the shrinking of the end-to-end distance through which the external force does work is

$$\Delta x = \int_0^L [1 - \cos\theta] ds. \quad (5.6)$$

The number of turns n measured (or applied) by the apparatus is equal to the excess link ΔLk stored in the rod, which is composed of the stored twist ΔTw , the writhe of the curve describing the rod axis Wr and any changes to the natural-internal twist of the filament ΔTw_o (i.e unwinding of a B-DNA double helix when there is a phase transition):

$$\Delta Lk = \Delta Tw + Wr + \Delta Tw_o \quad (5.7)$$

The stored twist is the sum of the local twist strain over the entire rod:

$$2\pi\Delta Tw = \int_0^L u_3 ds. \quad (5.8)$$

The changes in the natural twist ΔTw_o are given in an analogous manner through a shift in the local internal twist u^o :

$$2\pi\Delta Tw_o = \int_0^L u^o ds, \quad (5.9)$$

while for the writhe Wr we can use the expression given by [113] since we constrain our analysis to the high force-regime where $\theta(s)$ is small:

$$2\pi Wr = \int_0^L \frac{\mathbf{e}_3 \cdot (\mathbf{t} \times \mathbf{t}')}{1 + \mathbf{e}_3 \cdot \mathbf{t}} = \int_0^L [\psi'(1 - \cos\theta)] ds. \quad (5.10)$$

Combining Eq. (5.7) with Eq. (5.3) :

$$E = \int_0^L \hat{I}(f, M_3, \theta(s), \theta(s)') ds \quad (5.11)$$

where the stored energy density is, after completing the square in the u_3 terms:

$$\begin{aligned} \hat{I}(F, M_3, \theta, \theta') &= \frac{A}{2} (\theta'^2 + \psi'^2 \sin^2\theta) - \frac{M_3^2}{2C} - F \cos\theta \\ &- M_3 \psi'(1 - \cos\theta) + \frac{C}{2} \left[u_3 - \frac{M_3}{C} \right]^2 - u^o M_3 + F. \end{aligned} \quad (5.12)$$

We point out that the last two terms in the expression above are usually dropped in the literature because these are just constants that do not affect the response functions $\langle \Delta x \rangle$ and $\langle n \rangle$. We keep these terms because they are important in solving for the coexistence of phases in DNA (see Appendix

D).

We constrain ourselves to tensile forces for which the minimum energy configuration of the system is straight. The forces and moments are assumed to be such that the torsional buckling instability of classical mechanics does not occur. This requires $K_B F - M_3^2/4 > 0$ ([148]). Therefore, $\theta_{min} = \theta'_{min} = 0$ are the values corresponding to the minimum energy configuration at zero temperature for a rod subjected to an applied torque M_3 and tension F . Performing the minimization with respect to the twist u_3 we get the zero temperature twist:

$$u_3^{min} C = M_3 \quad (5.13)$$

such that the energy can be written as:

$$\begin{aligned} E &= \int_0^L \left[\frac{A}{2} (\theta'^2 + \psi'^2 \sin^2 \theta) - \frac{M_3^2}{2C} - F \cos \theta \right] ds \\ &+ \int_0^L \left[M_3 \psi' (\cos \theta - 1) + \frac{C}{2} (\delta u_3)^2 - u^o M_3 + F \right] ds \end{aligned} \quad (5.14)$$

where $\delta u_3 = u_3(s) - u_3^{min}$. The minimum energy and the linking number of the zero temperature chain are therefore given by:

$$E_{min} = - \int_0^L \left[\frac{M_3^2}{2C} + u^o M_3 \right] ds, \quad 2\pi n = 2\pi \Delta L k = \int_0^L [u_3^{min} + u^o] ds, \quad (5.15)$$

where $C(s), u^o(s)$ and $u_3^{min}(s)$ are not constants for a heterogeneous rod. Now we use the small angle approximation up to quadratic order:

$$\sin^2 \theta \sim \theta^2, \quad \cos \theta \sim 1 - \theta^2/2, \quad \frac{1 - \cos \theta}{1 + \cos \theta} \sim \theta^2/4. \quad (5.16)$$

We substitute these into equation Eq. (5.14) and after completing the square in the variable ψ' we obtain an approximate energy density:

$$\hat{I}(f, M_3, \theta, \theta') = \frac{A}{2} \left[\theta'^2 + \theta^2 \left(\psi' - \frac{M_3}{2A} \right)^2 \right] + \left[\frac{F}{2} - \frac{M_3^2}{8A} \right] \theta^2 - \frac{M_3^2}{2C} + \frac{C}{2} (\delta u_3)^2 - u^o M_3. \quad (5.17)$$

Next we define:

$$\Psi' = \psi' - \frac{M_3}{2A}, \quad (5.18)$$

$$v_x = \theta \sin(\Psi), \quad v_y = \theta \cos(\Psi), \quad (5.19)$$

such that

$$v_x^2 + v_y^2 = \theta^2, \quad (5.20)$$

$$v'_x = \theta' \sin \Psi + \theta \Psi' \cos \Psi, \quad v'_y = \theta' \cos \Psi - \theta \Psi' \sin \Psi, \quad (5.21)$$

and the bending curvature of the rod is given by:

$$\kappa^2 = v'^2_x + v'^2_y = \theta'^2 + \theta^2 \Psi'^2 = \theta'^2 + \theta^2 \left(\psi' - \frac{M_3}{2A} \right)^2. \quad (5.22)$$

Finally, the energy of the rod can be decomposed as:

$$\begin{aligned}
E &= \int_0^L \left[\frac{A}{2} v_x'^2 + \left(F - \frac{M_3^2}{4A} \right) \frac{v_x^2}{2} - \frac{M_3^2}{4C} + \frac{C}{4} (\delta u_3)^2 - u^o \frac{M_3}{2} \right] ds \\
&+ \int_0^L \left[\frac{A}{2} v_y'^2 + \left(F - \frac{M_3^2}{4A} \right) \frac{v_y^2}{2} - \frac{M_3^2}{4C} + \frac{C}{4} (\delta u_3)^2 - u^o \frac{M_3}{2} \right] ds,
\end{aligned} \tag{5.23}$$

which is quadratic in the kinematic variables v_x, v_y and δu_3 . Since v_x and v_y are independent we are confronted with the calculation of the partition functions of two identical (discretized) rods. We now follow the method described in [49] where the partition function of a discretized rod is calculated.

5.2.1 Discretized version of the energy

In this section for notation purposes we let $v(s)$ be the kinematic variable representing either v_x or v_y . We discretize our rod into N segments of length l to get a chain. The energy of each chain in discrete form is then given by:

$$\begin{aligned}
E_F &\approx \sum_{i=1}^{N-1} \frac{A_i}{2} \left(\frac{v_{i+1} - v_i}{l} \right)^2 l + \frac{F}{2} \sum_{i=1}^N v_i^2 l - \frac{M_3^2}{8} \sum_{i=1}^N \frac{v_i^2 l}{A_i} \\
&- \frac{M_3^2}{4} \sum_{i=1}^N \frac{l}{C_i} + \sum_{i=1}^N \frac{C_i}{4} (\delta_i u_3)^2 l - \frac{M_3}{2} \sum_{i=1}^N u_i^o l
\end{aligned} \tag{5.24}$$

where N is the number of segments (e.g. base pairs), l is the length of each segment and the total length of the system is $L = \sum_{i=1}^N l$. The expression above can be rearranged such that:

$$E_F \approx \sum_{i=1}^{N-1} \kappa_i (v_{i+1} - v_i)^2 + \sum_{i=1}^N v_i^2 q_i - \sum_{i=1}^N m_i + \sum_{i=1}^N c_i (\Delta_i)^2 \tag{5.25}$$

where we define

$$f = \frac{Fl}{2}, \quad \kappa_i = \frac{A_i}{2l}, \quad q_i = f - \frac{M_3^2}{16\kappa_i}, \quad m_i = \frac{M_3^2 l}{2C_i} + \frac{M_3}{2} u_i^o l, \quad c_i = \frac{C_i}{4l}, \quad \Delta_i = \delta_i u_3 l. \tag{5.26}$$

Note that the same analysis can be carried out for $l_i \neq l$ for all $i \in [1, N]$. Since u_3 has units (Length)⁻¹, we pick Δ_i as above to make the kinematic variable dimensionless.

Rearranging terms in Eq. (5.25), introducing $\kappa_0 = \kappa_N = 0$ and $\beta^{-1} = k_B T$ (where T is the absolute temperature and k_B is the Boltzmann constant) we can express the energy as:

$$\begin{aligned}
\beta E_F &\approx \beta \left[\sum_{i=1}^N \{ (\kappa_{i-1} + \kappa_i + q_i) v_i^2 \} - 2 \sum_{i=1}^{N-1} \kappa_i v_i v_{i+1} - \sum_{i=1}^N m_i + \sum_{i=1}^N c_i \Delta_i^2 \right] \\
&= \vec{\Omega} \mathbf{K} \vec{\Omega}^T + \vec{\Lambda} \mathbf{T} \vec{\Lambda}^T - \beta \sum_{i=1}^N m_i,
\end{aligned} \tag{5.27}$$

where $\vec{\Omega} = [v_1, v_2, \dots, v_N]$, $\vec{\Lambda} = [\Delta_1, \Delta_2, \dots, \Delta_N]$, \mathbf{T} is an N -dimensional diagonal matrix:

$$T_{ij} = \beta c_i \delta_{ij}, \tag{5.28}$$

and \mathbf{K} is the N -dimensional tridiagonal matrix which can be written in compact form as:

$$[\mathbf{K}]_{ij} = \left\{ \beta (\kappa_{i-1} + \kappa_i + q_i) \delta_{ij} - \beta \kappa_t \cdot \delta_{(|i-1|,1)} \right\}, \quad (5.29)$$

where $t = \min(i, j)$ and δ is the Kronecker delta.

5.2.2 Statistical mechanics of the chain

At fixed T , F and M_3 the partition function of the chain is given by:

$$Z = \sum_v \exp[-\beta E_F] \quad (5.30)$$

where the summation is over all the configurations that satisfy the constraints at the ends and $\beta = 1/k_B T$. The ensemble average of the energy is:

$$\langle E_F \rangle = Z^{-1} \sum_v E_F \exp[-\beta E_F] = -\frac{\partial \ln Z}{\partial \beta} \quad (5.31)$$

Since the energy of the chain is quadratic in the kinematic variables v_i and Δ_i (see Eq. (5.27)), the equipartition theorem can be readily applied to obtain:

$$\langle E_F \rangle = \frac{D}{2} k_B T - \left\langle \sum_{i=1}^N m_i \right\rangle = \frac{D}{2} k_B T - \sum_{i=1}^N m_i \quad (5.32)$$

where D is the number of degrees of freedom of the system. Combining equation Eq. (5.32) with Eq. (5.31) and integrating with respect to β gives:

$$\ln Z = -\frac{D}{2} \ln \beta + \left(\sum_{i=1}^N m_i \right) \beta - W(F, M_3) \quad (5.33)$$

where $W(F, M_3)$ is an unknown function of F and M_3 . All the thermo-mechanical quantities can be derived as functions of $W(F, M_3)$ and its first and second partial derivatives as shown by [49]. The free energy of the chain can be expressed as:

$$G(T, F, M_3) = -2 \ln Z / \beta, \quad (5.34)$$

where the factor of 2 accounts for the two identical discretized rods (see Eq. (5.23)). The entropy of the chain is given by:

$$S = -\frac{\partial G}{\partial T} \Big|_{F, M_3} = D k_B (1 + \ln \beta^{-1}) - k_B W(F, M_3) \quad (5.35)$$

The extension of the chain $\langle x \rangle$ is given by:

$$\langle x \rangle = L - \frac{\partial G}{\partial F} \Big|_{T, M_3} = L - k_B T \frac{\partial W}{\partial F} \quad (5.36)$$

Note that $\langle x \rangle$ is linearly decreasing with the temperature T , and this relation is independent of the heterogeneity of the chain. The variance of extension scales as T^2 and is given by:

$$\langle (\Delta x)^2 \rangle = k_B T \left(\frac{\partial \langle x \rangle}{\partial F} \right)_{T, M_3} = - (k_B T)^2 \frac{\partial W^2}{\partial^2 F} \quad (5.37)$$

Note that since $\langle (\Delta x)^2 \rangle \geq 0$ then $\partial W^2 / \partial^2 F \leq 0$, such that $\partial W / \partial F$ is a decreasing function of F (shrinking). The link of the chain $\langle 2\pi n \rangle$ is given by:

$$\langle 2\pi n \rangle = - \frac{\partial G}{\partial M_3} \Big|_{T, F} = \sum_{i=1}^N \left[\frac{M_3 l}{C_i} - u_i^o l \right] - k_B T \frac{\partial W}{\partial M_3} \quad (5.38)$$

The variance of the link also scales as T^2 and is given by:

$$\langle (\Delta(2\pi n))^2 \rangle = 2\pi k_B T \left(\frac{\partial \langle n \rangle}{\partial M_3} \right)_{T, F} = \sum_{i=1}^N \frac{l}{C_i} - (k_B T)^2 \frac{\partial W^2}{\partial^2 M_3} \quad (5.39)$$

5.2.3 General expressions for the heterogeneous chain

In single molecule experiments where magnetic tweezers are used to impose tension and twist, the boundary conditions of the chain correspond to a partially-clamped set up. In these experiments one end of the chain is fixed at the origin (clamped), but the other end is free to have transverse displacements instead of being constrained to lie in the the same axis as the origin (see figure 5.1). However, moments are applied such that $\theta^2 = v_x^2 + v_y^2$ at the two ends are zero, and hence we impose the constraint :

$$v_1 = v_N = 0. \quad (5.40)$$

The energy Eq. (5.27) can be rewritten in terms of $N - 2$ variables ranging from v_2 to v_N :

$$\begin{aligned} \beta E_F &\approx \beta \sum_{i=2}^{N-1} \{ (\kappa_{i-1} + \kappa_i + q_i) v_i^2 \} - 2\beta \sum_{i=2}^{N-2} \kappa_i v_i v_{i+1} + \vec{\Lambda} \mathbf{T} \vec{\Lambda}^T - \beta \sum_{i=1}^N m_i \\ &= \beta \sum_{i=1}^{N-2} \{ (\kappa_i + \kappa_{i+1} + q_{i+1}) v_{i+1}^2 \} - 2\beta \sum_{i=1}^{N-3} \kappa_{i+1} v_{i+1} v_{i+2} + \vec{\Lambda} \mathbf{T} \vec{\Lambda}^T - \beta \sum_{i=1}^N m_i \\ &= \vec{\Theta} \mathbf{M} \vec{\Theta}^T + \vec{\Lambda} \mathbf{T} \vec{\Lambda}^T - \beta \sum_{i=1}^N m_i \end{aligned} \quad (5.41)$$

where $\vec{\Theta} = [v_2, \dots, v_{N-1}]$ and \mathbf{M} is $N - 2$ -dimensional tridiagonal matrix:

$$\mathbf{M} = \beta \begin{bmatrix} (\kappa_1 + \kappa_2 + q_2) & -\kappa_2 & 0 & \dots & 0 \\ -\kappa_2 & (\kappa_2 + \kappa_3 + q_3) & -\kappa_3 & \dots & 0 \\ 0 & -\kappa_3 & \dots & \dots & 0 \\ 0 & 0 & \dots & (\kappa_{N-3} + \kappa_{N-2} + q_{N-2}) & -\kappa_{N-2} \\ 0 & 0 & 0 & -\kappa_{N-2} & (\kappa_{N-2} + \kappa_{N-1} + q_{N-1}) \end{bmatrix} \quad (5.42)$$

which can be written in compact form as:

$$[\mathbf{M}]_{ij} = \{ \beta (\kappa_i + \kappa_{i+1} + q_{i+1}) \delta_{ij} - \beta \kappa_{i+1} \cdot \delta_{(|i-1|, 1)} \} \quad \text{for } 1 \leq (i, j) \leq N - 2. \quad (5.43)$$

The partition function is the integral:

$$Z = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp[-\beta E_F] d\vec{\Theta} d\vec{\Lambda}, \quad (5.44)$$

where $d\vec{\Theta} = [dv_2, dv_2, \dots, dv_{N-1}]$ and $d\vec{\Lambda} = [d\Delta_1, d\Delta_2, \dots, d\Delta_N]$. In general, the exact limits of the integral should be $v_i = \pm\pi$, but note that the term βF reaches its minimum value at $v_i = 0$. So, applying the Laplace method to approximate the integral by extending its limits to $\pm\infty$, we get

$$Z = \exp \left[\beta \left(\sum_{i=1}^N m_i \right) \right] \int_{-\infty}^{+\infty} \exp \left[- \left(\vec{\Theta} \mathbf{M} \vec{\Theta} + \vec{\Lambda} \mathbf{T} \Lambda^T \right) \right] d\vec{\Theta} d\vec{\Lambda} \quad (5.45)$$

$$= \exp \left[\beta \left(\sum_{i=1}^N m_i \right) \right] \sqrt{\frac{\pi^{N-2}}{\det \mathbf{M}}} \sqrt{\frac{\pi^N}{\det \mathbf{T}}}. \quad (5.46)$$

We define:

$$\det \mathbf{T} = \beta^N \times \prod_{i=1}^N c_i = \beta^N \times \det \hat{\mathbf{T}}, \quad (5.47)$$

$$\det \mathbf{M} = \beta^{N-2} \times \prod_{i=1}^{N-2} \lambda_i = \beta^{N-2} \times \det \mathbf{A}, \quad (5.48)$$

where λ_i is a sequence containing the information about the bending modulus κ_i and the parameter q_i

$$\lambda_1 = \kappa_1 + \kappa_2 + q_2, \quad \lambda_i = \kappa_i + \kappa_{i+1} + q_{i+1} - \frac{\kappa_i^2}{\lambda_{i-1}}, \quad (i = 2, 3, \dots, N-2). \quad (5.49)$$

Then it is a straightforward calculation to show that the logarithm of the partition function is given by:

$$\begin{aligned} \ln Z &= \left(\sum_{i=1}^N m_i \right) \beta + \frac{1}{2} \log(\pi^{2N-2}) - \frac{(2N-2)}{2} \log \beta \\ &- \frac{1}{2} \log \left(\prod_{i=1}^{N-2} \lambda_i \right) - \frac{1}{2} \log \left(\prod_{i=1}^N c_i \right), \end{aligned} \quad (5.50)$$

so that the free energy is given by Eq. (5.34) and the average extension $\langle x \rangle = L - \partial G / \partial F$ is:

$$\langle x \rangle = L - \frac{1}{\beta} \frac{\partial \log(\det \mathbf{M})}{\partial F} = L - \frac{1}{\beta} \frac{\partial \log(\det \mathbf{A})}{\partial F} = L - \frac{1}{\beta} \sum_{i=1}^{N-2} \left[\frac{\partial \lambda_i}{\partial F} \lambda_i^{-1} \right]. \quad (5.51)$$

Similarly, the linking number (average number of turns) is $\langle 2\pi n \rangle = -\partial G / \partial M_3$:

$$\langle 2\pi n \rangle = \sum_{i=1}^N \left[\frac{M_3 l}{C_i} + u_o l \right] - \frac{1}{\beta} \frac{\partial \log(\det \mathbf{A})}{\partial M_3} = \sum_{i=1}^N \left[\frac{M_3 l}{C_i} + u_o l \right] - \frac{1}{\beta} \sum_{i=1}^{N-2} \left[\frac{\partial \lambda_i}{\partial M_3} \lambda_i^{-1} \right]. \quad (5.52)$$

5.2.4 Results for the homogeneous chain

Next we present the results for a chain with homogeneous properties subjected to a force F and external torque M_3 . In writing the expressions below we have taken the limit as $N \rightarrow \infty$ and $l \rightarrow 0$ such that $Nl = L$. The complete derivations are shown in D.1. The average extension for the chain

in the torque controlled ensemble is given by :

$$\langle x \rangle = L - \frac{1}{\beta} \frac{\partial \log(\det \mathbf{A})}{\partial F} = L - \frac{k_B T}{2} \left(\frac{L}{Q} \coth \left[L \frac{Q}{A} \right] - \frac{A}{Q^2} \right), \quad (5.53)$$

where

$$Q^2 = AF - \frac{M_3^2}{4}, \quad (5.54)$$

which for large F (equivalently large Q) becomes:

$$\langle x \rangle = L - \frac{k_B T}{2} \frac{L}{Q}. \quad (5.55)$$

In a similar fashion the linking number vs. applied torque relation for a homogeneous chain is given by:

$$\langle 2\pi n \rangle = \frac{M_3 L}{C} + u^\circ L + M_3 \frac{k_B T}{4A} \left(\frac{L}{Q} \coth \left[L \frac{Q}{A} \right] - \frac{A}{Q^2} \right), \quad (5.56)$$

which for large F (equivalently large Q) becomes:

$$\left\langle \frac{2\pi n}{L} \right\rangle = M_3 \left(\frac{1}{C} + \frac{k_B T}{4QA} \right) + u^\circ. \quad (5.57)$$

Equations Eq. (5.55) and Eq. (5.57) from our model are the same as those obtained for the helical worm like chain by [40, 56] in the high force regime (setting the shift of the internal link $u^\circ = 0$).

Our model allows us to obtain similar relations in the case of an ensemble where $n = \Delta L k$ is the controlled variable and the energy of the system is given by:

$$E = \int_0^L \frac{1}{2} [A\kappa^2 + Cu_3^2] ds + F\Delta x + \lambda \left(2\pi\Delta L k - \int_0^L u_3 + u^\circ + [\psi'(1 - \cos\theta)] ds \right), \quad (5.58)$$

$$E = \int_0^L \hat{I}_n ds + 2\pi\Delta L k \lambda \quad (5.59)$$

$$\hat{I}_n = \frac{A}{2} (\theta'^2 + \psi'^2 \sin^2 \theta) - \frac{\lambda^2}{2C} - F \cos \theta - \lambda \psi'(1 - \cos \theta) + \frac{C}{2} \left[u_3 - \frac{\lambda}{C} \right]^2 - u^\circ \lambda + F. \quad (5.60)$$

where λ is the Lagrange multiplier enforcing the link constraint, which is obtained from minimizing the energy with respect to u_3 :

$$\lambda = Cu_3^{min}. \quad (5.61)$$

To obtain the response functions one can carry out the same analysis as done for the controlled torque ensemble, replacing M_3 by λ . In that case the average extension of a homogeneous chain is still given by the same expression Eq. (5.53) but Q is replaced by Q_n :

$$Q_n^2 = AF - \left(\frac{C\pi n}{L} - \frac{Cu^\circ}{2} \right)^2, \quad (5.62)$$

while the ensemble average torque is given by :

$$\frac{\langle M_3 \rangle}{2\pi} = \left[\frac{C}{L} - k_B T \frac{K_t^2}{4AL^2} \left(\frac{L}{Q_n} \coth \left[L \frac{Q_n}{A} \right] - \frac{A}{Q_n^2} \right) \right] n - \frac{Cu^\circ}{2\pi}. \quad (5.63)$$

Our model has the advantage that it can be useful not only for long DNA filaments, but for

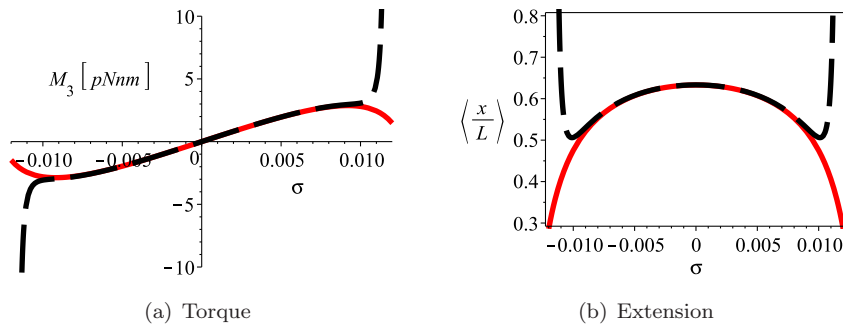


Figure 5.2: Comparison of our model with those valid in the high force long chain limit. a. External torque M_3 as a function of $\sigma = 10.5\Delta TwLk/N$, where N is the number of base pairs and σ is normalized linking number (see D.4). Graph corresponds to a short DNA molecule of $N=500$ base pairs subjected to a low tension force $F = 0.1\text{pN}$. b. Normalized extension $\langle x \rangle / L$ as a function of σ for a moderately short DNA molecule of $N=500$ base pairs subjected to a $F = 0.1\text{pN}$. The red solid line represents the full solution given by expressions Eq. (5.53) and Eq. (5.56), the black dashed line corresponds to the high force limit formulae Eq. (5.55) and Eq. (5.57). We have used the moduli corresponding to B-DNA (bending modulus $A = 45k_B T$ and twisting modulus $C = 100k_B T$). We have plotted the solution for $Q_n > 0$. As $Q_n^2 \rightarrow 0$, the solutions predicted by Eq. (5.55) and Eq. (5.57) using the high force (long chain) approximation diverge toward ∞ .

moderately short chains. Fig.5.2 shows the comparison of a fluctuating chain under tension and twist using the full solution from our methods and the approximate high force-long chain solution [56, 134, 149]. For small values of σ (normalized linking number) the two models agree very well as they should, but for larger σ the models of Marko and Moroz & Nelson diverge, where as our's produces the correct result. Note that the plot of $\langle x/L \rangle$ vs. σ for given F looks symmetric about $\sigma = 0$. Moroz and Nelson fitted the value of the twisting modulus of DNA using this type of extension-rotation curves. The symmetry is lost when twist-stretch coupling is taken into account. In experiments the asymmetry becomes obvious at large forces. We illustrate the effect of twist-stretch coupling by plotting the results for double-stranded RNA (dsRNA).

Over the past decade there has been an increasing interest in the study of dsRNA due to the its new found role interacting with numerous proteins (helicases, polymerases and nucleases) [23]. Although dsRNA is the genetic material in a wide variety of viruses and has a regulatory role in the cell, our understanding of its mechanical behavior is still lacking. Single molecule experiments are now being applied to dsRNA with the objective of measuring its mechanical properties. [23] provided measurements of the bending persistence length of dsRNA using two different single molecule techniques – atomic force microscopy and magnetic tweezers. [24] studied the overstretching transition of both molecules, dsRNA and dsDNA, and found that dsRNA has a lower stretch modulus than B-DNA and that its mechanical properties are highly dependent on electrostatic contributions. More recently there has been some interest in probing the mechanical response of dsRNA with a torsional constraint, where it has been seen that dsRNA shortens when overwound [150]. To our knowledge there is no documented twist modulus for dsRNA from single molecule experiments, hence we use the dsRNA twist modulus obtained in [25] using molecular dynamic simulations. We focus on predicting the qualitative behavior of dsRNA as a function of the twist-stretch coupling. In Fig. 5.3 we present the results for a homogeneous chain subjected to external tension F and turns n under partially clamped boundary conditions (see equations Eq. (D.84) and Eq. (D.85) in the Appendix). When the twist-stretch coupling modulus g is positive the slope of the extension-rotation curve near $\sigma = 0$ is negative and vice versa. Similarly, for positive g the curve for $\langle M_3 \rangle$ vs. σ shifts to the right and vice versa, It will be interesting to see if experiments agree quantitatively with our predictions.

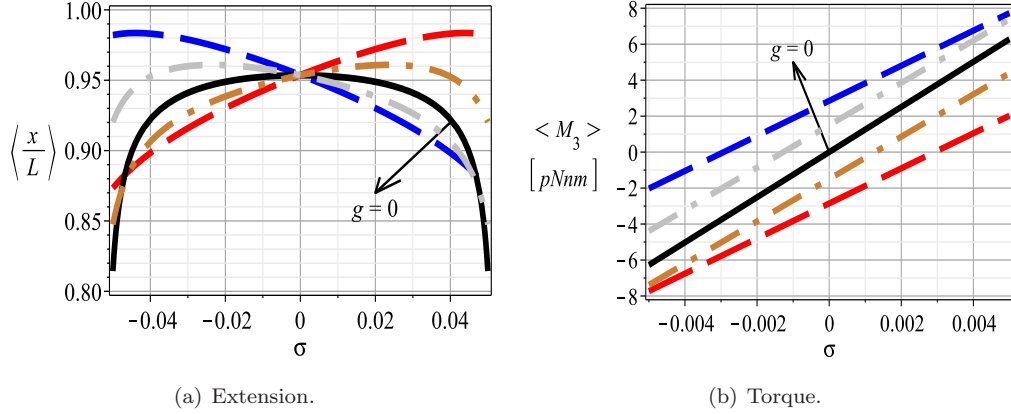


Figure 5.3: Duplex RNA behavior as a function of the twist-stretch coupling term g . In (a) we plot the normalized extension $\langle x \rangle / L$ as function of the normalized linking number σ (see D.4) and in (b) we plot the average torque $\langle M_3 \rangle$ as a function of σ . In the calculations we have used $F = 5\text{pN}$ which is a physiologically relevant force, $A = 63k_B T$ and $l_{RNA} = 0.279\text{nm}$ [23]. For the stretch modulus we use $S = 500\text{pN}$ which corresponds to $A = 61k_B T$ at 150mM monovalent salt concentration as documented in [24]. For C we use the value from molecular dynamic simulations $C_1 = 191k_B T$ [25]. The black solid lines in (a) and (b) correspond to a dsRNA chain where $g = 0$. In (a) the blue (dashed) curve ($g = 300\text{pNm}$) and the gray (dash-dot) curve ($g = 160\text{pNm}$) show a negative slope of the extension as function of the number of turns with the slope increasing as g increases. The red (dashed) curve ($g = -300\text{pNm}$) and brown (dash-dot) curve ($g = -160\text{pNm}$) show a positive slope of the extension as function of the number of turns with a larger magnitude of the slope as g decreases. Double stranded B-DNA has a negative twist-stretch coupling $\approx -85\text{pNm}$ ([26]). In (b) we see that for positive g the torque $\langle M_3 \rangle$ slightly shifts to the right and for negative g the torque $\langle M_3 \rangle$ shifts slightly to the left. As the magnitude of g increases, the slope of $\langle M_3 \rangle$ as a function of σ decreases. This is expected since the effective torsional modulus is $\tilde{C} = C - g^2/S$.

5.2.5 A two state heterogeneous chain in the n-ensemble.

In this section we present the results for a special case of a heterogeneous chain consisting of only two states using the general energy expression Eq. (5.1) which includes the effects of bending, twist and stretch. We give analytical expressions for the average extension and torque for a chain subjected to tension F , applied turns n and under partially clamped boundary conditions. The derivations are presented in D.3. The subscripts $j = 1, 2$ denote the state to which each variable correspond. A_j is the bending modulus, C_j is the twist modulus, S_j represents the stretch modulus, g_j the twist-stretch coupling modulus and $\tilde{C}_j = C_j - g_j^2/S_j$. The total length $L = L_1 + L_2$ is given by:

$$L_1 = Nl_1x_1, \quad L_2 = Nl_2x_2, \quad x_1 + x_2 = 1 \quad (5.64)$$

where N is the total number of segments, l_j is the length per segment and x_j is the fraction of segments in each state.

The average extension $\langle x \rangle$ and torque $\langle M_3 \rangle$ are:

$$\langle x \rangle = L - \frac{\partial E_{min}}{\partial F} - k_B T \left(X - \frac{1}{2} \left[\nabla_1 \frac{A_1}{Q_1^2} + \nabla_2 \frac{A_2}{Q_2^2} \right] \right), \quad (5.65)$$

$$\langle M_3 \rangle = \frac{1}{2\pi} \frac{\partial E_{min}}{\partial n} + k_B T \left[T_n - \tilde{\Lambda}_1 \frac{A_1}{2Q_1^2} - \tilde{\Lambda}_2 \frac{A_2}{2Q_2^2} \right], \quad (5.66)$$

where:

$$\begin{aligned}
-\frac{\partial E_{min}}{\partial F} &= \frac{L_1}{C_1} \left[\left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) F - \frac{2\pi n - U^o}{\varpi} \right] \left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) \\
&+ \frac{L_2}{C_2} \left[\left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) F - \frac{2\pi n - U^o}{\varpi} \right] \left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) \\
&+ \left(\frac{L_1}{S_1} + \frac{L_2}{S_2} \right) F - 2\pi n \frac{\Phi}{\varpi},
\end{aligned} \tag{5.67}$$

$$\frac{1}{2\pi} \frac{\partial E_{min}}{\partial n} = \frac{L_1}{\varpi \tilde{C}_1} \left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) F + \frac{L_2}{\varpi \tilde{C}_2} \left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) F + \frac{2\pi n + F\Phi - U^o}{\varpi}, \tag{5.68}$$

$$\begin{aligned}
X &= \frac{\left[\left(L_1 \nabla_1 \frac{Q_2}{Q_1} + L_2 \nabla_2 \frac{Q_1}{Q_2} \right) \cosh(h_1) \cosh(h_2) + (L_1 \nabla_1 + L_2 \nabla_2) \sinh(h_1) \sinh(h_2) \right]}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]} \\
&+ \frac{\left(\nabla_1 \frac{A_1}{Q_1} \right) \sinh(h_2) \cosh(h_1) + \left(\nabla_2 \frac{A_2}{Q_2} \right) \sinh(h_1) \cosh(h_2)}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]},
\end{aligned} \tag{5.69}$$

$$\varpi = \frac{L_1}{\tilde{C}_1} + \frac{L_2}{\tilde{C}_2}, \quad \Phi = \frac{g_1 L_1}{\tilde{C}_1 S_1} + \frac{g_2 L_2}{\tilde{C}_2 S_2}, \quad U^o = u_1^o L_1 + u_2^o L_2 \tag{5.70}$$

$$\nabla_j = \left[1 - \frac{\Phi}{2A_j} \left(\frac{2\pi n - U^o + F\Phi}{\varpi^2} \right) \right], \quad h_j = L_j \frac{Q_j}{A_j}, \quad Q_j^2 = A_j F - \frac{1}{4} \left(\frac{2\pi n}{L} + \frac{g_j F}{S_j} \right)^2 \tag{5.71}$$

for $j = 1, 2$. And T_n is equal to $-X$ replacing ∇_j with $\tilde{\Lambda}_j$:

$$\tilde{\Lambda}_j = \frac{1}{2A_j} \left(\frac{2\pi n + F\Phi - U^o}{\varpi^2} \right) \tag{5.72}$$

5.3 Applications: Comparison with experiment and predictions

5.3.1 Small molecule EtBr binding to B-DNA

[27] recently performed torsionally constrained experiments where small-molecules (drugs) are bound to the double stranded B-DNA at moderate to low forces $0.2 < F < 10$ pN and biologically relevant torques. Here we focus on one of their experiments involving the ligand intercalator EtBr (Ethidium Bromide), which has been shown to interfere in several biological processes such as recombination, replication, gene expression and has been used in the production of some anticancer pharmaceuticals [6, 27, 28]. Experiments show that when EtBr binds, it lengthens, softens and unwinds the DNA molecule by $\sim 27.3 \pm 1^\circ$ per intercalation event. Furthermore, experiments show that EtBr delays the accumulation of torsional stress, which can be explained by a reduced torsional modulus and torque dependent binding [27]. Evidence of a reduced twist modulus C comes from the broadening of the extension-rotation curves [27, 28]. In this section we apply our model of a heterogeneous chain with two states to intercalation experiments, where state 1 is made of B-DNA and state 2 is made of EtBr bound B-DNA. We aim to provide some insight into the mechanics of ligand binding for a better understanding of the effects of small molecule drugs in the cellular environment.

We let $x_2 = a$ be the fraction of segments where the small molecule is bound to the basepairs. From the experiment [27] it is clear that a is a function of the concentration $c_o [\mu M]$ of EtBr. When

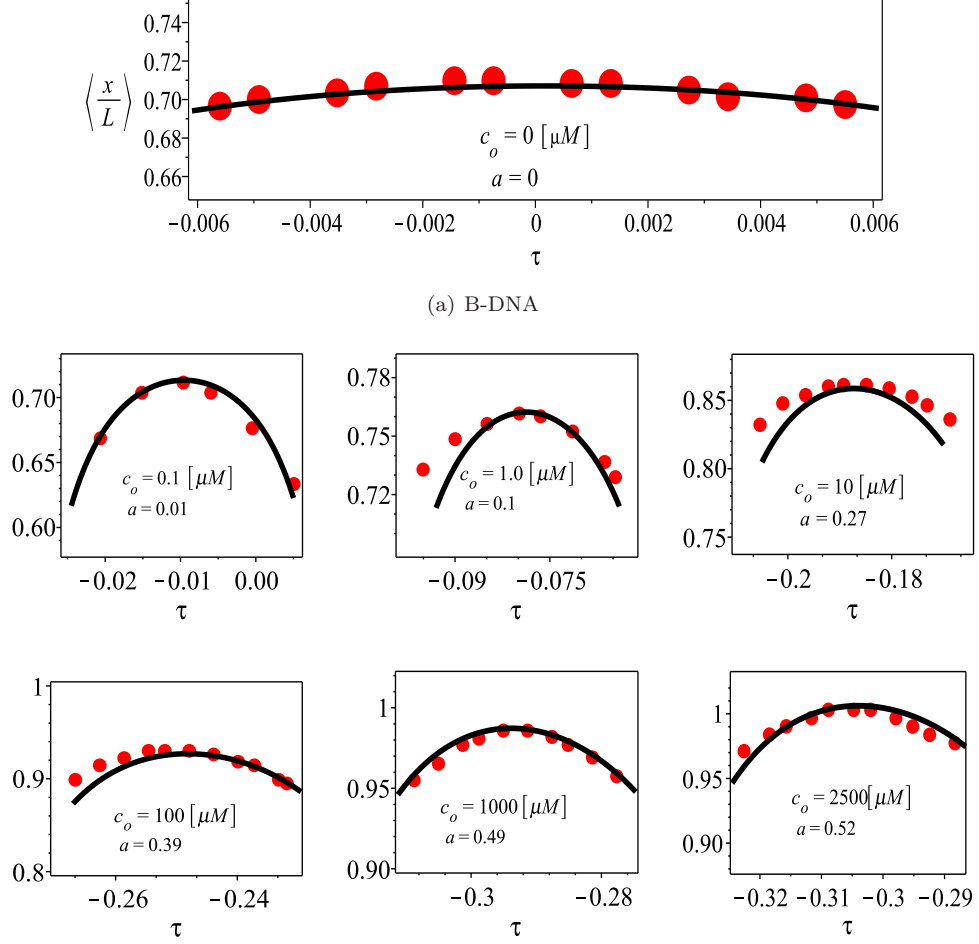


Figure 5.4: Normalized extension $\langle x \rangle / L$ as a function of the normalized twist τ . (a) The top graph shows the result of our model using the properties of B-DNA in table 5.1. We plot the solution up to the point where Lipfert et al. saw the formation of plectonemic structures. In the remaining graphs we compare the results of our two-state model (black solid line) with the experiments in [27] (red circles). The experiments were performed using $N = 20.6\text{kp}$ and $F = 0.25\text{pN}$. The fraction a of the DNA where EtBr has bound is presented in each figure as a function of c_o . The values of the properties of each state, B-DNA and EtBr bound DNA, are presented in table 5.1. Using our model we have fitted the shift in internal twist per intercalated EtBr molecule as $\Delta t w_{EtBr} \approx -0.48\text{rad}$ per intercalation (-28°) which agrees with the values presented in [27].

a EtBr molecule intercalates, it increases the length of DNA by $\Delta z_{EtBr} = 0.34\text{nm}$ per intercalation [27, 28]. Therefore, $l_2 = l_{EtBr} = l_1 + \Delta z_{EtBr} = 0.68\text{ nm}$ per segment.

We define the local twist in state 1 as $u_{3,1}^{min} = w_o \tau$, where $w_o = 2\pi / (10.5 l_1) = 1.77\text{nm}^{-1}$, such that τ is a normalized twist (10.5 is approximately the number of base pairs per helical turn in B-DNA). Since state 1 is our reference state we set $u_1^o = 0$, and since the binding of EtBr molecules to the DNA base pairs induces a shift in the local twist we have $u_{3,2}^{min} = w_o \tau - u_2^o$, such that :

$$n = 2\pi \Delta L k^{EtBr} = u_{3,1}^{min} L_1 + (u_{3,2}^{min} + u_2^o) L_2 = w_o \tau L(c_o), \quad (5.73)$$

$$L(a) = L_1 + L_2 = L_o + a(c_o) \Delta z_{EtBr} N = [1 + a(c_o)] L_o, \quad (5.74)$$

since $\Delta z_{EtBr} \approx 0.34\text{nm}$ which is the length l_1 of B-DNA. $L_o = Nl_1$ is the length of the unperturbed B-DNA and $L(a)$ is the length of the DNA with intercalated EtBr molecules for a given fraction $a(c_o)$ and Δz_{EtBr} (which are independent of the applied force F). We point out that $L(a)$ is not the average extension $\langle x \rangle$ measured in the experiments which includes the effects of stretch (force) and fluctuations. We make this distinction to differentiate $a(c_o)$ (fraction of DNA to which EtBr molecules have bound independent of the force) in Eq. (5.74) with the definition of γ_f given by [6], for $F > 10\text{pN}$:

$$\langle x(F, c_o) \rangle = (1 + \gamma_f) \langle x(F, 0) \rangle, \quad (5.75)$$

where $\langle x(F, c_o) \rangle$ is the extension of the drug-DNA complex at a given force and concentration and $\langle x(F, 0) \rangle$ is the extension in the absence of the drug. This definition of γ_f (which [6] call the fractional number of molecules intercalated per base pair) is effectively lumping the effects of fluctuations, force and the drug concentration. We do not work with γ_f and rather define $a(c_o)$ to be the fraction of DNA segments to which EtBr molecules have bound for a given concentration. We do so because in torsionally constrained DNA using γ_f does not capture the experimental data. Let us illustrate this by using the torsional experiments in [27].

It is well known that EtBr intercalation leads to unwinding of the DNA helix. This is evident in a negative offset in the number of turns in the extension-rotation experimental curves [27, 28]. As mentioned in [27] the offset in extension-rotation curves is a function of the concentration of EtBr, but does not vary with the force F . From the data series, with constant $c_o = 2500\mu\text{M}$, for different forces in Figure 4 in Lipfert et al., the fit of Eq. (5.75) leads to $\gamma_f \approx 0.5$ at $F = 0.25\text{pN}$ and $\gamma_f \approx 0.7$ at $F = 10\text{pN}$. This increase in γ_f as a function F would mean that more molecules are intercalated per base pair. Due to the unwinding caused by EtBr this would lead to an extension-rotation curve at $F = 10\text{pN}$ with a noticeable larger offset with respect to the extension-rotation curve at $F = 0.25\text{pN}$, which is not seen [27]. The same conclusion about the inappropriateness of using γ_f can be reached by looking at the data of [28] for $F = 0.3\text{pN}$ and $F = 0.8\text{pN}$. The above discussion holds if one takes the unwinding shift Δtw_{EtBr} per intercalation event to be a constant independent of the loading parameters. Letting Δtw_{EtBr} vary as a function of F could explain the same shift in the extension-rotation curves, but we do not have any supporting evidence that this might be the case.

So, instead of using γ_f we say that EtBr bound DNA has a different stretch modulus and twist-stretch coupling modulus from B-DNA. It has been suggested that in the presence of intercalators there is ligand-ligand repulsion that is likely mediated by ligand-induced structural deformations of the dsDNA [6]. Due to the structural deformations, the force can in theory relieve this long-range ligand-ligand repulsion, leading to a stretching of the drug-DNA complex. Several groups [6, 27, 28] have studied the mechanical properties of EtBr-saturated B-DNA using single state modified versions of the worm like chain (or WLC) model [48]. From their fits they have concluded that the bound ligand reduces the bending, twisting and stretching modulus significantly. But we note that the values of the mechanical properties obtained from those fits significantly disagree. For instance, in [27] the authors presented the bending modulus A_2 ranging from $30 - 50[k_B T]$ (depending on c_o), [6] fitted $A_2 \sim 5.7k_B T$ and [28] $A_2 \sim 54k_B T$. Although the one state WLC model does seem to capture the qualitative mechanical response, we think that making use of our heterogeneous rod model can give a more accurate picture of dsDNA in the presence of this drug.

We next describe the procedure used to reproduce the experimental data in [27]. Since not all the properties of DNA in the presence of EtBr are known we needed to fit certain parameters. To do so we divided the fitting procedure in three steps:

- First fit the value of a to the torsionally-unconstrained experimental data presented in Figure 2a in [27] using expression Eq. (5.65) but setting $C_j = g_j = 0$. In the fitting we have used B-DNA properties given in the literature [56, 151] (see table 5.1 for values) and for the EtBr state we have used $A_2 = 30k_B T$. We use this as the bending modulus since it was obtained in [27] at $c_o = 2500\mu\text{M}$ where the intercalation of EtBr to DNA has reached saturation, and

it corresponds to the low-to-moderate force range $F \leq 10\text{pN}$. Also since the constant force used in the torsionally-unconstrained experiments in [27] was very low $F = 0.25\text{pN}$, the stretch effects are almost negligible. We found that the fit of $a(c_o)$ did not vary (by much) by using a wide range of $S_2 \in [0, 1220]\text{pN}$.

- Next we moved to the torsionally constrained problem, while still in the low force regime but with a as a function of c_o . We fitted expression Eq. (5.65) to the $F = 0.25\text{pN}$ data set at $c_o = 2500\mu\text{M}$ (see Figure 3 in [27]). From the fit we obtained $u_2^o = u_{EtBr}^o = -0.71$ rad per nm, a reduced twist modulus $C_2 = 30k_B T$ (as expected). The shift in the internal linkage per intercalated EtBr molecule $\Delta t w_{EtBr} = u_{EtBr}^o l_{EtBr} \approx -0.48\text{rad}$ per intercalation (-28°) which agrees with the values presented in [27].
- To see the effects of stretch we next used the $F = 10\text{pN}$ data set at $c_o = 2500\mu\text{M}$ in [27] from which we got $S_2 = 76\text{pN}$ and a positive $g_2 = 42\text{pNnm}$. The value of g_2 obtained in the fit falls within previous measurements of B-DNA at large forces [152].

After recording all the properties of state 2 (see table 5.1) and fraction a as a function of c_o , we use equation Eq. (5.65) to *predict* the rest of the data series for $c_o = [0.1, 1, 10, 100, 1000]\mu\text{M}$ at $F = 0.25\text{pN}$ and for $F = [0.5, 1, 3]\text{pN}$ at $c_o = 2500\mu\text{M}$ in [27] with no fitting. The results comparing our model with experiments in [27] are shown in Fig. 5.4 (for varying c_o and constant $F = 0.25\text{pN}$) and in Fig. 5.5 (for varying F and constant $c_o = 2500\mu\text{M}$). Furthermore in Fig. 5.6 we present the results for both the average extension and torque from our theory using equations Eq. (5.65) and Eq. (5.66) with no fitting parameters and compare it with the experimental results in [28]. It is evident from Figs. 5.4, 5.5 and 5.6 that our torsionally constrained two-state model accurately describes a wide range of drug concentrations and forces.

For the sake of completeness we have also performed calculations allowing $a(c_o, F, \tau)$ to be a function of the controlled parameters. Sample plots of the results are shown in D.6. As mentioned earlier, for $F = 0.25\text{pN}$ the effects of the stretching are not evident (see, for instance, figure D.1 in the Appendix where $S_2 = 1220\text{pN}$ and compare it to figure 5.4 where $S_2 = 76\text{pN}$ for the same a). But for larger F the stretch modulus becomes important. Hence, when using $S_2 = 1220\text{pN}$ the fraction a needs to vary as function of F (see Fig. D.2 in Appendix.) to describe the data accurately.

Property	State 1 - B-DNA	State 2 - EtBr
l_j [nm/bp]	0.34	0.68
A_j [$k_B T$][nm]	50	30
C_j [$k_B T$][nm]	90	30
g_j [pNnm]	-84	42
S_j [pN]	1220	76

Table 5.1: Properties of B-DNA and EtBr bound DNA. [6] used a WLC model to fit their experimental data of a torsionally unconstrained molecule and obtained a reduced stretch modulus $S_2 \sim 220\text{pN}$.

Predictions for other small molecules. Besides EtBr, there is a large number of other small molecules that bind to B-DNA with similar effects. For instance, [153] recently studied Rad51-DNA interactions and found stretching and underwinding of DNA in the presence of the bound protein leading to similar behavior as seen in the EtBr-DNA experiments. [27] also performed torsionally constrained experiments using netropsin (minor groove binder) and TPT (topotecan is a topoisomerase IB inhibitor). The extension-rotation curves with netropsin and TPT had analogous characteristics to the ones using EtBr. To our knowledge a detailed study of these molecules, using different forces and torques, is not yet available in the literature. Nevertheless, we would like to extend our treatment of small molecule binding by making predictions for DNA behavior in the presence of netropsin, because its binding leads to overwinding of B-DNA [27, 154].

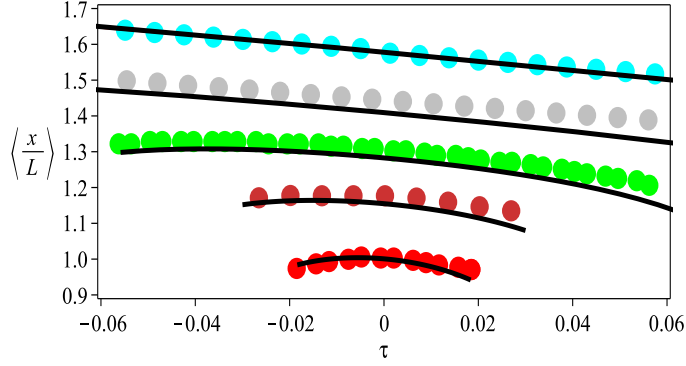


Figure 5.5: EtBr effects on rotation extension behavior at low and high forces. The graphs shows the normalized extension as a function the normalized twist τ for a set of different forces. Circles correspond to data from [27]. Forces used going from bottom to top of the graph are $F = [0.25, 0.5, 1, 3, 10]$ pN. We have plotted the solution up to the point where there is no formation of plectonemes as given by [27]. Properties used are presented in table 5.1.

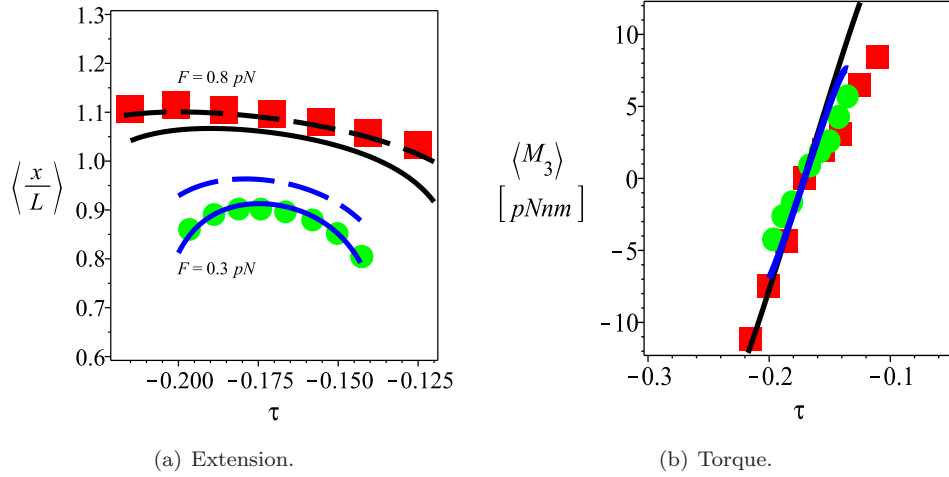


Figure 5.6: Blue lines correspond to $F = 0.3$ pN and black lines to $F = 0.8$ pN. Green circles and red squares are the data series from [28] for $F = 0.3$ pN and $F = 0.8$ pN respectively. For the solid lines we have used the same properties as in Fig. 5.4. Dashed line calculated using $A_2 = 54k_B T$ which is the value obtained for $c_o = 10\mu M$ fitted by [28] using a single state WLC model. In (b) the blue solid line representing the torque $\langle M_3 \rangle$ at $F = 0.3$ pN has a slightly smaller slope than the black solid line (torque at $F = 0.8$ pN). Here we have not fitted any parameters; we use the values of a from comparing our model to [27] experiments. The extension $\langle x \rangle$, torque $\langle M_3 \rangle$ and the shift in τ seen in the experiments of [28] are predicted accurately with our model.

The netropsin experiments in [27] were performed using a 20.6kbp DNA molecule at a force of $F = 0.25$ pN. At $c_o = 100\mu M$ the solution is saturated with netropsin and at this point the shift in the number of turns due to overwinding is measured to be $\Delta T w_o = 40$ [27]. The experiments show that there is no significant change in extension, so we assume $l_2 \approx l_1$, and no significant change in the bending persistence length, so we take $A_2 \approx A_1$. [154] reported that netropsin overwinds the molecule by $\Delta t w_{ne} = l_2 u_2^2 = 0.06$ radians per intercalation event. Using this value for the shift in internal link we obtain the fraction of the DNA to which netropsin molecules have bound at

saturation:

$$a_{ne} = \frac{2\pi\Delta T w_o}{\Delta t w_{ne} N} \approx 0.2. \quad (5.76)$$

Since netropsin overwinds DNA by a small fraction, we expect that the twisting modulus C_2 of the EtBr-DNA complex to be slightly larger than the modulus C_1 of B-DNA, while the stretch modulus S_2 and the twist-stretch modulus g_2 should not change significantly. In Fig. 5.7 we make some predictions for the extension-rotation curves in the presence of netropsin at saturation for different forces. In agreement with [27] at $F = 0.25\text{pN}$ the netropsin-DNA curve is shifted to the right. At such a low force the effects of the stretch modulus are not apparent. The effects of a modified S_2 or g_2 become significant at higher forces. If the mechanical properties of the netropsin-DNA molecule do not significantly change with respect to B-DNA we expect to see the blue-dashed extension-rotation curves in experiment (see fig. 5.7), while if $S_2 > S_1$ we expect to see the gray-dash-dotted curves. Finally, we point out that since B-DNA already has a very large twist modulus $C_1 \approx 90k_B T$, increasing the value of $C_2 \sim 2C_1$ does not modify the extension-rotation curves by much. In fact, since $a_{ne} = 0.2$ at netropsin saturation (small fraction), the slope of the torque M_3 vs. turns n , in the presence or absence of netropsin is approximately the same (not shown).

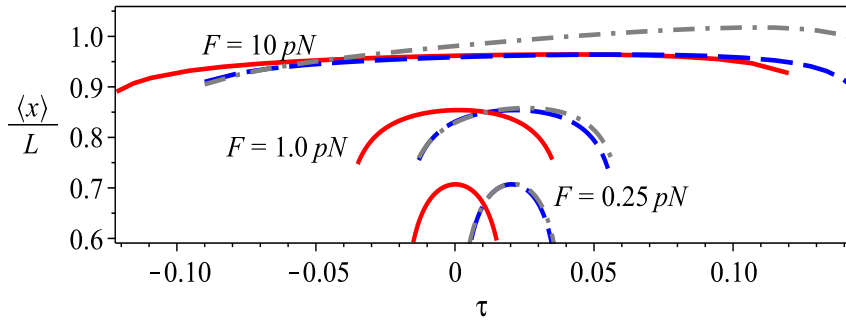


Figure 5.7: Predictions of the effects of netropsin on rotation extension behavior of DNA at low and high forces. The graphs show normalized extension as a function the normalized twist τ for three different forces. State 1 is B-DNA and state 2 is DNA-netropsin at saturation $c_o = 100\mu M$ [27]. At saturation the fraction of the DNA chain where netropsin has bound is $a_{ne} = 0.2$. The red-solid curves are calculated for state 1 (properties for B-DNA are presented in table 5.1). For state 2 we have used $A_2 = A_1$ and $l_2 = l_1$ since netropsin has no significant effect on the extension and bending persistence length [27]. The blue-dashed line represents state 2 using $g_2 = g_1$, $S_2 = S_1$ and $C_2 = 110k_B T$ slightly larger than C_1 (since we expect the overwound DNA to have a larger twisting resistance). The gray-dot-dashed line represents state 2 using $g_2 = g_1$ and $C_2 = 110k_B T$ but $S_1 > S_2 = 100\text{pN}$. If after netropsin binds to DNA all the properties of the complex do not significantly change, then the prediction is the blue-dashed curves. If the stretch modulus changes (gray-dash-dot curve), experiments at low forces would not be able to capture the effects (blue-dashed curve and gray-dash-dotted curve are almost on top of each other at $F = 0.25\text{pN}$), but at large forces $F \sim 10\text{pN}$ the extension rotation curves would show a significant difference.

5.3.2 Analysis of L-DNA mechanics

Using single molecule techniques, [57] reported a phase transition from B-DNA to a structure with -13 bp per turn, which they designated L-DNA because of its net left handed helicity. They saw that for negative n , B-DNA underwent a phase transition where the average torque during the transition was $\langle M_3 \rangle \approx -10\text{pNnm}$. They suggested that L-DNA could in fact be a combination of denaturated DNA and Z-DNA (a left handed structure with -12 bp per turn [102]). More recently, [31] continued a more detailed study of the L-DNA transition including effects of salt concentration and force. [31] analyzed the transitions of B-DNA to Z-DNA and from B-DNA to strand-separated DNA. They concluded that it is hard to reconcile L-DNA as being only a strand-separated structure. Some of

their findings also support the idea of L-DNA as a mixture of states. In a separate study [29] reached similar conclusions for the L-DNA transition. These authors made use of the worm like chain (or WLC) model [134] and a theory of phase transitions to extract the effective mechanical properties of L-DNA. The studies of [29, 31, 57] provided useful insight into this problem, but the question of whether L-DNA is a single state or a mixture remains unanswered. In this section we study both possibilities and suggest an experiment that could resolve the question. We start by assuming L-DNA as single state and extract its mechanical properties as done in [29], but including effects of stretch and twist-stretch coupling in the L-DNA phase. Next, based on the results of our fit to L-DNA as a single state and recent studies on the competition between strand-separated DNA and S-DNA [32, 33] we formulate the hypothesis that L-DNA is a mixture of Z-DNA and S-DNA. We also pursued the idea that L-DNA is a mixture of strand-separated DNA and Z-DNA, but with the documented mechanical properties of strand separated DNA in the literature we could not reproduce the experimental data for L-DNA.

L-DNA effective mechanical behavior: single state model. In this section we are going to focus on the regime where L-DNA has been formed and treat it as a single state. Hence, we do not use subscripts when writing L-DNA properties. In D.4.1 when dealing with the phase transition we let B-DNA be state 1 and L-DNA state 2. To be consistent with the experimental studies of [29] and [31], we present the results using the degree of supercoiling σ instead of the twist. Relations between ΔLk and σ are given in the D.4.

Now we outline the fitting procedure that we follow to obtain the mechanical properties of L-DNA presented in table 5.2:

- In [29] the authors set S to infinity as they mention that it could not be determined accurately over the range of forces explored. We take a different approach and take $A \approx 5.5k_B T$, $C \approx 20k_B T$ and $l_{L-DNA} \approx 1.35l_{B-DNA}$ as fitted in [29]. By doing so we find the value of the stretch modulus $S \approx 610\text{pN}$ by fitting the L-DNA curve that appears in Fig 4 (b) in [43]. At large forces the effect of bending modulus is not evident (since fluctuations are expected to be small). We found that using a slightly large $A = 9k_B T$ provided a better fit for $F < 30\text{pN}$ in the L-DNA experimental curve of [43] (not shown).
- From extension-rotation graphs in [29] we notice a trend that the hat curves (for $\sigma < \sigma_{end} \approx -1.8$) had a positive slope. This slope near the maximum point in the extension-rotation graphs is present for all the large forces. Although [29] use a small bending persistence length $\sim 3 - 5\text{nm}$, which would decrease the effective length of the molecule due to the higher contribution of the fluctuations, the theoretical hat-curves (at large $F > 24\text{pN}$) obtained by them are mostly flat at their peaks (see Supplementary Material in [29]). We found by using our model that the negative slope in the hat curves can be captured by considering a negative twist-stretch coupling $g = -100\text{pNnm}$ similar to that known for B-DNA [26].
- We plot the predictions of our model using Eq. (5.53) and Eq. (5.63) in Fig.5.8, where we have used $\sigma^o = -1.7$ (-15 helical repeat) [31]. As seen from the graph an effective single state model for L-DNA captures the mechanical behavior after DNA has made the transition into the new left handed structure.

Using the fitted mechanical properties of L-DNA from the single state model section we predict some features of the transition from B-DNA to L-DNA using a theory of phase transitions as described in D.4.1. We find $\sigma_{start} \sim 0.015$ the critical number of turns at which B-DNA starts to make the transition into L-DNA, $\sigma_{end} \sim 1.7$ the critical number of turns at which transition ends and the value of the torque $\langle M_3 \rangle \sim 10\text{pNnm}$ during the transition. Our prediction of the transition variables agree with the experimental evidence reported in [31, 57] and [29].

Using these for L-DNA mechanical properties, next we explore the idea that L-DNA is a mixture of two other states.

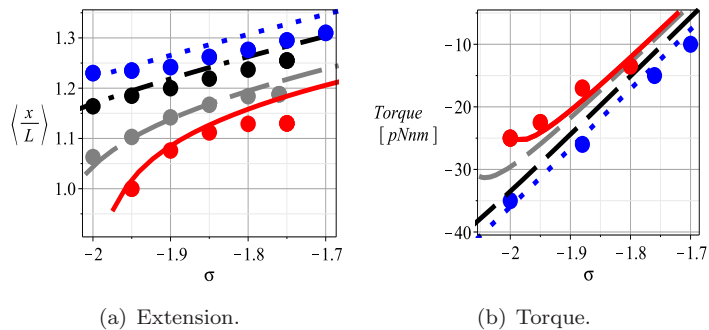


Figure 5.8: L-DNA as a single effective state. Circles correspond to the experimental data in [29]. Red-solid line corresponds to $F = 8.5\text{pN}$, gray-dashed to $F = 12\text{pN}$, black-dashed to $F = 24\text{pN}$ and blue-dotted to $F = 36\text{pN}$. We plot the solutions from $\sigma_{end} = -1.7$ (critical number of turns at which L-DNA is fully formed after the transition) up to $\sigma^* \approx 2$ where [29] saw the external torque plateauing again (probably the start of a new phase transition). We use only large forces $F > 8.5\text{pN}$, which is the regime where effects of stretch are noticeable and to avoid the formation of plectonemic structures upon the addition of twist. In (b) we only present the experimental torque for $F = 8.5\text{pN}$ and $F = 36\text{pN}$ to avoid over-crowding the graph. The other two experimental curves that are not shown fall within the blue and red lines.

L-DNA effective mechanical behavior: a two state model After the transition from B-DNA to L-DNA has begun there is no plateau in the torque vs. σ curve. Thus a phase transition driven by M_3 is not evident [29, 31]. Furthermore, from the rotation-extension curves in [29] it is possible to conclude that for $\sigma < \sigma_{end}$ (once L-DNA is formed) there is no $F_{critical}$ that is the signature of a first order phase transition in the $F - \langle x \rangle$ space. This is more evident from the L-DNA force vs. extension curve presented in [43]. Hence, if L-DNA is comprised of two states, the transition between them is not a function of the mechanical loadings. Although, the fractions x_1 and x_2 are independent of F and M_3 , these could in theory be functions of the salt concentration c_s . Next we present some arguments that lead us to hypothesize that L-DNA is a mixture of Z-DNA and S-DNA:

- Based on the fact that L-DNA is a left handed structure we consider Z-DNA as the main candidate for one of the phases because of its tight left-handed helicity $\sigma_1^o = -1.9$ (-12 bp/turn) [31, 102]. The measured parameters that we found for Z-DNA in the literature include its bending modulus $A_1 = 200k_B T$ and $l_1 = 0.37$ [31]. Based on our fitted parameters for L-DNA (in the previous section), if we take Z-DNA as one possible phase, then we must find another phase with $\sigma_2^o > \sigma_1^o$, $A_2 < A_1$ and $l_2 > l_1$. These characteristics are found both in melted DNA [29, 155] and S-DNA [30, 32, 156–158]. According to [31] melted DNA is destabilized by high ionic concentrations, while L-DNA is stabilized. Furthermore, [33] and [32] performed torsionally unconstrained experiments and found that as the ionic concentration increased S-DNA is strongly favored compared to melted DNA (or DNA bubbles). In the light of all these observations we decided to consider S-DNA as the second state in the mixture. We point out that L-DNA experiments of [29] were performed at a moderate salt concentration $c_s = 150\text{mM}$ NaCl.
- For S-DNA properties we use $A_2 \sim 7 - 12k_B T$, $S_2 \sim 3030\text{pN}$, $l_2 \sim 0.58\text{nm/bp}$ [30, 156, 157] and for the change in internal link we used $\sigma_2^o = -0.72$ (37.5 bp/turn) [32, 159].
- The rest of the parameters C_1, C_2, S_2, g_1 and g_2 were fitted to experiments keeping in consideration the properties of L-DNA as an effective single state. Since S-DNA has a slightly overwound ladder-like structure [30, 32] we expect it to have a reduced torsional stiffness $C_2 = 10k_B T$. For the stretch and twist stretch coupling moduli we fit the values to be $S_1 = 270\text{pN}$, $g_1 = -81\text{pNm}$ and $g_2 = -100\text{pNm}$.

- We plot the predictions of our model using Eq. (5.65) and Eq. (5.66) in Fig.5.9, where $x_1 = 0.77$ and $x_2 = 0.23$. As seen from the graph L-DNA = Z-DNA +S-DNA (a two state model) agrees well with experiment.

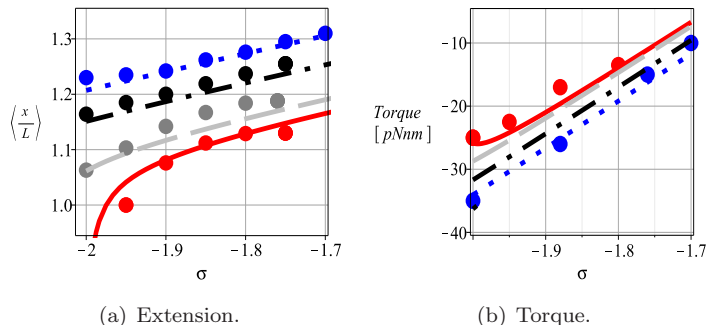


Figure 5.9: L-DNA as two states: S-DNA + Z-DNA. Circles correspond to the experimental data in [29]. Color code is the same as in Fig. 5.8. The properties used for S-DNA and Z-DNA are presented in table 5.2. In D.6 we show the results using $A_2 = 11.6k_B T$ (S-DNA bending modulus) as measured by [30].

Property	Single state - L-DNA	State 1 - Z-DNA	State 2 - S-DNA
l_j [nm/bp]	0.46	0.37	0.58
A_j [$k_B T$][nm]	9	200	7
C_j [$k_B T$] [nm]	20	23	10
g_j [pNm]	-100	-81	-122
S_j [pN]	610	3030	270
σ_i^o	-1.7	-1.9	-0.72
x_i	1	0.77	0.23

Table 5.2: L-DNA properties as a single state or as a mixture of two states: S-DNA + Z-DNA.

Just as in section 5.3.2 next we predict $\sigma_{start}, \sigma_{end}$ and $\langle M_3 \rangle$ for the B-DNA to L-DNA transition as described in D.4.1, but here we use the fitted mechanical properties of L-DNA as a mixture of S-DNA and Z-DNA. The method used in the calculations is described in D.4.1. From the calculation we get $\sigma_{start} \sim 0.013$, $\sigma_{end} \sim 1.7$ and $\langle M_3 \rangle \sim 10\text{pNm}$ during the transition. Our predictions of the transition variables using a two-state model is in good agreement with the experimental evidence reported in [31, 57] and [29].

From our analysis we have determined that L-DNA mechanical behavior can be accurately captured using a single state model that includes the effects of thermal fluctuation and twist stretch coupling. It can also be described by a two state model as a mixture of Z-DNA and S-DNA. Based on the experimental evidence presented in [29, 31, 43], in our model of L-DNA as a mixed state the fractions of Z and S comprising the mixture are independent of the mechanical variables F and M_3 . So, L-DNA is a mixture where the transition from Z to S might be driven by ionic concentration. We think that a plausible experimental method to test the prediction of L-DNA as a mixed state is to perform torsionally constrained single molecule experiments as in [29] and [31] at various salt concentration c_s . Assuming that the mechanical properties of Z and S DNA are approximately independent of c_s , if the extension-rotation and torque-rotation curves show a significant change then one can conclude that the fraction of S and Z has changed. By performing other experiments, it can be established if in fact the elastic properties of S and Z-DNA are approximately independent of c_s . This is certainly the case for B-DNA over the physiological range of ionic concentrations (0.1-0.5)M [24, 31, 95, 96].

From the experimental evidence and our study of L-DNA mechanical behavior, we predict that as the ionic concentration increases the fraction of S-DNA should increase, which in turn would lead to smaller σ_{end} (smaller left handed helicity), larger extension $\langle x \rangle$ and a slightly smaller slope in the torque-rotation graphs (see Fig. 5.10). We base our predictions on the following analysis. As c_s increases ΔG_Z^0 (ground state energy difference between Z and B-DNA, see D.4.3) decreases ([31]). For $c_s = 1.8M$ NaCl and $F = 3.2pN$ [31] measured $\Delta G_Z^0 = 0.04 \pm 0.02$ kcal/mol per bp which is about an order of magnitude smaller than at low ionic concentrations. Using these values we obtain B-Z transition torque $\langle M_3^{B-Z} \rangle \approx -0.25pNnm$, which is in good agreement with [31] experiments. In the case of S-DNA in a range of $c_s = 20 - 500mM$, there are no significant differences in ΔG_S^0 (ground state energy difference between S and B DNA) [30, 32]. Therefore, for a high ionic concentration at $F=10pN$ we use $\Delta G_Z^0 = 0.04 \pm 0.02kcal/mol$ per bp in the B-L transition as described in D.4.3. Keeping the fractions of Z-DNA $x_1 = 0.77$ and S-DNA $x_2 = 0.23$ as before, the transition torque comes out to be $\langle M_3^{B-L} \rangle \approx -5pNnm$ which is about half of the observed value in [31]. Using our two-state model we find that we require the S-DNA fraction to increase up to $x_2 = 0.4$ in order to recover the transition torque observed in [31] at high ionic concentration (see Fig. 5.10).

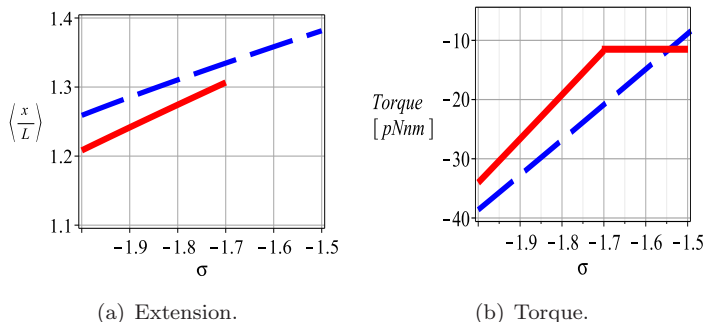


Figure 5.10: Predictions from our two state model of L-DNA as ionic concentration c_s varies. For moderate salt $c_s \sim 0.1M$ we use S-DNA fraction to be $x_2 = 0.23$ as fitted to experiments in [29, 31] and for high salt $c_s \sim 1.8M$ we use $x_2 = 0.4$ as fitted to [31] experiments. Red line correspond to moderate salt concentration and blue-dashed line corresponds to high salt. For illustrative purposes we have used one of the large force $F = 36pN$ used in [29] experiments. For moderate salt $\sigma_2 \approx -1.7$ while for a high salt $\sigma_2 \approx -1.5$. From (a) we see that the $\langle x \rangle$ peak increases as the concentration increases and in (b) we see a slightly smaller slope for higher c_s . The properties used for S-DNA and Z-DNA are presented in table 5.2.

5.4 Conclusions

We present a one-dimensional model for the statistical mechanics of a heterogeneous rod including the effects of bending, stretch, twist. Following the methods of Su and Purohit [49] we are able to account for heterogeneity and boundary effects efficiently when calculating the thermo-mechanical properties of fluctuating molecules. Using a special case of our theory, we have successfully applied a two-state model to experiments where a DNA molecule can be regarded as a mixture. We tackle two problems – one in which the reason for heterogeneous mechanical properties is the binding of EtBr (ethidium bromide) to DNA [27], and another, where the reason is a phase change from right-handed B-DNA to a left-handed DNA structure known as L-DNA [31].

The addition of EtBr leads to a reduction in the bending, torsional and stretching stiffness of DNA. So we treat the system as a two state model consisting of B-DNA and EtBr bound DNA where the fractions of the two states is a function only of EtBr concentration. We accurately reproduce experimental data [27, 28] for a wide range of concentrations and forces. Besides EtBr, there is a large number of other small molecules that bind to B-DNA with similar effects. For instance, single

molecule experiments using Rad51-DNA [153], netropsin and TPT [27] reveal modifications in DNA behavior that are similar to the ones found for EtBr binding. Our model can provide some insight for future studies using these and other ligands. For example, Lipfert et al. [27] performed experiments at low forces $F \approx 0.25\text{pN}$ and found that the bending and stretching properties of the DNA molecule remain unchanged after the addition of netropsin. Since the experiments were performed in the low force regime the complete picture of the netropsin-DNA mechanical behavior remains unclear. We have used our model to predict the behavior of the netropsin-DNA complex at larger forces. These predictions can be easily tested in new experiments.

In the case of L-DNA we started by using a single state model to capture the effective behavior of its left-handed structure. We found a set of effective L-DNA properties that gave strong agreement between our theory and the experiments of Sheinin et al. [29] and Lipfert et al. [27]. But, the studies of Bryant et al. [57], Oberstrass et al. [31] and Sheinin et al. [29] suggest that L-DNA is a mixture of DNA states. Using a two-state model we propose that L-DNA is in fact a mixture of Z-DNA and S-DNA. We determined that this two state model can also accurately describe the mechanical behavior of L-DNA. From experimental evidence in Bustamante et al. [43] and Sheinin et al. [29], we think that the fractions of Z and S-DNA comprising the mixture are independent of the force and torque. So, L-DNA is a mixture where the transition from Z to S might be driven by ionic concentration. To test the prediction that L-DNA is a mixture we suggest that one could perform torsionally constrained single molecule experiments [29] where the salt concentration is varied. From the experimental work in [30–32] we predict that as the ionic concentration increases the fraction of S-DNA should increase. A larger S-DNA fraction would lead to a structure with a smaller left handed helicity and a shorter plateau in the torque-rotation graphs. As the ionic concentration increases another signature to confirm our prediction is a larger peak in the extension-rotation graphs and a slightly smaller slope in the torque-rotation graphs (once L-DNA is fully formed).

Finally, we point out that our model is applicable only for sufficiently long chains; it can not describe the behavior of a few dozen DNA base-pairs. Large curvatures (or kinks) can develop in the molecule at these short length scales so that a quadratic energy is a poor approximation [160, 161]. Furthermore, at this length scale the double-stranded structure and the base-pair details become much more relevant, necessitating the use of rigid-base models [139]. These models can account for more complexity in the DNA structure by taking into account stereochemical relations between individual bases [139], but this requires a high computational cost. In contrast to describing the deformations of the filament by base (base-pair) frames as done in the rigid-base (basepair) models [138, 139, 141], our discrete model corresponds to a segmented twistable worm like chain in which the number of links and the link length are arbitrary. By segmenting the chain into two pieces we have provided analytical expressions for a two-state heterogeneous chain which are relatively easy to compute. Our two-state model is suitable for interpreting DNA single molecule experiments in which the molecules are a few hundred nanometers long.

Chapter 6

Equilibrium and kinetics of DNA overstretching modeled using a quartic energy landscape

Under physiological conditions inside the cell, the prevalent dsDNA conformation is the B-DNA form, a right-handed double helix with ~ 10.5 base-pairs (bp) per helical turn and approximately 0.34nm per base-pair. When B-DNA is subjected to external stress conditions it is known that it can undergo conformational changes into other DNA forms [162]. A key experiment used for the study of DNA response is one in which a tensile force is applied on the DNA, leading to an elastic regime of increasing extension and eventually to what has been denominated in the literature as an overstretching transition from B-DNA to an elongated form [43].

The force-extension curve of DNA has been extensively studied in the literature by both, experimental and theoretical methods, and we refer the reader to Bustamante et al. [43] for a more detailed review regarding the advances in the study of DNA under tension. Some of the pioneering experimental work regarding the overstretching transition was performed by Smith et al. [163], and Bloomfield and coworkers [7, 8, 164]. Parallel to the experimental discovery of the overstretching transition, Cluzel et al. [165] presented a theoretical model analogous to the helix-coil transition model. Since then, the majority of the theoretical efforts discussing the DNA overstretching transition are based on models such as the Zimm-Bragg theory or Ising models, where the theory is adapted to include entropic elasticity effects. Ahsan et al. [166] and Marko [134] included entropic elasticity by means of the worm-like-chain (WLC), where both states were assumed to have the same flexural rigidity of B-DNA. Rouzina and Bloomfield [7] combined the Zimm-Bragg model with both the WLC model and the Freely Joint Chain (FJC) to describe overstretching, while Storm and Nelson [156, 167] and Cizeau et al. [168] presented an Ising-type model where each state can have arbitrary elastic constants. Besides helix-coil type models, simulations of a dynamical Langevin model using a Landau-Ginzburg landscape [169] and simulations of a dynamical model using Peyrard-Bishop-Dauxois (PBD) mesoscopic model for the energy potential [170], have also been proposed to describe the DNA transition at high forces.

With advances in experimental techniques the interest in the overstretching transition has peaked in recent years, and it has been argued that B-DNA can undergo more than one overstretching transition. Yan and coworkers [30, 32, 158] and King et al. [33] have been using single molecule methods to study the overstretching transition, reaching the conclusion that DNA can undergo three types of transitions: B-DNA to peeled DNA (ssDNA), B-DNA to Melted DNA (M-DNA, inside strand separation in the form of bubbles) and B-DNA to S-DNA (non-hysteretic transition). The experiments can be done in such a fashion that peeling is not topologically allowed in what

the authors call end-closed set up [32]. Furthermore, it was established that even in the case of end-opened DNA (which does not preclude peeling topologically), unpeeling is suppressed with increasing ionic strength [33]. It has also been shown that when the content of (AT) tracts is high, a force-induced melting transition is prevalent, while sequences with a higher GC content undergo a non-hysteretic overstretch transition into the S-form [33, 36]. Moreover, the different overstretching transitions seem to be characterized by different levels of cooperativity and kinetic behavior [171]. Bianco et al. [35] and Bongini et al. [34] studied the transition kinetics of the overstretching transition using force-steps in pulling experiments of λ DNA (~ 48.5 kbp), where they found that the cooperative length of the B-to-S transition is approximately (22-25)bp.

Cooperativity during a phase transition is defined as the phenomenon in which some property changes gradually (in a sigmoidal way) as a function of the external controlled parameter [172]. This means that certain regions of the system are somehow tied together in such a way that the driving force on a region to undergo the phase transition is directly influenced by whether other regions have undergone the conformational changes [173]. A system in which n identical subunits undergo a phase transition in perfect unison is denominated as perfectly (maximum) cooperative. If the transition is less cooperative, the gradual change of the parameter during the transition would be less steep. Therefore, during a phase transition one can define a cooperative unit for the transition [173]. This unit of cooperativity reflects the size of the subunits that undergo the conformation changes completely independently of one another. In the case of DNA filament undergoing a two state phase transition, the cooperative unit size is the number of base-pairs that will transition as a single unit.

In the present work we focus on the regime where peeling is not allowed, and we present a theoretical model applicable to the B-to-S and B-to-M transitions, where we assume that at high forces the bending effects can be neglected. In our cooperative system the property that varies in a sigmoidal fashion is the extension of the DNA molecule as a function of the applied force, where the cooperative unit size is defined by the length l of the number of basepairs of independent segments n that make up the total DNA chain. The main objective of our work is to obtain the value of the cooperative unit l . To do so we developed an analytical continuous two-state model for the DNA overstretching transition, in which we account for the global statistical fluctuations of the system due to thermal effects. We compare our model to overstretching experiments in order to obtain the values of l . Our results are in excellent agreement with independent experimental measures of the cooperative unit l for dsDNA overstretching transitions [34, 35, 171]. Furthermore, in agreement with recent experimental evidence [32–34], our predictions of l support the notion that the overstretched state is a mixed DNA form at the temperature range and ionic concentrations used in various experiments. We find that l varies depending on whether the B-DNA undergoes a transition closer to a pure S form, or a transition closer to the M form. Finally, we connect our model with Kramers rate theory for stochastic systems with double-well potentials [174] to describe the kinetics of the system. We find that the resulting analytical model using our predicted l values accurately reproduces the kinetic behavior seen in experiments of Bianco et al. [35] and Bongini et al. [34].

6.1 Model Description

The DNA in the overstretching experiments is modeled as a continuous and extensible rod, where s is the arc length along the centerline of the rod. This rod (dsDNA filament) is subject to thermal fluctuations and high tensions depicted in the inset of Fig 6.1. Our problem is to evaluate the partition function and free energy of such a rod assuming that the energy can be expressed as a quartic function of the order parameter u . The variable $u(s)$ is the stretch of the rod that we map to the spatial coordinate (arc-length s). The order parameter $u(s)$ not only characterizes the rod configuration, but is also the reaction coordinate describing the onset of a phase transition driven

by the applied tension F .

6.1.1 Energy of the system

In the present model the DNA is assumed to be torsionally unconstrained such that there is no twist in the molecule. In the undeformed initial configuration the DNA is in the B state, in which its length is given by $L = N \cdot bp$, with N being the number of basepairs and $bp \approx 0.34\text{nm}$ the length of each base-pair. We constrained our model for a regime of moderate to high forces $F > \sim 15pN$. In this regime, the shortening of DNA due to thermal bending fluctuations is negligible. Next, we proceed with a mean field Hamiltonian for the energetics of the system:

$$H = [V(u) + Cu]L, \quad V(u) = A_4u^4 - A_2u^2. \quad (6.1)$$

Here $u(s)$ plays the role of the order parameter and it is a measure of the filament's stretch. A discrete version of the Hamiltonian given in Eq. (6.1) has been used in DNA overstretching dynamic simulations [169]. A_4, A_2 and C are phenomenological parameters to be evaluated by comparing to the overstretching experiments. The value of A_4 must be positive to satisfy the conditions of stability of $u(s)$.

6.1.2 Homogeneous Equilibrium Solutions

The homogeneous equilibrium stretch value u^* simply corresponds to the value of u that minimizes the potential $H(u)$, where C plays the role of the external field:

$$\left. \frac{dH}{du} \right|_{u=u^*} = g(u^*) = 4A_4(u^*)^3 - 2A_2(u^*) + C = 0. \quad (6.2)$$

For $C = 0$, expression Eq. (6.2) has one unstable solution $u_b = 0$ and two stable minima $\pm u_o$:

$$u^*|_{C=0} = \pm u_o = \pm \sqrt{\frac{A_2}{2A_4}}. \quad (6.3)$$

6.1.3 DNA overstretching transitions: changes in temperature and salt concentration

Within the Landau framework of phase transitions any of the parameters A_4 , A_2 or C , appearing in the phenomenological potential $V(u)$, can be a function of the controlled external variables [88]. In the DNA stretching experiments that we will consider, the controllable external variables are the force F , ionic solution concentration I and temperature T . Since $u(s)$ is the stretch, it is clear that the external field F must contribute through the linear term in the Hamiltonian, such that $C = -F +$ other terms. Furthermore, the effects of temperature T and salt concentration I will come through the linear term as well. This is analogous to what is seen in the Landau model of liquid-vapor systems [175]. Therefore, we assume a general form for $C(F, T, I)$:

$$C = f(T, I) - F. \quad (6.4)$$

The exact form of $f(T, I)$ is taken from the phenomenological relationships found in experiments [30, 32, 34, 34, 35]. In the experiments one of the variables T or I is constant while the other one can be varied. At the onset of the transition $C = 0$ and $F_c = f(i)$, where $i = \{I, T\}$, depending on whether the controlled variable is the temperature T or the salt concentration I . The value of the critical force F_c at which the transition takes place depends strongly on I and T [30, 32, 34]. Instead of using a general phenomenological equation for all experimental data, we have used the $f(i) = F_c$ as measured in each experiment. We do so because different groups reported different

phenomenological behavior for similar experimental conditions [30, 34]. A brief description of the empirical relations found by experimental groups between the critical force F_c , temperature T and concentration I is given in SI Appendix E.2.

6.1.4 Experimental observables: DNA extension z

There are two types of stretching experiments – extension controlled and force controlled experiments. We will focus mainly on recent experiments where F is controlled and the end-to-end extension z is measured [30, 32, 34]. The relationship between the order parameter $u(s)$ and the extension of the molecule z is given by:

$$z = (1 + u^* - u^*|_{F=\hat{F}}) \text{bp} \quad (6.5)$$

where $\text{bp} \approx 0.34\text{nm}$ is the base-pair length in the B-DNA state and $u^*(F)$ is the global minimum in Eq. (6.2). The order parameter u is a measure of stretch with reference to a midpoint between compact (B) and extended states (S or M) as shown in Fig. 6.1 inset. To set the reference state with zero stretch to be the B-form, where $z = \text{bp}[\text{nm}]$, u^* must be shifted by a constant. Therefore, the shifting constant $u^*|_{F=\hat{F}}$ (u^* evaluated at $F = \hat{F}$) ensures the condition of zero stretch in the system at the initial B-DNA configuration ($L = N \cdot \text{bp}$). The value of the force at which $z \approx 0.34\text{nm}$ is seen in DNA extension experiments is $\hat{F} \sim 15\text{pN}$. Note that $\hat{F} \gg 1\text{pN}$ since thermal fluctuations would effectively be shortening the DNA length in the small force regime. To distinguish between the B and overstretched forms, the length of the overstretched form after the transition is complete is labeled \bar{L} .

Next we describe the procedure to fit the values of the parameters A_2 and A_4 to the experimental overstretching curves $z(F)$. The first equation is provided by the difference in value of z (between the two states) at the midpoint of the transition ($C = 0$). The change in extension δz between the compact and extended form is approximately :

$$\delta z \approx (\bar{L} - L)/N = |2u_o| \text{bp}, \quad (6.6)$$

where u_o is the equilibrium solution of the order parameter given by Eq. (6.3) and it is a function of A_2 and A_4 . The second equation is provided by an analysis of the low force regime ($F \ll F_c$), where the filament is in the B-form. Far away from the transition point, the slope of $z(F)$ is not affected by the statistical fluctuations of u and it can be obtained from the parameters A_2 and A_4 using the zero-temperature model. Analogously to the stretch modulus in the WLC theory [163, 176], at high forces ($F > 15\text{pN}$) where bending effects are small, A_2 and A_4 quantify the change in extension $z(F)$. We found that $A_4 \approx 500\text{pN}$ consistently provided a good fit for the force-extension slope for all the experimental curves used in this study. Therefore, in the following sections we will use $A_4 \sim 500\text{pN}$, and let A_2 be the parameter that dictates the change in extension δz between the B-state and the overstretched state. To illustrate this procedure, in Fig 6.1 we fit the values of A_2 and A_4 to one of the Zhang et al. [32] experimental data sets.

So far we have described the zero-temperature model which neglects thermal fluctuations. In section 6.2 we include the contributions of the global statistical fluctuations of the order parameter $u(s)$ due to thermal effects and describe an analytical procedure to compute the average extension $\langle z \rangle$ as a function of the controlled parameters and the cooperative unit l .

6.2 Statistical Mechanics of the Chain

We model our DNA filament as a chain made of n segments of length l , where each segment is identical to others and each segment can be easily identified and labeled along the arc-length s . The Hamiltonian of a single segment j is given as:

$$H_j = [A_4 u_j^4 - A_2 u_j^2 + C(F, T, I) u_j] l. \quad (6.7)$$

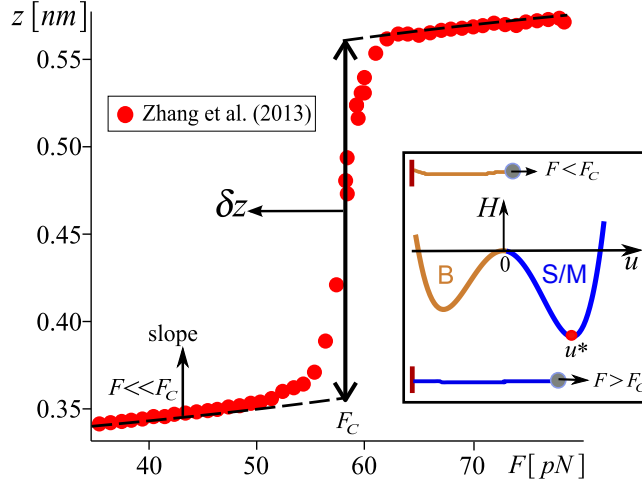


Figure 6.1: Procedure to fit values of A_2 and A_4 . Data points correspond to Zhang et al. [32] at $I = 3.5\text{mM}$ and $T = 12\text{C}$, where the critical force $F_c \approx 57.5\text{pN}$. Dashed lines correspond to extension z assuming no thermal fluctuation (see Eq. Eq. (6.5)). We fit $A_2 = 93\text{pN}$ and $A_4 = 500\text{pN}$ to the slope (far from the transition point) and to the change in extension δz between the compact and extended states. Inset: Schematic of double-well potential modeling the overstretching experiment. For $F < F_c$ system is in the B-DNA state. As the force increases, for $F > F_c$ the right well of the potential H is deeper as shown in the figure and the molecule transitions into the overstretched state. The right well represents the global minimum u^* in the overstretched form (S/M stands for S-DNA or M-DNA) and the left well corresponds to the local equilibrium of u in the B-form.

Then the total energy E of the n non-interacting distinguishable segments is $E = \sum_{j=1}^n H_j$.

We assume that each segment is an independent subsystem with its own specific set of boundary conditions that does not interact with other subsystems. We focus on a single subsystem where the phase change occurs by the passage of a single phase boundary (maximum cooperativity within each subsystem). Therefore, the length l is a measure of the cooperativity in our model.

If we further assume that the subsystems are identical, then the partition function of the chain of identical but distinguishable non-interacting segments is given by [177]:

$$Z_s = Z^n = \left[\sum_j \exp(-\beta H_j) \right]^n, \quad \beta = \frac{1}{k_B T}. \quad (6.8)$$

where k_B is the Boltzmann constant.

6.2.1 Partition function Z of a segment of length l

The order parameter in expression Eq. (6.7) is a continuous variable, therefore taking the sum over all possible configurations of $u_j(s)$ in the continuous limit yields the partition function of a single segment to be:

$$Z = \int_{-\infty}^{+\infty} \exp[-\beta l (A_4 u^4 - A_2 u^2 + C u)] du, \quad (6.9)$$

where, for convenience, we have dropped the j index. Next, to evaluate expression Eq. (6.9) we follow the methods in Tuszynski et al. [175] and define the variables:

$$a = A_4\beta l, \quad b = -A_2\beta l \quad \text{and} \quad c = C\beta l. \quad (6.10)$$

Then the partition function of the system can be rewritten by performing a Taylor expansion on the linear term of the Hamiltonian:

$$Z = \sum_{n=0}^{\infty} \frac{(-c)^n}{n!} \int_{-\infty}^{+\infty} u^n \exp[-au^4 - bu^2] du. \quad (6.11)$$

It is clear from Eq. (6.11) that when n is odd, $Z = 0$. When $n = 2m$ is even the solution to each one of the integrals in the summation of Eq. (6.11) can be found in [178], such that $Z|_{n=2m} = Z_m$:

$$Z_m = \sum_{m=0}^{\infty} \frac{(-c)^{2m}}{(2m)!} \frac{\Gamma(m + 1/2)}{(2a)^{(2m+1)/4}} \exp\left(\frac{q^2}{4}\right) U(m, q), \quad (6.12)$$

where $U(m, q) = D_{-m-1/2}(q)$ is the parabolic cylinder function [179]. We can simplify this expression further by making use of the properties of the $\Gamma(x)$ function:

$$\Gamma\left(m + \frac{1}{2}\right) = \frac{\sqrt{\pi}}{4^m} \frac{(2m)!}{m!}, \quad (6.13)$$

such that the partition function of an individual segment of length l is¹:

$$Z_m = \left[\frac{\pi^2}{2A_4\beta l} \right]^{1/4} \sum_{m=0}^{\infty} \left[\frac{Q^m}{m!} \exp\left(\frac{q^2}{4}\right) U(m, q) \right], \quad (6.14)$$

where

$$Q = (\beta l)^{3/2} \frac{[C(F, T)]^2}{4\sqrt{2A_4}} \quad \text{and} \quad q = A_2 \sqrt{\frac{\beta l}{2A_4}}. \quad (6.15)$$

Finally, the ensemble free energy of each subsystem is:

$$G = -\beta^{-1} \ln Z_m. \quad (6.16)$$

6.2.2 Force-extension relation for the chain undergoing the overstretching transition

Once the free energy G is known, one can compute the average value of conjugate variable $\langle u \rangle$ to the external field C :

$$\langle u \rangle = -\frac{1}{l} \frac{\partial G}{\partial C}. \quad (6.17)$$

The procedure to obtain $\langle u \rangle$ follows from using expression Eq. (6.14) and Eq. (6.16):

$$\beta G = \frac{1}{2} \ln \pi - \frac{1}{4} \ln(2a) + \ln \Psi, \quad (6.18)$$

where

¹Expression Eq. (6.14) is an exact result, but near the transition midpoint ($C \sim 0$) evaluating the sum up to $m = 2$ suffices to capture the transition behavior.

$$\Psi = \sum_{m=0}^{\infty} \left[\frac{Q^m}{m!} \exp\left(\frac{q^2}{4}\right) U(m, q) \right]. \quad (6.19)$$

Next, by means of the chain rule:

$$\begin{aligned} \beta \frac{\partial G}{\partial C} &= \Psi^{-1} \frac{\partial \Psi}{\partial c} \frac{\partial c}{\partial C} \\ &= \frac{cl\beta}{\Psi \sqrt{8A_4}} \sum_{m=1}^{\infty} \left[\frac{Q^{m-1}}{(m-1)!} \exp\left(\frac{q^2}{4}\right) U(m, q) \right]. \end{aligned} \quad (6.20)$$

We can simplify this expressions further by making use of the definition $\Gamma(x) = (x-1)!$ and the relations in Eq. (6.10), such that the average value of the order parameter is :

$$\langle u \rangle = \left[\frac{(l\beta)^{1/2} C}{\Psi \sqrt{8A_4}} \right] \sum_{m=1}^{\infty} \left[\frac{Q^{m-1}}{\Gamma(m)} \exp\left(\frac{q^2}{4}\right) U(m, q) \right], \quad (6.21)$$

where Ψ is given by Eq. (6.19) and Q, q are defined in Eq. (6.15). Finally, replacing u^* with the expected value $\langle u \rangle$ in equation Eq. (6.5) we can compute the average end to end extension:

$$\langle z \rangle = (1 + \langle u \rangle - u^*|_{F=\hat{F}}) bp. \quad (6.22)$$

In section 6.3 we compare our model to the experiments with C as defined by the phenomenological expression Eq. (6.4).

6.3 Applications: comparison with experimental data

Our strategy to obtain the cooperative unit l from the experimental data is the following. First, we fit A_2 and A_4 using a zero-temperature model as described in section 6.1.4. Next, for convenience, we shift the extension experimental curves by the value of the critical force F_c corresponding to each experiment. Then the extension $\langle z \rangle$ can be expressed as a function of $\Delta F = F - F_c$. Finally, combining expressions Eq. (6.21) and Eq. (6.22), we fit l (the only unknown in the system) to the experimental curves. In Fig. 6.2(a) we show the fitting of our statistical model to the same experimental curve presented earlier in Fig. 6.1. The force-extension prediction with $l \approx 25$ bp agrees very well with the data set in [32]. Through different methods and at different ionic conditions $I = 150$ mM, Bianco. et al. [35] and Bongini et al. [34] measured the cooperativity length to be $l \in [22, 25]$ bp. They found that this value of l is essentially independent of temperature for $T \in [10, 25]C$. Another group [171], reported that the B-to-S transition is characterized by a slightly lower value of $l \approx 10$ bp and that the B-to-M transition is much more cooperative where $l \approx 100$ bp. Since $l = 25$ bp is between the two reported values for B-to-S and B-to-M in [171] we consider the possibility that the overstretched state, in the experimental data in Fig. 6.2(a) and the experiments in [34, 35], is a mixture of S and M. In fact, recent experimental findings strongly support the idea of a mixed overstretched form. In the next section we briefly review some of the experimental results behind this idea.

6.3.1 DNA overstretching transitions: B-to-S, B-to-M or both

While there is conclusive evidence that S-DNA exists and that it is favored at high salt concentration and low temperatures, and that M-DNA can also be present in force induced transitions for lower salt concentrations and high temperatures [30, 32, 33, 36], there are still some unanswered questions

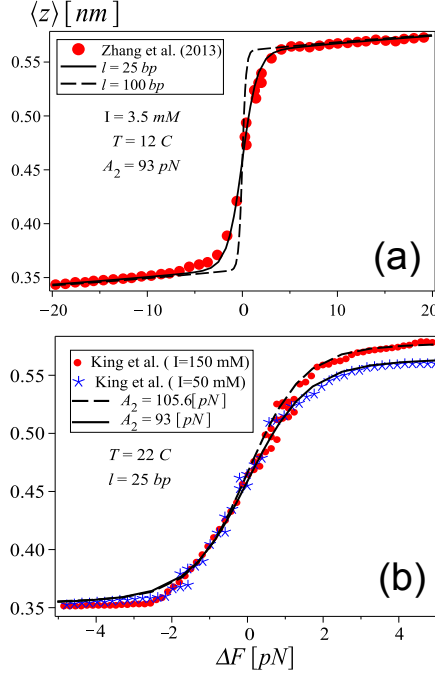


Figure 6.2: Force-extension relation during dsDNA overstretching. (a) Solid ($l = 25\text{bp}$) and dashed ($l = 100\text{bp}$) theoretical predictions for Zhang et al. [32] experiments (same conditions as in Fig. 6.1). (b) Data points correspond to two experiments in King et al. [33], while lines are theoretical predictions from our model using $l = 25\text{bp}$. Here $F_c = 69.5\text{pN}$ at 50 mM and $F_c = 63.5\text{pN}$ at 150 mM .

regarding the structure of the overstretched state. In table 6.1 we list some of the experiments that present the possibility of a mixed S and M state after the overstretching transition. For instance, M-DNA was present in a non-hysteretic process [33], whereas in Zhang et al. [32] hysteresis was always seen in a significantly shorter M-DNA. As pointed out by the authors in [32] the difference could be explained if there was a mixture of S and M in the overstretching experiment in [33]. Furthermore, in [34, 35] the overstretched length was ~ 1.7 times the B-DNA form for both hysteretic and non-hysteretic transitions and the authors kinetic data suggests two processes during the overstretching transition. The idea of M and S coexistence is further supported by the gradual change in extension of the overstretching curves in [32] as a function of ionic concentration. Based on these experimental facts we think our model will be useful in using the cooperative length l to quantify the mixing of S and M DNA after the transition.

6.3.2 Cooperativity length l predictions

Next, we present some ideas that emerged when we used our model to analyze the DNA overstretching curves of various groups:

(i) Regardless of temperature T , length L or ionic concentration I , for experimental curves where the change in extension is $\bar{L} \approx 1.7L$ the cooperativity of the system is given approximately by $l \in [22 - 25]\text{bp}$. Although T, I and L are different among the data sets presented in Fig. 6.2, the three sets are quantitatively reproduced by our model using $l \approx 25\text{bp}$. Similarly, Fig. 6.3(a) shows that using $l \approx 22\text{bp}$ accurately reproduces Zhang et al. [32] experiments at $I = 150\text{mM}$ and $T = 22\text{C}$. Furthermore, using $l \approx 22\text{bp}$ in Fig.E.1, we show that Bianco et al. [35] and Bongini et al. [34] experiments at $I = 150\text{mM}$ and $T = 25\text{C}$ are consistent with experiments in [32, 33].

Experiment	T [°C]	I [mM]	L [kbp]	Description	\bar{L}
Zhang et al. [32] Fig. 3b	24	150	~ 7.3	n-h (S-DNA)	~ 1.7
Zhang et al. [32] Fig. 3b	24	1.0	~ 7.3	h (M-DNA)	~ 1.5
Zhang et al. [32] Fig. 3c	12	3.5	~ 7.3	n-h (S-DNA)	~ 1.7
King et al. [33] Fig 1A	22	50	~ 48.5	n-h (M-DNA)	~ 1.7
King et al. [33] Fig S4	22	150	~ 25	n-h (S-DNA)	~ 1.7
Bongini et al. [34]	10-25	150	~ 48.5	n-h and h	~ 1.7
Zhang et al. [30]	10-20	500	~ 48.5	n-h (S-DNA)	~ 1.7

Table 6.1: Summary of recent DNA overstretching experiments used throughout this study. L is the length in the B-DNA state and \bar{L} is the length of the DNA filament after the transition measured in L units. The abbreviations n-h and h stand for non-hysteretic and hysteretic respectively.

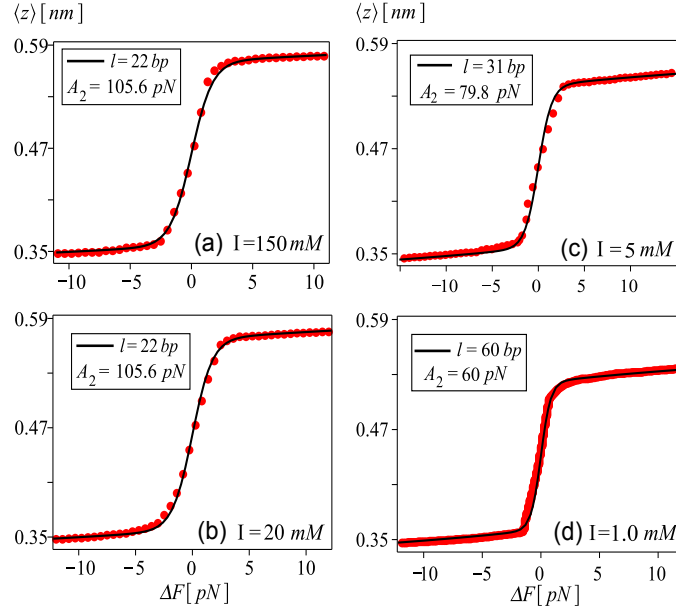


Figure 6.3: DNA overstretching at $T = 24\text{C}$ for different ionic concentrations. Red Markers correspond to Zhang et al. [32] experiments and solid lines are theoretical predictions from our model. As the ionic strength decreases going from (a) trough (d), A_2 decreases and l increases. We use $A_4 \sim 500\text{pN}$ and F_c as measured in experiments: $F_c = [68.3, 63.5, 58.8, 50.5]\text{pN}$ going from panel (a) through (d).

(ii) The experimental data in [32] showed a gradual change in $\langle z \rangle$ as a function of ionic concentration. This is shown in Fig. 6.3 where we fit data from figure 3b in [32]. In Zhang et al. [32], at the low ionic concentrations, there is an asymmetric pattern in the force-extension curves when the force increases or decreases. This *hysteretic* behavior observed during overstretching is due to the slow convergence to equilibrium [159, 169]. The asymmetric hysteresis at $I = 1\text{mM}$ and $I = 5\text{mM}$ in [32], where the system is out of equilibrium mainly during unloading, is consistent with previous stretching experiments that depict marked hysteresis during the decreasing force regime [34, 35, 164, 170, 180, 181]. During the loading phase in some of these experiments there are no hysteretic effects, while in others, if hysteresis is present, its effect is significantly less pronounced than during unloading. Therefore, for the low $I = 1$ and 5mM we fit only to the pulling data in [32]. As shown in Fig. 6.3, l decreases gradually with increasing I , ranging from $l \approx 60\text{bp}$ at $I = 1\text{mM}$ down to $l \approx 22[\text{bp}]$ at $I = 150\text{mM}$. The smaller cooperativity values are consistent with $\bar{L} \rightarrow 1.7L$. Given that experiments [32, 33, 171] confirm that the S-form is preferred at high salt concentrations

and that the S form is mainly responsible for the 1.7 times change in extension [34], we conclude based on our fittings that for $l < 30\text{bp}$ the predominant state in the overstretched form is S-DNA.

(iii) As the transition becomes less cooperative, the increase in l of each subsystem has the same theoretical effect as drastically decreasing the temperature T , as evidenced from the definition of the partition function given in Eq. (6.9). A change of 4 times in the value of l from 25bp to 100bp is equivalent to a decrease of the absolute temperature from room temperature to $T = 75\text{K}$, which effectively yields less global statistical fluctuations. Thus, one can expect the sigmoidal curve to sharpen and become closer to an abrupt first order phase transition. In figure 6.2(a) we have plotted the curve for $l = 100\text{bp}$ next to the 25bp solution, so that the difference in width of the curves is apparent.

(iv) Although moderate changes in temperature $T \in [10 - 25]\text{C}$ affect the critical force F_c value [32, 34], the change of temperature in this range does not seem to affect the extension of the molecule up to $T = T_M$ at fixed I (see Fig 3(c) in [32]). But once $T \geq T_M$, there is a sudden change in the extension profile of the overstretching curves [32]. Bongini et al. [34] and Zhang et al. [30] data support the idea that at a fixed I for a range of temperatures $T \in \sim [10 - 25]$, the extension of the overstretched form remains approximately the same. This would imply that given a fixed I , there is a single transition class to the S-form (or at least closer to pure S) for $T < T_M$ and a melting transition for $T > T_M$. Hence, we think of l as independent of T for each transition class, and making use of the phenomenological model for $C(F, T)$ as described in section 6.1.3, we predict the behavior of the overstretching transition as a function of T . A sample of the results is shown in Fig. E.3, where we show that introducing the temperature effects through $C(F, T)$ captures what is seen in experiment.

(v) At higher ionic concentrations ($I = 500\text{mM}$ [30]), we found that although $l = 22\text{bp}$ is a good average fit to the experimental data, the curve is not symmetric about the midpoint of the transition and the data is better fit by $l \approx 15\text{bp}$ near the overstretched state (See fig 6.4). Similar behavior is found in King et al. [33] overstretching curve at $T = 22\text{C}$ and $I = 1\text{M}$. But, this asymmetric aspect of the overstretching transition is much more evident in Fig. 6.5 where we present the comparison of Zhang et al. [30] variance measurements with our theoretical predictions. Since the n segments making up the entire chain are independent of one another, the system is analogous to a random walk of n steps. In this analogy the average step size of the walker is $\langle ul \rangle$ and the variance of each step equals the variance of one segment of length l :

$$\sigma_l^2 = l^2 \left(\langle u^2 \rangle - \langle u \rangle^2 \right), \quad (6.23)$$

where $\langle u^2 \rangle = d(\ln Z_m)/db$ is the second moment of the partition function and b is defined in expression Eq. (6.10). Then the variance for the entire chain (n -steps) is [146] :

$$\sigma_v^2 = n\sigma_l^2 = L \left(\langle u^2 \rangle - \langle u \rangle^2 \right) l. \quad (6.24)$$

For a given force, since L is fixed, the variance grows linearly with the cooperativity length. In Fig. 6.5 gray circles correspond to Zhang et al. [30] experimental measurements of the variance at $I = 500\text{mM}$. Lines correspond to our theoretical predictions for different values of l using Eq. (6.24). The red solid line (for $l = 15\text{bp}$) agrees strongly with the experimental data for $\Delta F > 0$ (right side of the graph), while on the left side of the graph the blue solid line (for $l = 30\text{bp}$) provides a better fit. Black solid line ($l = 22\text{bp}$) in Fig. 6.5 is shown as an average fit for both sides of the graph.

(vi) An alternative method to quantify the cooperativity of the DNA overstretching transition is to use the Zimm-Bragg parameter σ_F [7]:

$$\sigma_F = \exp(-2\beta E_s), \quad (6.25)$$

where $2E_s$ is the energetic cost involving two junctions (this is the definition given in [182]). In

l [bp]	Fig.	$\sigma_F \cdot 10^{-3}$	δF [pN]	δz [nm]
55	Fig. 6.3(d)	0.3	1.8	0.16
31	Fig. 6.3(c)	1.0	2.8	0.19
25	Fig. 6.2(a)	1.6	3.1	0.20
22	Fig. 6.3(a)	2.0	3.4	0.21
15	Fig. 6.4(b)	4.3	4.5	0.23

Table 6.2: Calculation of the Zimm-Bragg cooperativity σ_F as a function of l using relation Eq. (6.26). As a reference value, Rouzina and co-workers [7, 8] measured $\sigma \approx 10^{-3}$ in DNA overstretching experiments at room temperature and $I = 150\text{mM}$.

qualitative terms, large values of a cooperative unit l are analogous to small values of σ_F [173], but a quantitative relation can be obtained by the following procedure. The parameter σ_F reflects the width δF of the overstretching transition in terms of the force [7, 182]. δF can be determined by the midpoint slope of the plot $P_S vs. F$ [7, 182], where P_s is the fraction of the filament in the overstretched state. Then the force transition width is [7]:

$$\delta F = \left. \frac{\partial f}{\partial P_s} \right|_{F=F_c} = 4\sigma_F^{1/2} \frac{k_B T}{\delta z}, \quad (6.26)$$

where δz is the change in extension per basepair during the transition. Instead of using δz as approximated before in the zero-temperature calculations, the model presented in section 6.2 with statistical fluctuations allows to provide a more exact estimate for δz :

$$\delta z = \langle u_f \rangle - \langle u_i \rangle, \quad (6.27)$$

where the subscripts i and f stand for initial and final point of the transition respectively. Using the definition of P_s given by Eq. (6.39) we can directly compute δF by the left equality in expression Eq. (6.26). Next, making use of expression Eq. (6.21) we can evaluate $\langle u_i \rangle$ and $\langle u_f \rangle$ at $F_i = F_c - \delta F/2$ and $F_f = F_c + \delta F/2$ respectively. By doing so, expression Eq. (6.26) directly links our methods to the Zimm-Bragg cooperativity model, and we can calculate the parameter σ_F as a function of l . In table 6.2 we present δF and σ_F for several sets of experimental data used throughout this chapter, where we see that σ_F is of the order of 10^{-3} in agreement with the reported values in [7, 8].

Since the model based on subsystems of cooperative length l accurately describes the quasi-static overstretching experiments, in section 6.4 we extend our methods to study the kinetics of a system with sharp interfaces, meaning the phase transition takes place in a spatially homogeneous way.

6.4 Kinetics of the chain: Sharp interface

Next, we consider the kinetics of a single chain unit of length l , where the order parameter u is the relevant macroscopic variable describing the dynamic process over time t . Due to the effects of thermal fluctuations in the fast changing microscopic variables, the evolution of $u(t)$ is stochastic in nature and it obeys the Langevin equation [174]. Bongini et al. [34] found that the kinetic mechanism during the transition involves viscosity dependent delocalized motions at low frequency. Hence, we consider the case of spatial-diffusion-limited rate theory [174], where the kinetic equation of the over-damped system can be simplified to:

$$\partial_t u(t) = \gamma [-\partial_u H_j(u)] + \sqrt{2\gamma k_B T} \bar{\xi}(t), \quad (6.28)$$

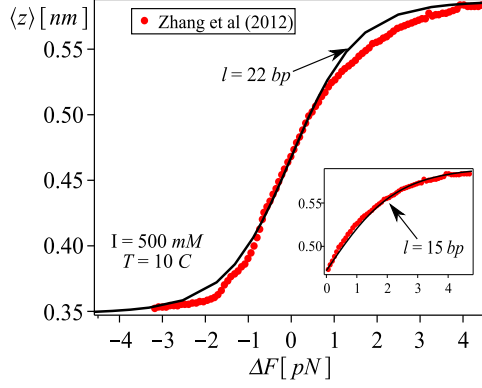


Figure 6.4: Force-extension curve fit to Zhang et al. [30] experiments at higher salt concentration. We have used $A_2 = 120\text{pN}$ and $F_c = 67.2\text{pN}$. Although $l \approx 22\text{bp}$ provides a good overall fit, using $l \approx 15\text{bp}$ provides better agreement in the upper-right section of the curve as shown in the inset.

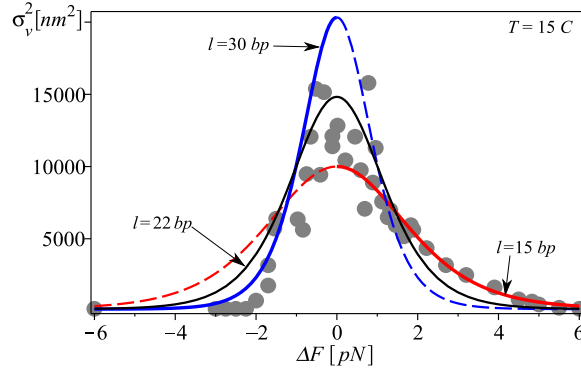


Figure 6.5: Variance σ_v^2 as a function of the force. Points correspond to Zhang et al. [30] experimental measurements at $I = 500\text{mM}$. Lines correspond to our theoretical predictions for different values of l using Eq. (6.24). The red solid line $l = 15\text{bp}$ agrees strongly with the experimental data for $\Delta F > 0$ (right side of the graph), while on the left side of the graph the blue solid line $l = 30\text{bp}$ provides a better fit. The asymmetric fitting to the variance is in agreement with the results shown for the force extension curves at the same high ionic conditions in Fig. 6.4. Black solid line $l = 22\text{bp}$ is shown as an average fit for both sides of the graph.

where H_j is the potential of a subsystem of length l given by Eq. (6.7), γ is the kinetic coefficient which we assume to be constant independent of the external parameters, and the Gaussian noise term $\bar{\xi}(t)$ is defined in terms of the Dirac delta function $\delta(x)$:

$$\langle \bar{\xi}(t_1) \bar{\xi}(t_2) \rangle = \delta(t_1 - t_2). \quad (6.29)$$

Eq. (6.28) describes the classical problem of the diffusion of a particle (unit) $j \in n$, with probability density function $\rho(u, t)$, which is moving in a potential field H_j . The inset of Fig. 6.6 shows the potential for $C = 0$, which is characterized by the two stable minima $\pm u_o$ and an energy barrier E_l with maximum value at the unstable solution $u_b = 0$. For the over-damped case, the time evolution of $\rho(u, t)$ is governed by the Smoluchowski equation [174] and following Kramers methodology for $E_l \gg k_B T$ [183], we can determine the steady-state escape rates. The details of the procedure to obtain the rate from B-to-overstretched (k_L) and overstretched-to-B (k_R) using the potential $H_j(u)$ are shown in SI Appendix E.3. The final results are:

$$k_L = \sqrt{\frac{k_B T}{2\pi}} \left(\frac{w_b \gamma}{Z_1 + Z_2} \right), \quad k_R = \sqrt{\frac{k_B T}{2\pi}} \left(\frac{w_b \gamma}{Z_1 - Z_2} \right), \quad (6.30)$$

where w_b is the curvature that results from linearizing the potential $H_j(u)$ about the unstable solution u_b [174] and

$$Z_1 = \sum_{m=0}^{\infty} \frac{(\beta C l)^{2m}}{(2m)!} \frac{\Phi(m)}{2}, \quad (6.31)$$

$$Z_2 = \sum_{\hat{m}=0}^{\infty} \frac{(\beta C l)^{2\hat{m}+1}}{(2\hat{m}+1)!} \frac{\Phi(\hat{m} + \frac{1}{2})}{2}, \quad (6.32)$$

$$\Phi(x) = \frac{\Gamma(x + 1/2)}{(2\beta A_4 l)^{(2x+1)/4}} \exp\left(\frac{q^2}{4}\right) U(x, q). \quad (6.33)$$

The parameter $q^2 = 2\beta E_l$ is given by Eq. (6.15). Therefore, the rates can be cast in the familiar Arrhenius type form:

$$k_i = \Upsilon(F, T) \exp(-\beta E_l) \quad (6.34)$$

where $i = [L, K]$, $E_l \gg k_B T$ is the energy barrier evaluated at $C = 0$ (Eq. Eq. (E.3)), and Υ is a function of the external parameters F and T .

Next, we let n_B be the number of j segments in B-state and n_S the number of segments in the overstretched state. Then at any instant in time, the total number of segments $n = n_B + n_S$ is conserved. Therefore, the change of n_B as a function of time is given by:

$$\frac{dn_B}{dt} = -k_L n_B + k_R n_S = -r n_B + k_R n, \quad (6.35)$$

where $r = k_L + k_R$ is the relaxation rate [174]. The solution of the first order ODE Eq. (6.35) is:

$$\frac{n_B}{n} = \left(\frac{k_R + \alpha \exp[-rt]}{r} \right), \quad (6.36)$$

$$\frac{n_S}{n} = \left(\frac{k_L - \alpha \exp[-rt]}{r} \right), \quad (6.37)$$

where α is the integration constant and the equilibrium steady state values are:

$$\frac{\hat{n}_B}{n} = \frac{k_R}{r}, \quad \frac{\hat{n}_S}{n} = \frac{k_L}{r}. \quad (6.38)$$

Finally, the equilibrium probability of segments in the overstretched state can be expressed as:

$$P_s = \frac{\hat{n}_S}{n} = \frac{1}{2} - \frac{Z_2}{2Z_1}. \quad (6.39)$$

where Z_2 and Z_1 are given by equations Eq. (6.31) and Eq. (6.32).

In Fig. 6.6 we compare Bongini et al. [34] experimental measurements of the relaxation rate r with the predictions of our model. Using the same values of A_2, A_4 and l that resulted from the force-extension curve analysis at the same experimental conditions (see Fig. E.1(a)), we are able to accurately reproduce their experimental findings. The only fitting parameter is the kinetic coefficient which we find to be $\gamma = 200$. Using $l = 22\text{bp}$ and assuming γ constant, our predictions for the relaxation rate r_{min} (corresponds to $F = F_c$) as a function of T , yield a range of values $r_{min} = 5.5s^{-1}$ at $T = 25C$ to $r_{min} = 3.2s^{-1}$ at $T = 10C$. These are in agreement with the

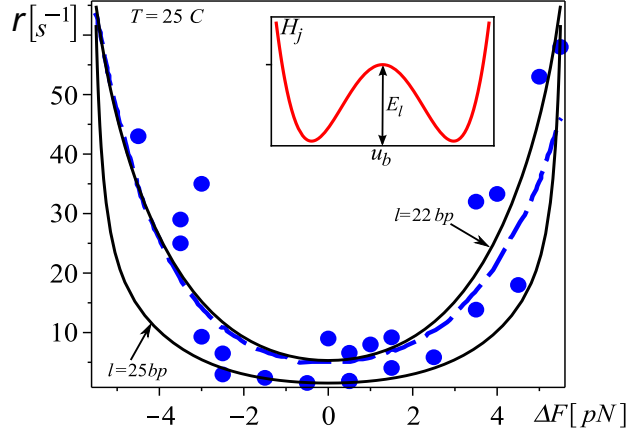


Figure 6.6: Relaxation Rate $r = k_L + k_R$ in dsDNA overstretching experiments at $I = 150\text{mM}$. Blue markers correspond to 2pN force-step experiments in [34], where $F_c \approx 66\text{pN}$. Blue dashed line is the fit used in [34], where the authors combined their experimental measurements with Kramer-Bell theory. Black solid lines correspond to theoretical predictions from expressions in Eq. (6.30) using $l = 22$ and 25bp . Here we have used $\gamma \sim 200$.

values reported in [34], where $r_{min} \in [3.5, 7.1]\text{s}^{-1}$ is approximately constant for the same range of temperatures. Using the same parameters as in Fig. 6.6, in Fig E.4 we show a fit to another set of data in Bongini et al. [34], where the authors have used a different loading protocol.

One of the protocols used by Bongini et al. [34], denominated by the authors *square wave protocol*, consisted in instantaneously applying and reversing a large force step S_F . During the loading phase, F is instantaneously changed from $F^{(1)} \ll F_c$ (entire molecule is in the B form), up to a value of $F^{(2)} \approx F_c$ (midpoint of the transition). After the system reaches equilibrium, at time t_2 it is unloaded using an instantaneous change in force of the same magnitude from F_c to $F^{(3)}$ [34]. In conditions at which the M form is energetically unfavorable in comparison to the S form ($T = 10\text{C}$ and $I = 150\text{mM}$) [32, 33], Bongini et al. [34] studied the transient kinetics using the *square wave protocol* going from $F^{(3)} = F^{(1)} = 47\text{pN}$ to $F^{(2)}$ slightly larger than F_c . Their findings for the lengthening and shortening responses in the absence of melting coincide with the behavior described by our theoretical predictions. In Fig. 6.7 we present the theoretical solution for a dsDNA molecule undergoing a B-to-S transition. We use $l = 22\text{bp}$ and $\bar{L} = 1.7L$, values corresponding to an overstretching transition close a pure S form (see section 6.3). To model the large force-step lengthening experiment, we set the initial condition at time t_1 to be $n_s(t_1) = 0$ (dsDNA in B-form initially), which yields $\alpha = k_L$ in equation Eq. (6.37). The red line in Fig. 6.7 shows the theoretical calculation of the evolution of the fraction of segments n_s/n in the loading phase for $F^{(1)} \ll F_c$. For $F^{(2)} = F_c$ the system reaches steady state in $\sim 1\text{s}$ in agreement with Bongini et al. [34] experiments and the time scale measurements of the B-to-S transition reported in [171]. During the unloading phase (shortening), the initial condition is $n_s(t_2) = n/2$, such that $\alpha = (k_L - k_r)/2$ in equation Eq. (6.37). After the force is the suddenly dropped to $F^{(3)} = F^{(1)}$ there is no exponential time course during the unloading, rather an abrupt instantaneous change in the extension of the molecule [34]. The blue solid line in Fig. 6.7 shows that our predictions for n_s/n during unloading and $F^{(3)} = F^{(1)}$, match the experimental evidence, by depicting a sudden (almost instantaneous) change from $n_2 = 0.5$ to 0. This behavior can easily be understood in the context of our theory in the following way. For any applied force step ($S_F = F^{(3)} - F_c$) < 0 , the left well in H_j will become deeper than the right well, and the population of segments in the right well will diffuse into the left well ($k_R > k_L$). For sufficiently small $F^{(3)} \ll F_c$ the expression n_s/n in Eq. (6.37) can be

approximated to be:

$$\frac{n_s}{n} \approx \frac{\exp(-k_R t)}{2} \quad \text{as} \quad \frac{k_L}{k_R} \rightarrow 0. \quad (6.40)$$

and for $t > 0$ the fraction $n/n_s \rightarrow 0$ rapidly. The dashed blue lines in Fig 6.7 show the time evolution of n_s/n during the unloading phase for different force steps ($S_F = F^{(3)} - F_c$).

Bongini et al. [34] also conducted the same *square wave protocol* experiment at $T = 25\text{C}$, where it is possible the M and S forms start to mix in the overstretched state, so the authors denominated this experiment the kinetics in a *partially melted molecule*. The shortening transient was characterized by two different processes: (a) stepwise shortening corresponding to the S-to-B transition [34] and (b) a much slower ($\sim 10\text{s}$) process likely to correspond to a rate-limiting re-annealing of the melted segments [34]. On the other hand, during the lengthening transient of the partially melted DNA at $T = 25\text{C}$, the overall exponential behavior characteristic of the B-to-S transition remains [34]. Moreover, as mentioned earlier, the hysteretic behavior during the overstretching transitions is mainly seen during unloading, while in the pulling phase the system is likely to be closer to equilibrium at all times. The difference in time that it takes the molecule to melt in comparison to the time taken by the molecule's recombination during unloading can explain the asymmetric hysteresis [170]. We already know that the kinetic two-state model presented in this study, accurately captures the time course of the non-hysteretic B-to-S transition at $T = 10\text{C}$, therefore, in the analysis of the lengthening transient at $T = 25\text{C}$ in Bongini et al., a two-state model with a *lumped* overstretched state where the S form is predominant is somehow justified. The inset of Fig. 6.7 shows the length change from B-to-overstretched ($\Delta L/\Delta L_e \approx n_s/\hat{n}_s$) as a function of time, where ΔL_e is the length measurement at steady state. Solid line corresponds to the theoretical solution at $T = 25\text{C}$ and $F^{(2)} = F_c$ with the same parameters as in Fig. 6.6 ($l \approx 25\text{bp}$), and red markers correspond to data from figure 1C in [34]. While for $F_c = F^{(2)}$ the system reaches steady state in $\sim 1.5\text{s}$, for $|F_c - F^{(2)}| > 0$, our model predicts the system will reach steady state faster. This last feature is also present in Bianco et al. [35] and Bongini et al. [34].

As a final remark, we point out that assuming $\gamma \approx 200$ remains independent of ionic concentration, the relaxation rate (r_{min}) for the four fits in Fig. 6.3 are of similar magnitude. Therefore, our methods predict that during a force-step pulling experiment, the time it takes the system to reach steady state at F_c ($\tau \approx 5r_{min}^{-1}$), is $\sim < 1\text{s}$ for all 4 experimental conditions in Fig. 6.3. Given that Zhang et al. [32] uses a pulling protocol of 1pN/s , this means according to our loading curve fits, that the molecule has enough time to reach equilibrium. We also calculated the effect of l on τ (see Fig E.7), which shows that τ increases very fast with l . In fact, when using the parameters from Fig 6.3(d), if l was to increase from 60bp to 80bp , τ would increase by factor of 10, and at a 1pN/s pulling rate, hysteresis would be seen. We conclude that for conditions when the M-form starts to become energetically more favorable, extremely low salt concentrations and/or very high temperatures $T > 30\text{C}$, it is possible that $l \rightarrow 100\text{bp}$ in dsDNA overstretching experiments. But, at the same time, we are likely to see hysteresis effects during both loading and unloading. We think this could be the case in the experimental curves in Zhang et al. [32] at $I = 3.5\text{mM}$ and $T \geq 34\text{C}$ where there is some hysteretic behavior even during the pulling phase. The presence of hysteresis during lengthening and shortening precludes the use of our methods to extract the l value from the overstretching curves. However, hysteretic effects can be greatly reduced by increasing the waiting times during pulling. In fact, times of $\approx 1\text{min}$ have been reported to be sufficient to completely reach equilibrium [159].

The results presented in this section provide insights on the kinetics of the system as a function of the imposed force F . However, the study of the kinetics of a phase transition can not be complete without an analysis of the nucleation and propagation of interfaces. In SI Appendix E.5 we present a summary of the kinetic analysis including domain walls near the transition point $F_c = F$. The results of SI Appendix E.5 on the kinetic analysis including propagating interfaces further validates the approximation of sharp interface presented in this section. Both solution methods share the same qualitative characteristics.

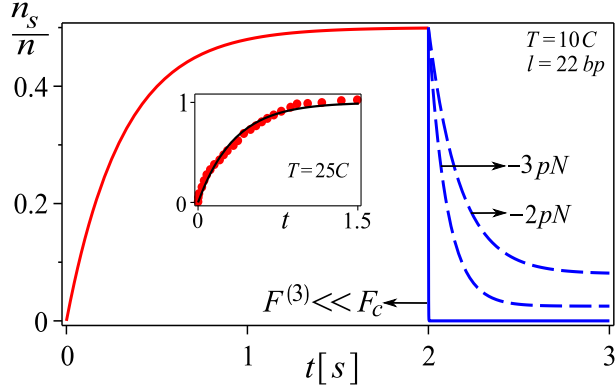


Figure 6.7: Transient Kinetics: B-to-S and S-to-B. Force-jumps loading and unloading a dsDNA molecule at $T = 10C$. As before we use $\gamma = 200$, $l = 22\text{bp}$ and $\bar{L} = 1.67L$. Red solid line depicts the exponential evolution of n_s/n in time during loading phase. At $t_1 = 0$, F is instantaneously changed from $F^{(1)} \ll F_c$ to $F^2 = F_c$. The system reaches steady state in $\sim 1\text{s}$. Blue lines depict evolution during the unloading phase. At $t_2 = 2\text{s}$, F is instantaneously dropped applying a force step $S_F = F^{(3)} - F_c$. For $S_F = -2$ and -3pN (blue dashed lines), there is exponential decay behavior, while for $F^{(3)} \ll F_c$ (blue solid) $n_s \rightarrow 0$ rapidly. Inset: Shows lengthening ($\Delta L/\Delta L_e$) as a function of time for a *partially melted molecule* (S+M). Red markers are Bongini et al. [34] data, and black line is our prediction using the same parameters as in Fig.6.6 ($l \approx 25\text{bp}$).

6.5 Conclusions

In the current work, we have tackled two problems in torsionally unconstrained DNA undergoing an overstretching transition – force-extension relation and kinetics. First, we introduced a model for the statistical mechanics of a chain composed of n segments of length l (the measure of cooperativity of the system), in which we neglect the spatial fluctuations of $u(s)$ within each segment of the chain. We assume that each segment is perfectly homogeneous at any instant in time and we focus on studying the global statistical fluctuations of $u(s)$. By doing so we are able to obtain an analytical expression for the end-to-end extension $\langle z \rangle$ of the filament as a function of force F , temperature T and the segment’s length l . In agreement with the Bongini et al. [34] conclusion from their experiments, we found by using our model to fit several sets of data from numerous groups [30, 32–35], that independently of length L , temperature T or ionic concentration I , if the extension of the overstretched state is $\sim 1.7L$, the cooperative length is $l \approx 25\text{bp}$. As the extension of the overstretched state decreases to $\sim 1.5L$, for conditions at which the M-form can compete with the S-form, we found $l > 30\text{bp}$ is required to fit the lengthening curves in [32]. There is, in fact, experimental evidence supporting the idea that for a range of temperatures and low to moderate salt concentration there would be a mixed overstretched state (S+M) [32–34]. The larger values $l > 50\text{bp}$ approach the measurements corresponding to the B-to-M transition ($l \sim 100\text{bp}$) reported by Rouzina et al.[171]. This feature would further support the idea that the $\sim 1.5L$ overstretched DNA is in fact M+S. Therefore, our model can be practical in using extension $\langle z \rangle$ and l as parameters to quantify the fractions in the mixed overstretched state.

In section 6.4 we extended our theory to encompass the kinetics of the chain under the assumption of an over-damped system based on experimental evidence [34]. Our kinetic results further consolidate our two-state cooperative model. We show that the theoretical predictions for the relaxation rates r and the system’s exponential time course are consistent with the results seen in experiments and simulations [34, 35]. Our calculations show that the time required to reach steady state in a transition to an overstretched state where the S-form is predominant is $\tau \sim 1\text{s}$, which agrees with the time scale reported in the B-to-S transition [171].

An important aspect of our theory is that the probability of segments in the overstretch state

P_s and the width of the transition δF depend on the cooperativity unit l , temperature T and the parameters A_2 and A_4 . But, as long as the ratio A_2/A_4 remains the same, the specific values of these parameters do not affect P_s and δF significantly. Therefore, our model can provide significant insight in the overstretching transition with knowledge of the relative extension of the overstretched phase $(\bar{L}/L - 1) \sim \sqrt{2A_2/A_4}$, which is an easy variable to determine in experiments.²

During DNA overstretching Bianco et al. [35] and Bongini et al. [34] found the transition state is almost midway between compact and extended states. This supports the use of a symmetric potential as done in this study. But, although our analytical model proves to be a useful tool by capturing some key features of the transition, there are still several interactions that could influence the transition, which our current model can not accommodate. For example, there is evidence that the percentage of GC vs AT tracts plays an important role in the transition [36], and for this level of detail, base-pair or rigid-base models are required [138, 139]. The slower relaxation rates of a partially melted molecule which result in different hysteretic behavior during the loading and unloading phase in overstretching experiments can not be captured by our symmetric model, and a dynamic mesoscopic Peyrard-Bishop-Dauxois potential has been proposed to account for this asymmetric behavior [170]. Similarly, we do not expect our model to capture the transition to single stranded DNA (peeled DNA), because thermal fluctuation will play a larger role in the extensibility of the peeled structure even at high forces, leading to markedly asymmetric force extension curves (softer after the transition). However, our methods can be combined with asymmetric potentials, such as piece-wise quartic energies with a cubic term (SI Appendix E.4), which could help to account for asymmetric behavior present in single molecule stretching experiments.

In addition to applications to DNA phase changes, two-state models have been used to interpret experiments on partially unfolded protein oligomers [49] and to study the mechanical behavior of small molecule binding to DNA. [59]. The binding of proteins or small-molecules to dsDNA can produce coupled conformational changes that affect the binding of subsequent proteins (molecules), often over a long range in the dsDNA [184]. Binding of the RecA protein to dsDNA lengthens the molecule by approximately 1.5 times and it has been shown that overstretching promotes RecA nucleation and polymerization along dsDNA [185]. Similarly, an approximately 50% increase in extension of DNA is observed upon interaction with EtBr and Rad 51 [27, 186]. Therefore, it is possible that in protein-dsDNA complexes that the conformational changes caused by small-molecule binding to be coupled with external mechanical forces. Our theoretical framework to study the dsDNA overstretching transition can be useful in describing certain aspects present during conformational changes in DNA-protein complexes.

²In Fig.E.5 we present solutions of P_s and $\langle u \rangle$ for different values of A_4 , while keeping the ratio A_2/A_4 constant.

Chapter 7

Concluding Remarks and Future Directions

The contents in this Thesis are a step toward building a better understanding of DNA behavior in response to mechanical loads. By modeling a wide range of DNA single molecule experiments (at low and high tensions, positive and negative twist, monovalent and multivalent ions at high and low concentrations, drug-DNA interactions, etc) we have contributed to the construction of a framework to study DNA conformational changes and DNA phase transitions. Here we give a brief summary of our accomplishments and point to future directions.

In chapter 2 we found that our mechanical model for supercoiled structures combined with a careful choice of electrostatic and entropic models allows for more accurate predictions of the onset of the transition to plectonemic DNA. In chapter 3, in the case of condensing agents, we find that for each tension there is a critical number of end rotations above which the supercoiled solution is preferred and below which toroids are the preferred state. In chapter 4 we developed a model for DNA supercoiling where we relax the constraint of a uniform helical angle and allow for a non-constant pitch super-helix. This model is relevant in DNA braiding problems where two double stranded molecules are wound around each other. For DNA braids we found that it is possible that the length of the super-helix is controlled by the maximum super-helical angle (geometrical constraint). One important limitation of our description of DNA compaction is that our methods are only applicable for static experiments. This is not always the case. Making use of a novel experimental method van Loenhout et al. [187] visualized the formation and diffusion of super-helices. This new set of experiments will surely serve as propelling factor in the development of dynamical models of supercoiled structures in the near future.

In chapter 6, we developed an analytical model for the overstretching transition that allows for the possibility that the overstretched state is mixture of two phases – M-DNA (melted DNA with portions of inner strand separation) and S-DNA (a phase that retains the base pair structure). Our model accurately reproduces both equilibrium and kinetic behavior seen in a large number of experiments conducted in various labs. We find that the cooperativity of the transition is a key parameter that characterizes the fraction of the phases in the overstretched state. Although our analytical model is a useful tool that can capture some key features of the overstretching transition, there are still several factors that could influence the transition which our current model can not accommodate. For example, the slower relaxation rates of M-DNA which result in different hysteretic behavior during the loading and unloading phase in overstretching experiments [34] can not be captured by our symmetric model. Similarly, our model can not describe the transition to single stranded DNA (peeled DNA), because thermal fluctuations play a larger role in the extensibility of the single strands even at high forces leading to markedly asymmetric force extension curves. Hence, future work should include the use asymmetric potentials together with the overstretching transition framework we have

provided in this Thesis.

In chapter 5 we study L-DNA both as a single state and a mixed state. Making use of a heterogeneous twisted-fluctuating-rod model, we have examined the hypothesis that L-DNA is a mixture of two relatively well-characterized DNA phases - S-DNA and Z-DNA. We found that the existing L-DNA experimental data from two groups ([29, 31]) can be explained under the assumption of a mixed L-DNA state. We think that if L-DNA is in fact a mixed form, then the fractions of Z-DNA and S-DNA can change as a function of ionic concentration. The results drawn from the mixed L-DNA hypothesis are not conclusive and more experimental data is necessary to verify it. Future experimental and theoretical work on the effects of ionic concentrations in L-DNA can help clarify the nature of L-DNA structure.

Another important aspect of DNA biological function is its interaction with proteins, small molecules and other polymers. Small molecules or drugs can bind or intercalate between its bases inducing conformational changes and modifying locally the DNA mechanical response to loads. For this reason, a system of DNA interacting with small molecules can be seen as mixed state, where one state is the region of pure DNA and the other state is the region where the small molecules bind. In chapter 5, we have used a heterogeneous twisted-rod model to describe DNA interactions with ethidium bromide (EtBr). The results of our model are in good agreement with single molecule experiments on the DNA-EtBr compounds [27]. Besides EtBr there are numerous other agents that mechanically interact with DNA such as netropsin, RecA, and Rad 51, to name a few. All these small-molecules are known to produce structural modifications to DNA and it is possible that some of the mechanically induced transitions (denaturation, S-DNA, Z-DNA) are intermediate states that facilitate the structural changes triggered by the small molecules (drugs). To extend our current understanding of DNA function and interaction inside the cell, it is important that future efforts focus on developing theoretical models for DNA-protein and DNA-drug complexes.

Appendix A

Appendix - Chapter 2

A.1 The end loop

At one end of the plectoneme there is a loop. The end loop is formed in the transition from the straight DNA configuration to a plectonemic DNA configuration. In the classical theory, the loop is formed when a rod subjected to tension and twist (applied number of turns) undergoes localized buckling at a critical torque $2\sqrt{K_b F}$ [78, 148, 188] up to a point where there is a dynamic jump into self-contact. The localized solution to the equilibrium equations of the rod is unstable, and consequently the perturbed rod jumps dynamically either to the straight rod or the loop configuration with self-contact [188, 189]. Formulations of the rod with contact points have been studied for both closed and open rods with applications to DNA supercoiling [66, 126]. As pointed out by Daniels et al.[74] in the DNA case, the transition happens due to free energy minimization and not due to instability or buckling. Due to thermal fluctuations the system can be perturbed sufficiently to go from the straight configuration into a lower and stable energetic state, which for a given applied torque M_{ext} (number of turns n), we predict to be the plectonemic state. So a DNA strand subjected to tension and controlled number of turns does not reach the classical critical buckling torque $2\sqrt{K_b F}$, and hence $M_{critical} < 2\sqrt{K_b F}$. The jump in the external moment can be characterized as the difference $\delta M = M_{critical} - M_3$, where M_3 is the torque in the plectonemic state. Strick et al.[44] show a measurement of the critical torque at the transition point based on the minimization of energy of an initial (circular) loop model. This calculation of $M_{critical} = (2K_b F)^{1/2}$ is approximate since it neglects the thermal fluctuations in the loop and assumes a circular geometry. We propose a better approximation to account for the end loop based on a localizing solution of the rod. To our knowledge the energy stored in the loop derived from an analysis including bending, twist and thermal fluctuation has not yet been carried out. Coyne [78] analyzed the formation of loops in twisted semi-infinite rods, providing expressions for the energy of the buckled-loop configuration without self-contact. In the limit when $M_{ext} = 0$, the Coyne expressions reduce to the expressions given by Kúlic et al.[79] without thermal fluctuations. We will assume that thermal fluctuations are negligible in the end loop [80]; this is a good approximation when the loop has small average radius of curvature. In the case of the planar homoclinic loop under tension with no moment applied at the ends, Kúlic et al. [79] show that the expressions for the free energy in the straight plus loop and straight configurations differ by an amount equal to the elastic energy present in the loop. Their result is given below and takes into account both the bending energy and the work against the end force F :

$$E_{loop} = (E_{o-bend} + F) L_o = 8\sqrt{K_b F} = 2FL_o \quad (\text{A.1})$$

where $L_o = 4\sqrt{\frac{K_b}{F}}$. In the absence of thermal fluctuations, but including twist, the free energy of the loop (E_{loop}) is given by [78, 148]:

$$E_{loop} = \left(\frac{M_{ext}^2}{2K_t} + E_{o-bend} + F \right) L_o, \quad (\text{A.2})$$

where

$$E_{o-bend} = F, \quad L_o = 4\sqrt{\frac{K_b}{F}} \left(1 - \frac{M_{ext}^2}{4K_b F} \right)^{1/2}.$$

The expression for the writhe present in the loop is [148]:

$$Wr_o = \frac{2}{\pi} \cos^{-1} \left(\frac{M_{ext}}{2\sqrt{K_b F}} \right). \quad (\text{A.3})$$

The expressions for E_{o-bend} and Wr_o that we pick have to satisfy the condition that the number of turns $2\pi n$ is conjugate to the applied torque M_{ext} . We see that $Wr_o = 1$ only for $M_{ext} = 0$, becoming a planar homoclinic loop as in the case analyzed by Kúlic et al. [79]. Based on this idea we will approximate the energy of the loop by decoupling the bending and twisting energy, such that the E_{o-bend} and L_o are given by Kúlic et al. [79] formulae and the twist energy of the loop is the first term in Eq. (A.2).

A.2 Testing the Internal Energy models

In Clauvelin et al. [73] the mechanical description is combined with different analytical theories of DNA-DNA interactions that can be found in literature. The work in Clauvelin et al. [73] picked two well established models. The first is U_{PB} , derived by Ubbink and Odijk [20] from the Poisson-Boltzman equation; the second is U_{cc} , derived by Manning [190] and is based on the counterion condensation theory. According to the results obtained in Clauvelin et al. [73], an approximation of $U_{PB}(r, \theta)$ provided better agreement with experiment. When the undulations in the radial direction are not restrained by electrostatics but only by the structure of the plectoneme, the variable d_r will not appear in the electrostatic expression $U_{PB}(r, \theta)$ [73]:

$$U_{PB}(r, \theta) = \frac{1}{2} k_B T \nu^2 l_B g(\theta) \sqrt{\frac{\lambda_D \pi}{r}} e^{-\frac{2r}{\lambda_D}},$$

$$g(\theta) = 1 + 0.83 \tan^2(\theta) + 0.86 \tan^4(\theta). \quad (\text{A.4})$$

Similarly r will replace d_r in the configurational entropy expression given by Ubbink and Odijk [20] as done in Marko and Siggia[52]:

$$U_{conf-MS}(r, \theta) = \frac{k_B T}{A^{1/3}} \left[\frac{1}{(p\pi)^{2/3}} + \frac{1}{r^{2/3}} \right], \quad (\text{A.5})$$

where the constants $c_r = c_p = 1$. Additionally, Marko and Siggia [52] developed an analytical model for the electrostatic interactions that has also been used in the study of DNA single molecule experiments [4, 72]. The expression for the Marko and Siggia electrostatic model $U_{MS}(r, \theta)$ is:

$$U_{MS}(r, \theta) = l_B k_B T \nu^2 \left[K_0 \left(\frac{2r}{\lambda_D} \right) + K_0 \left(\frac{\pi r \cot \theta}{\lambda_D} \right) \right], \quad (\text{A.6})$$

where $K_0(x)$ is the modified Bessel function of the second kind. A.1 summarizes the different models used to describe the internal energy interactions in our calculations, showing the figures where each one of them has been used in this section of the Appendix A.

Label	Model	Figures
U_1	$U_{PB}(r, \theta) + U_{conf-MS}(r, \theta)$	A.1, A.2
U_2	$U_{MS}(r, \theta) + U_{conf-MS}(r, \theta)$	A.1
U_3	$U_{PB}(r, \pi/6) + U_{conf-MS}(r, \pi/6)$	A.1
U_4	$U_{PB}(r, \theta, d_r) + U_{conf}(\theta, d_r)$	A.2, A.3
U_5	$U_{PB}(r, \theta)$	A.2

We obtain theoretical results under the experimental conditions of Forth et al. [12]. The experiments were performed in phosphate buffered saline with 150 mM NaCl at 23.5°C. Numerical calculations resembling the experiments were performed assuming $K_b = 50k_B T$ and $K_t = 95k_B T$. The values used for the electrostatic parameters are:

- The Bjerrum length is $l_B \approx 0.715$ nm [20].
- The Debye length $\lambda_D \approx 0.8$ nm.
- The effective charge $\nu = 8.06$ nm⁻¹, where an interpolation of values listed in Table 7 in Ubbink and Odijk [20] has been used. In the main text the effective linear charge ν is treated as a fitting parameter.

Internal energy models: effects and comparison

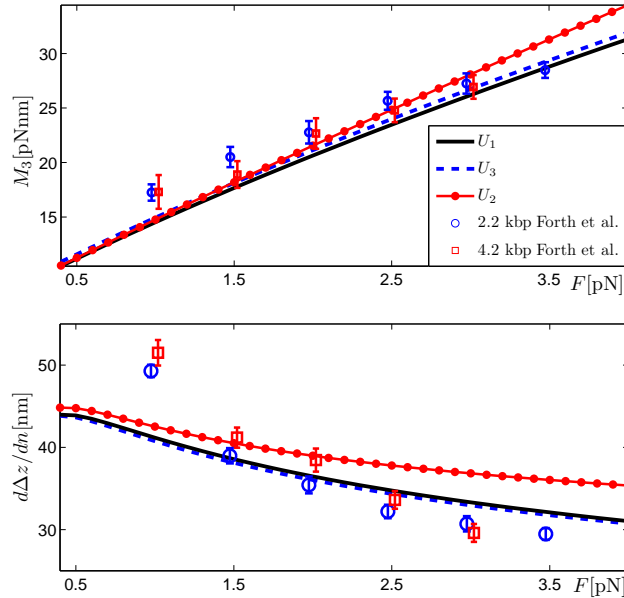


Figure A.1: Comparison between the solutions based on different internal energy models under the experimental conditions of Forth et al. [12]. The black solid line uses $U_1 = U_{PB}(r, \theta) + U_{conf-MS}(r, \theta)$, the red dotted line uses $U_2 = U_{MS}(r, \theta) + U_{conf-MS}(r, \theta)$ and the blue dashed line uses $U_3 = U_{PB}(r, \pi/6) + U_{conf-MS}(r, \pi/6)$ which is a function only of r . The value of the effective linear charge used is $\nu=8.06$ [nm⁻¹].

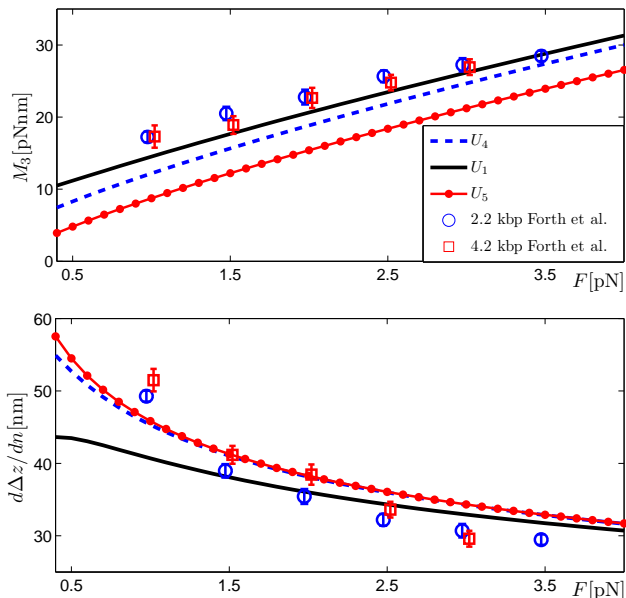


Figure A.2: Comparison between the solutions based on different internal energy models under the experimental conditions of Forth et al. [12]. The black solid line uses $U_1 = U_{PB}(r, \theta) + U_{conf}(r, \theta)$, the blue dashed line uses the internal energy model where the radial fluctuations in the plectoneme are constrained by the electrostatics $U_4 = U_{PB}(r, \theta, d_r) + U_{conf}(r, \theta, d_r)$ [20] and the red dotted line uses a model that neglects the configurational entropy contribution to the free energy $U_5 = U_{PB}(r, \theta)$. The qualitative behavior of M_3 and $d\Delta z/dn$ are independent of the internal energy models we have used, but the quantitative agreement with experimental data strongly depends on the choice of configurational entropy model. The value of the effective linear charge used is $\nu=8.06$ [nm⁻¹].

We show in Fig. A.1 a comparison of the results obtained using Marko and Siggia's $U_2(r, \theta)$ electrostatic model with variants of the Poisson-Boltzmann model: $U_1(r, \theta)$ and $U_3(r, \pi/6)$.

The helical angle θ , using the U_1 and U_2 models, increases slightly as a function of the applied force F and it is approximately equal to $\pi/6$ for large forces, while the helical radius r decreases as a function of F and approaches the crystallographic radius 1nm (not shown). Since θ does not vary dramatically as a function of the applied force note that the curves obtained using the approximation $U_3(r) = U_{PB}(r, \pi/6) + U_{conf-MS}(r, \pi/6)$ are very close to those obtained from the internal energy models with θ dependence. Using $\theta = 0$, where the angle dependence is neglected as done in Clauvelin et al.[45, 73] and Neukirch and Marko [4] lowers the predicted values of M_3 and $d\Delta z/dn$ (not shown). The analytical prediction of the slope $d\Delta z/dn$ using $U_3(r)$ and the prediction using $U_1(r, \theta)$ are almost identical to each other. It is clear from the graphs that the three approaches produce consistent results for the values of M_3 although the U_2 model gives slightly larger values of M_3 for $F > \sim 2.5pN$. The difference between the U_2 model and the U_1 model at moderate and large F is more evident in the predicted values of $d\Delta z/dn$, where the U_2 model predicts larger slopes. We also performed calculations for different salt concentrations and observed similar trends (not shown).

In Fig. A.2 we compare the solution obtained by using $U_1(r, \theta)$ with the solution obtained by using the undulation-enhanced free energy model $U_4(r, \theta, d_r)$ with empirically optimized coefficients by van der Maarel[84] $c_p = c_r = 2^{-8/3}$. Fig. A.2 also depicts the solution obtained by picking $U_5(r, \theta) = U_{PB}(r, \theta)$ with $c_p = c_r = 0$ such that the configurational entropy effects are neglected as done in Clauvelin et al. [73]. It is clear from Fig. A.2 that the final qualitative behavior of the curves is the same, but the quantitative agreement strongly depends on the choice of the internal energy, in particular, of the configurational entropy model and coefficients. When comparing U_1 and U_5 we can see that neglecting the entropy effects reduces the internal energy U , increases the value of $d\Delta z/dn$ (mainly at low forces $F < \sim 2pN$) and reduces the predicted value of M_3 (in the whole

range of F). The U_5 model neglecting entropy effects used by Clauvelin et al. [73] matches the slope predictions using the U_4 model, but the theoretical results for M_3 using U_5 are lower than the ones predicted using the U_4 model.

Now we are in position to understand how combining different approximations and assumptions can counteract each other. Starting with the U_2 model, if we next neglect the entropic effects as in $U = U_{MS}(r, \theta)$, we get larger slopes at low forces (matching qualitatively better the shape of the experimental trend in Fig. A.1), but we still overestimate the values of the slopes at larger F . This assumption also decreases the predicted values of M_3 . If we further assume $U = U_{MS}(r)$ where there is no angle dependence, the values of the slopes in the whole F range would decrease giving better quantitative agreement with experimental data of the slopes in Fig. A.1, and it will decrease the predicted values of M_3 even more. Therefore using the $U_{MS}(r)$ model would 'seem' to accurately match the experimental slopes but it will underestimate the values of M_3 for the whole range of F , especially at low values of $F < \sim 2$ pN.

As mentioned in the main text although the data sets for the slopes from different experimental groups and MC simulations agree quantitatively [5, 10–12], this is not the case for the direct and indirect measurements of the torque M_3 . The indirect torque measurements in Mosconi et al. [11] are significantly smaller than the direct measurements taken by Forth et al. [12], Lipfert et al. [27] and the MC simulations in Maffeo et al. [5]. The $U_{MS}(r)$ model combined with a mechanical description in Neukirch and Marko [4] seems to give good agreement with the indirect measurements of the torque M_3 in Mosconi et al. [11], specially at large forces and accurately describe the slope data of the same experimental group. In Maffeo et al. [5] the authors provide also an analytical model that matches the experimental data in Mosconi et al. [11], but does not match the predicted M_3 and r from their MC simulations. In their supplemental material, Maffeo et al. discuss the reasons for the success of their approach as well as its disadvantages and limitations of neglecting fluctuations and entropic terms. They conclude that the reason why their analytical predictions of M_3 are lower by ≈ 1.5 pNm than their MC simulations is due to neglecting configurational entropy and fluctuation effects. They show that neglecting the entropic effects and undulation enhancement decreases significantly the theoretical predictions of M_3 , while the effects in the theoretical slope predictions is not so drastic. The reason for the drastic increase in the analytical values of M_3 computed in Maffeo et al. [5] when using the entropic model proposed by Ubbink and Odijk [20] lies in the fact that the authors have used $c_p = c_r = 3/2^{8/3}$, which increases the entropic contribution by 3 times compared to the constants used in our $U_{PB}(r, \theta, d_r)$ model in the main text. The constants $c_p = c_r = 3/2^{8/3}$ were derived for a one dimensional worm-like chain confined in a harmonic potential. In the main text we have decided to use the values $c_p = c_r = 2^{-8/3}$ as van der Maarel [84] has suggested.

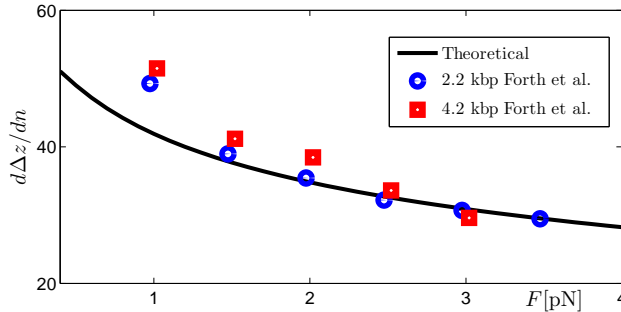


Figure A.3: Experimental data for the slope of the rotation-extension curve for two DNA templates taken from Forth et al. [12] for a 150mM salt concentration. We have used $\nu = 5.93\text{nm}^{-1}$.

From Figs. A.1 and A.2 we see that the experimental slopes of Forth et al. [12] are better described qualitatively by using the internal energy model $U_4(r, \theta, d_r)$, and that the theoretical predictions of M_3 using U_4 underestimate the experimental measurements of Forth et al. [12]. Using $\nu < 8.06\text{nm}^{-1}$

in $U_4(r, \theta, d_r)$ gives better quantitative agreement for the slopes, so in the main text we have used the effective linear charge as a fitting parameter. Fig. A.3 shows the results of $d\Delta z/dn$ from our theoretical model using $\nu = 5.93\text{nm}^{-1}$.

This survey of some of the internal energy models used to study DNA single molecule experiments leads to the conclusion that a careful choice of entropic and electrostatic parameters is needed to quantitatively match the experimental data. We have found that including angle dependence, configurational entropy and undulation enhanced effects due to thermal fluctuations in the helices are essential to have an accurate and complete model of plectonemic DNA. The $U_4(r, \theta, d_r) = U_{PB} + U_{conf}$ [20] model used in the main text provides the best results for a wide range of experimental data.

A.3 Critical torque $M_{critical}$ and the jump $\delta M = M_{critical} - M_3$

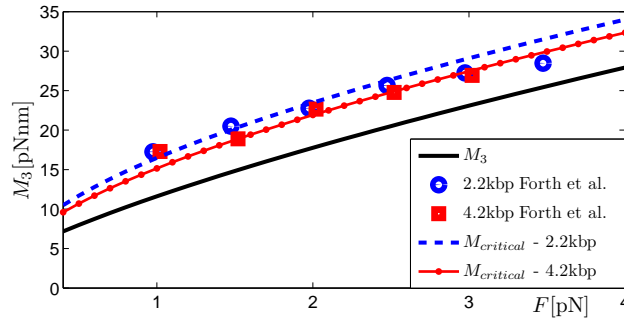


Figure A.4: Experimental data for two DNA templates taken from Forth et al. [12] for a 150mM salt concentration. The experimental data of the plectonemic torque in Forth et al. [12] agrees qualitatively with our predictions but seems to match quantitatively our predicted values of $M_{critical}$. We have used $\nu = 5.93\text{nm}^{-1}$.

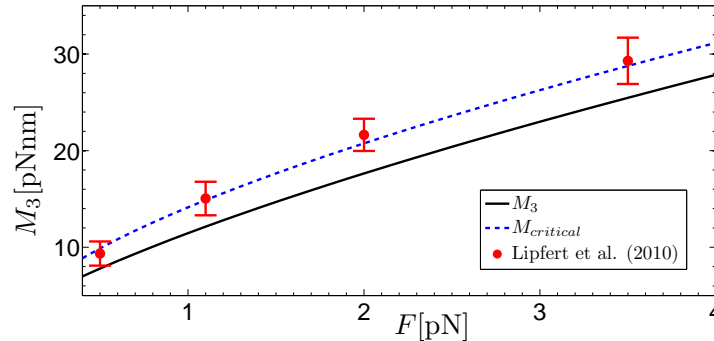


Figure A.5: External torque M_3 and critical torque $M_{critical}$ as a function of the external force F for a 7.9 kbp DNA template in a $c_o = 150\text{mM}$ salt concentration using the same parameters as in Fig. A.3. The data points correspond to the experimental values for the ‘buckling’ torque reported in Lipfert et al. [27].

The experimental data in Forth et al. [12] shows the jumps in the vertical extension but does not show a clear jump in the torque, and consequently there is no clear distinction between the torque before and after the transition. As shown in Fig.A.4 the experimental data of the plectonemic torque in Forth et al. [12] agrees qualitatively with our twisting moment predictions M_3 , but seems to match quantitatively our predicted values of the critical torque $M_{critical}$. Lipfert et al. [27] performed single molecule measurements in PBS buffer at $c_o \approx 150\text{mM}$ using a 7.9 kbp DNA template. The data in

Lipfert et al. [27] does not show the transition jumps in either torque or extension. Lipfert et al. [27] just report a ‘buckling torque’ by not making a distinction between the torque before the transition $M_{critical}$ and the plectonemic torque M_3 . Fig. A.5 shows excellent agreement between our predicted values of $M_{critical}$ and the ‘buckling’ torques in Lipfert et al. [27].

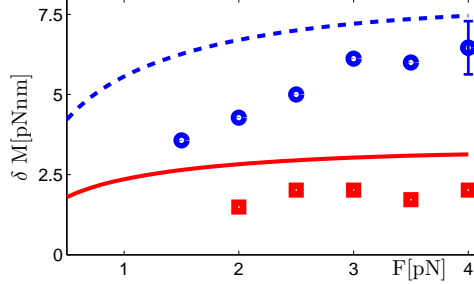


Figure A.6: Comparison of the the torque jump δM from our theory with the experiments in Brutzer et al. [10] at $c_o = 320\text{mM}$. Our model predicts that as the DNA length L increase δM decreases while as c_o decreases δM decreases. We have used $\nu = 10.00\text{nm}^{-1}$ as in the main text.

In Fig.A.6 we show the comparison between δM from the indirect measurements in Brutzer et al. [10] and our theoretical predictions. Similar qualitative trends are found in the indirect measurements of δM in Daniels et al. [74].

A.4 Indirect method for calculating external moment

Mosconi et al. [11] provide not only the direct measurements of the slopes $d\Delta z/dn$ of the rotation-extension curves of a single stretched and twisted DNA molecule using magnetic tweezers, but also an indirect measurement of the plectonemic torque M_3 . The theoretical predictions for the slopes $d\Delta z/dn$ presented in the main text match the experimental results in Mosconi et al. [11], but there seems to be constant 2.5pNnm offset between our theoretical predictions of M_3 and their reported indirect measurements. The process used by Mosconi et al. [11] computes the external torque M_3 from equation (23) in Zhang and Marko[85]:

$$M_3(F, n) = M_3(F_o, n) - \frac{1}{2\pi} \int_{F_o}^F \left(\frac{\partial \rho L}{\partial n} \right)_{\hat{F}} d\hat{F}, \quad (\text{A.7})$$

where F_o is the force corresponding to the initial reference rotation-extension curve. Eq. (A.7) is based on the ‘Maxwell’ type relation:

$$- \frac{1}{2\pi} \frac{\partial \rho L}{\partial n} \Big|_F = \frac{\partial M_{ext}}{\partial F} \Big|_n. \quad (\text{A.8})$$

The method to compute the external torque described in Zhang and Marko [85] assumes the existence of an equilibrium ensemble. So equation (23) in Zhang and Marko [85] is valid in both the straight and plectonemic states. But, at the transition point the system undergoes a dynamic jump from the straight to the plectonemic state or vice-versa. Hence, the method described by Zhang and Marko [85] can not be properly used since the first derivatives of the free energy become discontinuous at the jump. The process carried out to compute the external torques in the plectonemic regime in Mosconi et al. [11] neglects the presence of these jumps. The resolution of the experiments in Mosconi et al. [11] is such that the dynamic process at the transition point between the extended DNA configuration and the plectonemic configuration is not apparent. So, the rotation-extension experimental curves do not show a jump in the extension δz and consequently the external torque curves reported by

Mosconi et al. [11] show a smooth transition as a function of the degree of supercoiling $\sigma \propto n$. Accounting for the jumps precludes the use of Eq. (A.7) as done by Mosconi et al. [11].

We note, however, that the method described in Zhang and Marko [85] is valid in the plectonemic regime, as long as there is no dynamic jump. For a set of data containing rotation-extension curves at different values of applied F (for a fixed salt concentration), Eq. (A.7) can be used to compute a *change* in M_3 by holding n constant and integrating with respect to F . In figure 1 of Mosconi et al. [11] the change in torque M_3 between points A and B can be computed using Eq. (A.7), but a reference *absolute* value of the torque cannot be obtained using Eq. (A.7) due to the presence of the jump discontinuity. Since the method described in Zhang and Marko [85] can give the change in M_3 in the plectonemic regime, in Fig.A.7 we compare our theoretical predictions of the external torque M_3 with the indirect measurements of Mosconi et al. [11] after adding a constant value of 2.5pNm to the experimental data. The agreement is excellent.

Another possible explanation for the disagreement among the reported theoretical and experimental values of the external moment could be that the values of M_3 in the plectonemic regime are a function of the length of the DNA template, which is a parameter that is not captured in the existent theoretical models including our formulation.

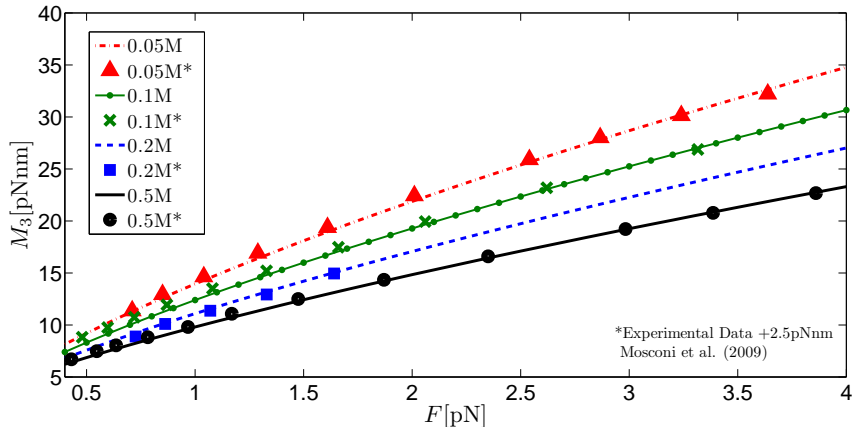


Figure A.7: External torque M_3 as a function of the external force F for the different salt concentrations in Mosconi et al. [11]. We show the experimental values from Mosconi et al. [11] after adding a shift of +2.5pNm. We have used ν values from Table2.1 in the main text.

A.5 Multivalent ions

In the main text we have explained how our model can be extended to the case of mixtures of high concentration of monovalent salt and low concentration of polyvalent salt. Here we present further comparison of the theoretical model with the experiments of Dunlap and co-workers (private communication) and show the behavior of the plectoneme radius r and plectonemic torque M_3 as a function of the applied force F . In our calculations we have used ν values from Table2.2 in the main text. Fig. A.8 shows r and M_3 values when using different c_{mu} concentrations of spermidine Sp^{3+} . As c_{mu} increases the supercoiling radius decreases, yielding more compact plectonemes. For $c_{mu} = 5mM$ and $c_{mu} = 10mM$ the value of r reaches the limiting interaxial spacing value $\sim 3nm$ as given in Todd et al.[13] and Raspaud et al.[14]. As explained in the main text we expect that due to a balance of attractive and repulsive interactions the plectoneme diameter stays approximately constant when it reaches the limiting interaxial spacing value. The values of M_3 decrease as a function of c_{mu} . The dashed lines in Fig. A.8 show M_3 as a function of F for $c_{mu} = 5mM$ and

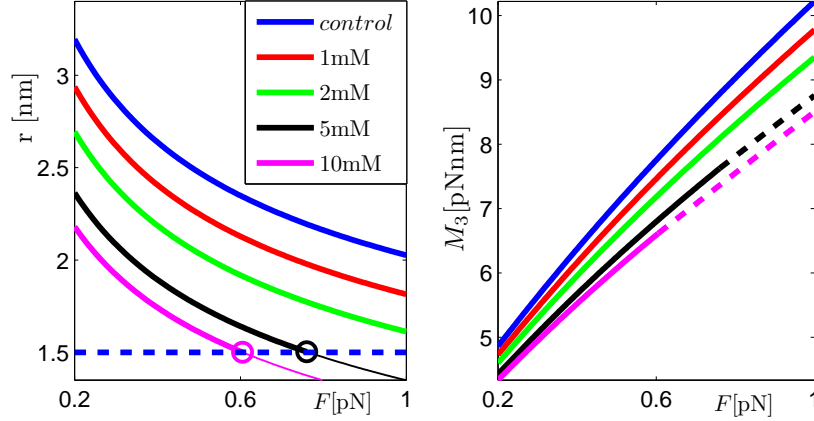


Figure A.8: Spermidine: M_3 and r . The dashed lines correspond to the limiting value of the interaxial spacing which is $2r = 3.0\text{nm}$ for spermidine. We have used ν values from Table2.2 in the main text.

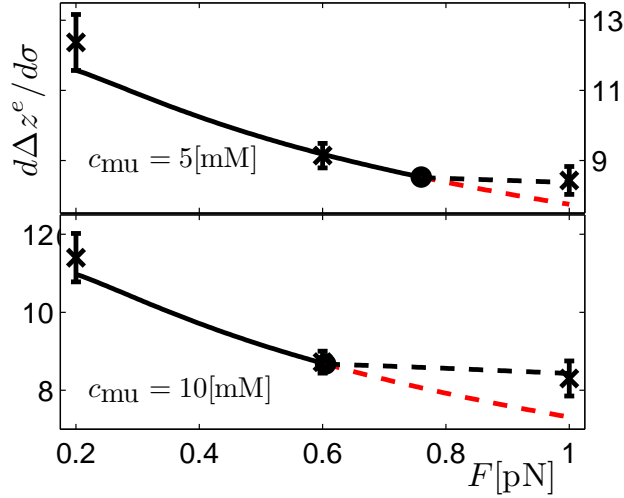


Figure A.9: Spermidine: the black solid curve shows the prediction of the slope $d\Delta z^e/d\sigma$ when $2r$ is larger than the interaxial spacing 3nm . The black dot shows the point where $2r = 3\text{nm}$, and from there on the black dashed line shows the prediction of the slope for $2r$ constant. The red dashed line shows the solution when $2r < 3\text{nm}$ is allowed to vary.

$c_{mu} = 10\text{mM}$ using $r \approx 1.5\text{nm}$ after the plectoneme diameter has reached the limiting interaxial spacing value. The theoretical solution of M_3 for $c_{mu} = 5\text{mM}$ and $c_{mu} = 10\text{mM}$ when r is not assumed to be constant does not differ significantly from the solutions plotted in Fig. A.8. The theoretical solution of the slope $d\Delta z^e/d\sigma$, for $c_{mu} = 5\text{mM}$ and $c_{mu} = 10\text{mM}$, when r is not assumed to be constant differs considerably from the solutions plotted in the main text in Figs.2.8 and 2.9 where r is constant after reaching the interaxial spacing value. If $2r$ is allowed to become smaller than the interaxial distance, the predicted slopes $d\Delta z^e/d\sigma$ underestimate the experimental data at $F = 1\text{pN}$ as shown by the red dashed line in Fig. A.9.

In Fig A.10 we present the critical degree of supercoiling $\sigma_{critical} \propto n_{critical}$ when using different c_{mu} concentrations of Sp^{3+} . As stated before the dashed lines in the $c_{mu} = 5\text{mM}$ and $c_{mu} = 10\text{mM}$ cases represent the solution when $2r$ reaches the interaxial spacing value.

Figs. A.11 and A.12 show the results obtained when using spermine Sp^{4+} . For Sp^{4+} we have only

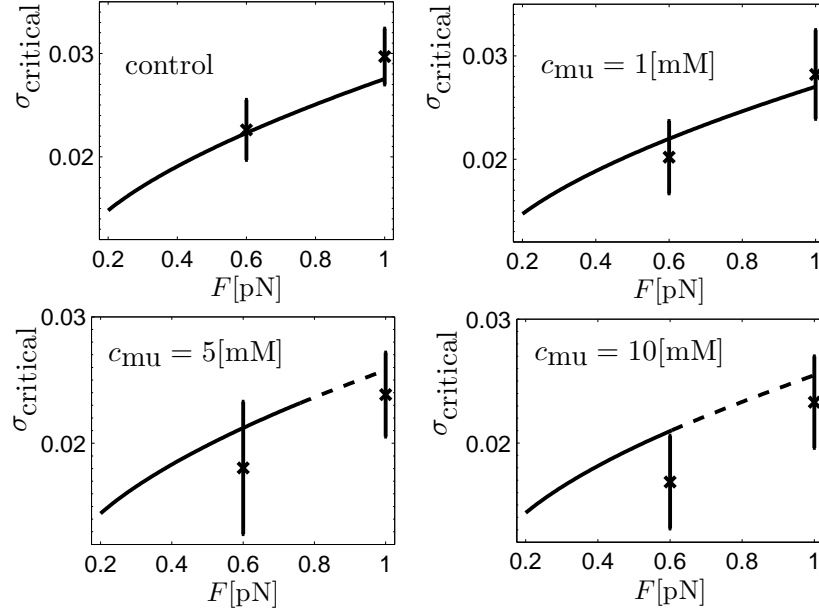


Figure A.10: Spermidine: critical degree of supercoiling.

fitted the value of ν to the experimental point $F = 0.6\text{pN}$ for the control and $c_{mu} = 0.2 - 0.75\text{mM}$ concentrations and obtained a curve for ν as a function of c_{mu} (see entries in Table2.2 in the main text). For $c_{mu} = 1\text{mM}$ and $c_{mu} = 2\text{mM}$ entries in Table2.2 shown in the main text we have extrapolated the value of ν from the curve obtained from the previous fitted values. As before, the dashed lines for the M_3 and $\sigma_{critical}$ function correspond to the solution when $2r$ is approximately constant and equal to the interaxial spacing $\sim 2.9\text{nm}$ for Sp^{4+} (Todd et al. [13], Raspaud et al. [14]). Our results show good quantitative agreement with the experimental values.

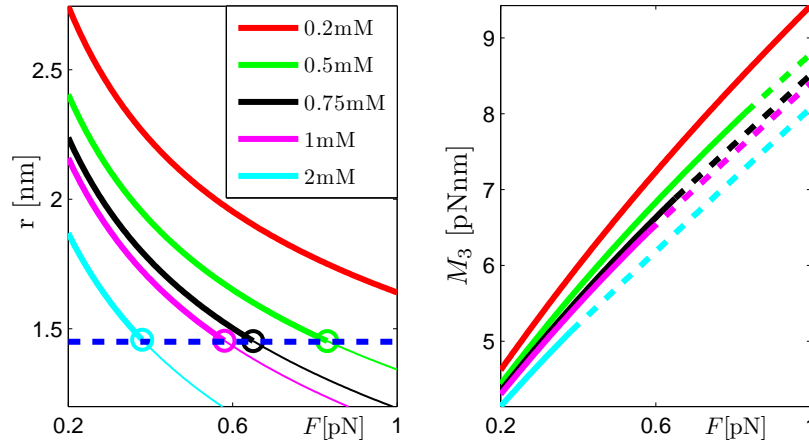


Figure A.11: Spermine: M_3 and r . The dashed lines correspond to the limiting value of the interaxial spacing which is $2r = 2.9\text{nm}$ for spermine.

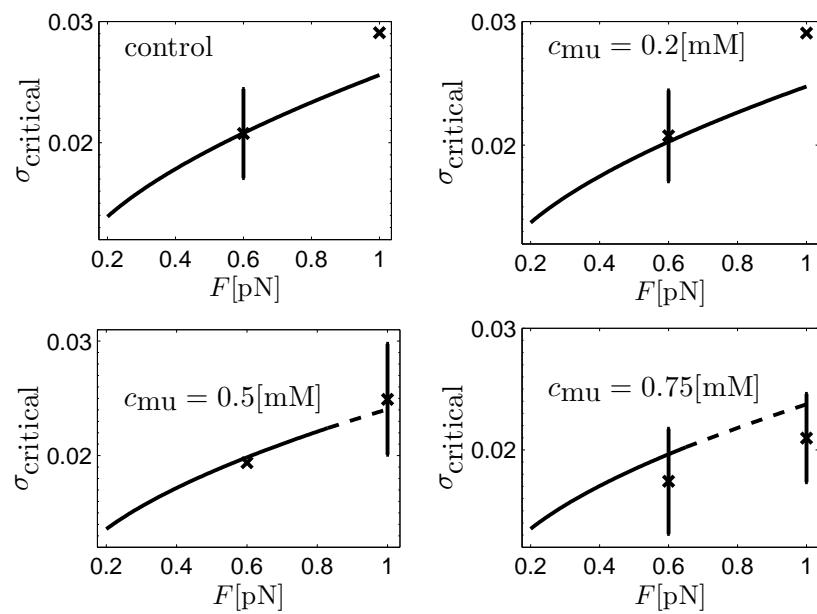


Figure A.12: Spermine: critical degree of supercoiling.

A.6 Experimental and Simulation Data

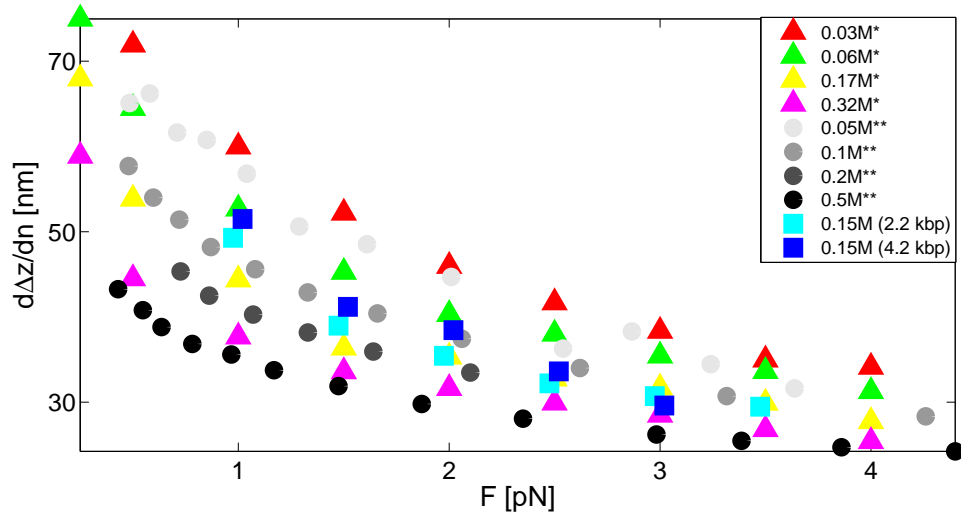


Figure A.13: Experimental data for the slopes $d\Delta z/dn$ for different salt concentrations. Triangles represent Brutzer et al. [10] data presented in reference [5]. We have denoted Brutzer et al. data with * next to the salt concentration value in the legend of the graph. Circles represent Mosconi et al. [11] data, which have been denoted with ** next to the salt concentration value in the legend. Squares represent Forth et al. [12] data for two different values of the DNA length. The data sets from Brutzer et al. and Mosconi et al. provide consistent slope values for the entire force range. The slopes from Forth et al.[4] are consistent with the rest of the data sets for moderate forces.

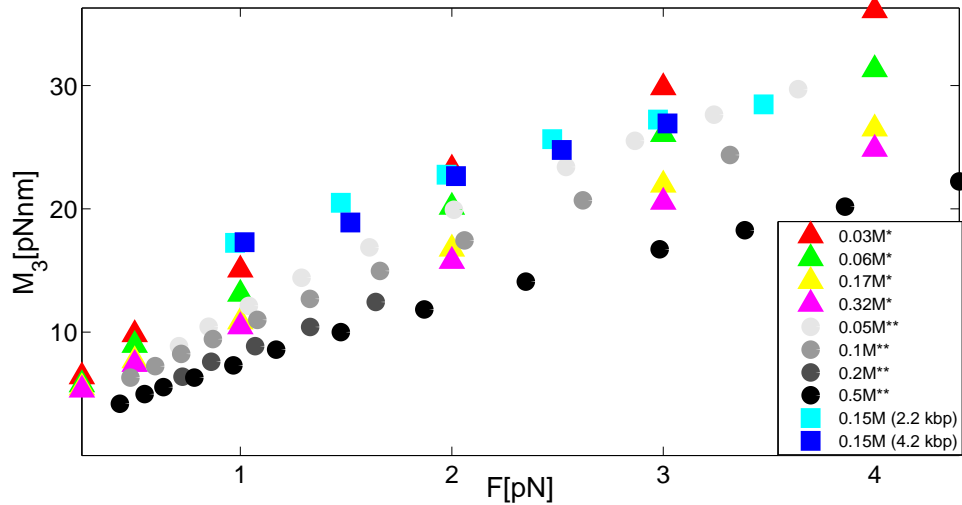


Figure A.14: Experimental and Simulations data for the torque measurements M_3 for different salt concentrations. Triangles represent Maffeo et al. [5] simulations data. We have denoted Maffeo et al. data with * next to the salt concentration value in the legend of the graph. Circles represent Mosconi et al. [11] data, which have been denoted with ** next to the salt concentration value in the legend. Squares represent Forth et al. [12] data for two different values of the DNA length. The various data sets show disagreement in the torque values.

Appendix B

Appendix - Chapter 3

B.1 Solution for the supercoiling parameters

Comparing the different types of lines in Fig. B.1 shows how the solution for the supercoiling variables M_3, θ, r and $d\Delta z/dn$ (slope of the extension vs. rotation curves) varies as a function of the constant parameters C_A, D_{eq}, K_b . The effect of C_A can be seen by comparison of the black solid line and the blue dashed-dot line, the effect of K_b can be seen by comparison of the black solid line and the green dotted line. The slope $d\Delta z/dn$ of the hat-curves (which can be experimentally measured [10–12]), showing the end to end distance as a function of the applied number of turns is equal to $\rho dl_p/dn$, where

$$\rho = 1 - \frac{1}{2\sqrt{K_3^2 - (1/32)}} + \frac{K_b}{k_B T K_3^2 l_t} \quad (\text{B.1})$$

is the effective mean extension due to thermal fluctuations [40], K_3 is given by Eq. (3.2) (replacing M by M_3) in the main text and dl_p/dn is given by Eq. (3.18) in the main text.

Fig.B.2 shows the toroid dimensions as a function of the number of turns n . The predicted values of R and a are of the same order as the ones observed in experiments with unconstrained DNA $R \sim 50\text{nm}$ and $a \sim 25\text{nm}$ [97, 106].

B.2 Alternative energy expressions for the toroidal condensates

We can also use the internal energy expression $U = H_{int}$ as suggested by Battle et al. [98] to obtain similar results for the scaling behavior of R_o and N_g^o as in the main text. In the case of the toroids we replace the electrostatic energy term $G_g L$ and the energy associated with the surface tension E_s in the energy expression Eq. (3.22) of the main text, with Battle et al. expressions such that:

$$V_g = \omega L \left(-3 + 2\sqrt{\frac{3}{N_g}} \right) + \epsilon L \frac{K_B}{R^2}, \quad (\text{B.2})$$

where we have used the approximation of large- N_g behavior [98] and ω is the surface tension parameter. The equilibrium variables are obtained by minimizing the energy with respect to the linking number constraint Eq. (3.24) in the main text and the length constraint $L = 2\pi R N_g$. In the case of

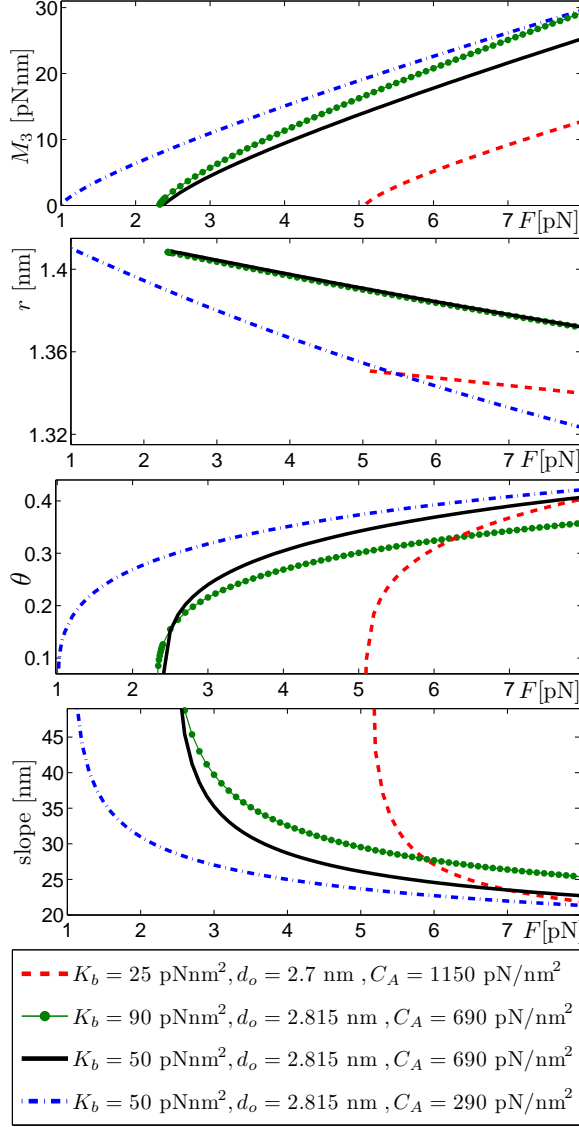


Figure B.1: An increase in C_A : (a) increases the threshold force F_T , (b) decreases M_3 and θ , (c) increases $d\Delta z/dn$, and (d) causes r to deviate less from the constant value $D_{eq}/2$. An increase in K_b : (a) slightly decreases F_T , (b) causes no significant difference in r (lines are almost on top of each other), (c) decreases θ , and (d) increases both M_3 and $d\Delta z/dn$. An increase in D_{eq} : (a) significantly decreases F_T , (b) shifts the solutions for θ and M_3 to the left along the force axis, (c) increases r , and (d) decreases $d\Delta z/dn$. We have used $K_t = 86k_B T$.

$M_g = 0$ we obtain:

$$R^o = \left(\frac{2\epsilon^2 K_B^2 L}{3\pi\omega^2} \right)^{1/5}, \quad N_g^o = \left(\frac{3\omega^2}{4\epsilon^2 K_B^2} \right)^{1/5} \left(\frac{L}{2\pi} \right)^{4/5}. \quad (\text{B.3})$$

The value of ω can be fitted to the case when $n = 0$ to obtain the experimental value of $F_{critical}$ in [15]. Using values of $\epsilon \in [1/2, 5/2]$, for spermine $\omega \sim 0.75\text{pN}$ and for cosep $\omega \sim 1.65\text{pN}$ gives good fits to experiment yielding toroids of radius $R \sim 30\text{nm}$ (Battle et al. uses $\omega = 1.6\text{pN}$).

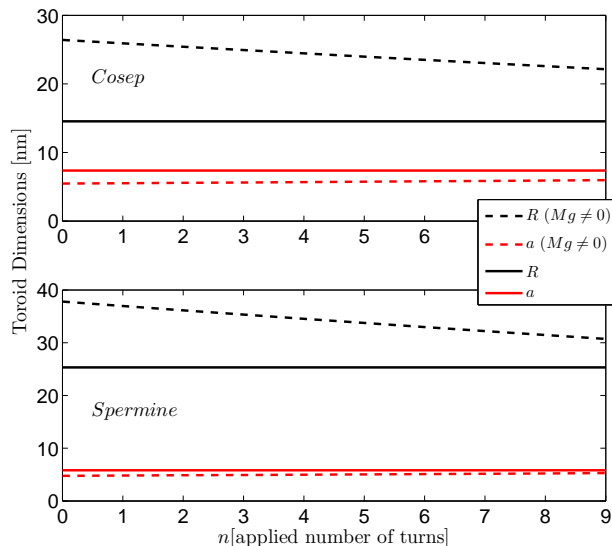


Figure B.2: Toroid dimensions as a function of the applied number of turns n . The solid lines correspond to the solution for toroids with no twist and dashed lines to toroids with twist.

B.3 Competition of Globules vs Plectonemes

For $n, F > 0$ there could be a competition between globules and plectonemes. The solution to the problem in thermodynamic equilibrium is therefore given by a comparison of the energies Eq. (3.9) and Eq. (3.22) (see main text) as a function of the state variables F and n . Figs. B.3 and B.4 present the energy density of competing structures as a function of the number of turns n for surface tension parameters $\alpha = 0.0013$ and $\alpha = 0.0003$ respectively. Figs. B.3 and B.4 are a schematic of the energy comparison calculations done to obtain Figs. 2-4 in the main text. The values of the energetic contributions $U(r, \theta)$, G_g and the surface tension contribution E_s are not precisely known. There are other energy terms that could be important in describing each configuration such as the energy cost due to defects in the toroid [117] or a more rigorous analysis of the energy cost upon confinement into condensed structures. Consequently, the comparison between these two states is not straight forward and could lead to more than one scenario. Below we describe the results obtained for some theoretically possible cases.

- We performed a calculation where we varied the persistence length $A \in [5, 150]$ nm. The bending energy in the plectonemic state decreases linearly with the persistence length, while the bending energy and the surface tension energy of the torus with no twist is proportional to $A^{1/5}$. So as A increases (decreases) the increase (decrease) of the supercoiled configuration energy V_{s-p} is larger than the increase (decrease) of the torus energy $V_g(M_g = 0)$. For small values of the persistence length A the equilibrium radius R^{tw} for toroids with twist ($M_g \neq 0$) is much larger the equilibrium radius R^o of the toroids with no twist ($M_g = 0$). Toroids with twist have a larger energy than the ones with no twist. But as A increases then $R^{tw} \rightarrow R^o$ and the gap in the energy between toroids with and without twist decreases.
- We also modified the parameter $\epsilon \in [1/2, 5/2]$ to account for configurational entropy cost of the toroidal condensates, where $\epsilon = 5/2$ if the energy cost is equal to confining a flexible chain in a sphere of radius R [116]. For larger values of ϵ (keeping all the other parameters fixed including α) the energy of the toroid increases, but note that the dimensions of the toroid without twist do not vary (see Eq. (3.29) in the main text). For larger values of ϵ we can still

accurately model Besteman et al. experimental data by fitting a smaller value of α (which is still in agreement with the range given in Refs. [116, 117, 120, 121]).

- We also computed the solution to the minimization of energy of the toroidal structure for different values of the surface tension parameter α . It is interesting to note that since $\alpha \propto c^{-5/2}$, for toroids with $M_g = 0$ and $M_g \neq 0$, the equilibrium radius R , the winding number N_g and the energy of the toroid V_g are independent of the value of the constant $c \in [1, 8\pi]$ (which stands for the shape imperfections of the toroid), but the thickness of the torus a is proportional to $c^{-1/2}$. As α decreases the toroid energy decreases and the the globular condensates become more energetically favorable. But as α decreases the globule winding number N_g^o decreases too, making toroids with out twist less likely to exist as a function of the number of turns n . Toroidal configurations start to compete with plectonemic structures for very small values of α (approximately an order of magnitude smaller than the ones used to compare to Besteman et al. experiments). For such values of α , the solution to the minimization of energy leads to equilibrium configuration where $R \sim 100\text{nm}$ for $n = 0$. For $n > n_{min}$, as F gradually decreases, even in the case where $\alpha \rightarrow 0$ (such that $N_g^o \rightarrow 0$ and $R \rightarrow 2\pi L$), straight DNA first transforms into a supercoiled-straight configuration and eventually can make a transition into a toroidal structure. So for a certain range of n the DNA makes a transition from a straight-supercoiled configuration into a toroid instead of becoming a pure plectoneme, meaning that there could be competition between pure plectonemes and toroids for a region in F vs. n space. Nevertheless, recall that smaller values of α mean larger cooperativity value in Zimm-Bragg Theory (less cooperative transition) and we would expect to see more than one nucleation site in experiment, which does not agree with the observations in [37].
- Increasing the value of the phenomenological parameter C_A and the exponential decay length λ in the DNA-DNA interactions Eq. (3.15) in the main text, makes both collapsed states ($S + P$ and T) become energetically more favorable with respect to the straight configuration. Consequently as C_A and λ increase the value of $F_{critical}$ (transition force between states) increases as function of n .

We conclude that for large number of turns $n > n_{min}$ it is likely that the DNA state corresponds to a supercoiled configuration.

B.3.1 Energy source contributions

In Fig. B.5 we show the relative dominance of each energetic source for the toroidal configuration including twist. Recall that the toroids energy is independent of the applied force F . The major energetic contributions correspond to surface energy and the DNA-DNA interactions. Bending energy is approximately constant as a function of the number of turns, while both the surface and twist energetic contributions increase as a function of the number of turns. In the case of the supercoiled-straight configurations, the energy is a function of both the force and the applied number of turns. For a given force F , the energy increases linearly with n , where l_p (length of the helical region) is the only variable as a function of n . In Fig. B.6 we plot the specific energy contributions in the supercoiled-straight configuration as a function of F .

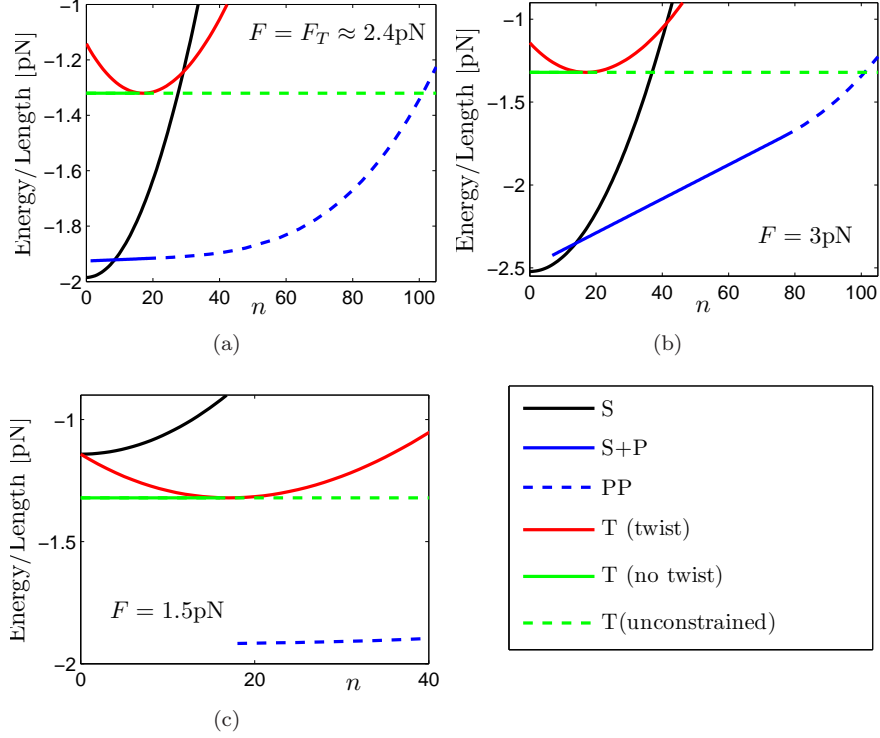


Figure B.3: Energy density of competing structures as a function of the number of turns n for surface tension parameter $\alpha = 0.013$. The parameters used correspond to the theoretical solution for the description of Besteman et al. experiments in spermine in Fig. 2 in the main text. The black solid line corresponds to the straight configuration, the blue solid line to the straight + supercoiled coexistence, the blue dashed line correspond to the pure plectonemic phase, the red solid line correspond to toroids with twist and the green solid line correspond to toroids with no twist. Notice that the solution for toroids with no twist is only valid up to the point where the applied number of turns n is smaller than or equal to the toroid winding number N_g^o . The point where $n = N_g^o$ corresponds to the intersection of the red line and the green solid line. We have also plotted the energy density for torsionally unconstrained toroids depicted by the green dashed line. We point out that the torsionally unconstrained toroids which carry no twist are not a realistic solution for this problem. The link constraint on the DNA molecule must be taken into account otherwise toroids (unconstrained) would become energetically more favorable than supercoiled structures as n increases (green dashed line crosses the blue dashed line). The value of $n_{critical}$ (transition $S \rightarrow S + P$) is given by the intersection of the black solid line and blue solid line, the value of n_{pp} (transition $S + P \rightarrow PP$) is given by the intersection of the blue solid line with the blue dashed line. We have plotted the energy density for three representative forces F : (a) $F = F_T \approx 2.4 \text{ pN}$, supercoils are the favorable state for $n \geq n_{critical}$; (b) $F = 3 \text{ pN} > F_T$, for larger forces the intersection between the black solid line and blue line moves toward larger values of $n_{critical}$; (c) $F = 1.5 \text{ pN} < F_T$, there is no solution for the coexistence of straight + supercoiled, toroids are the preferred state up to n_{pp} which is the point where the PP solution starts.

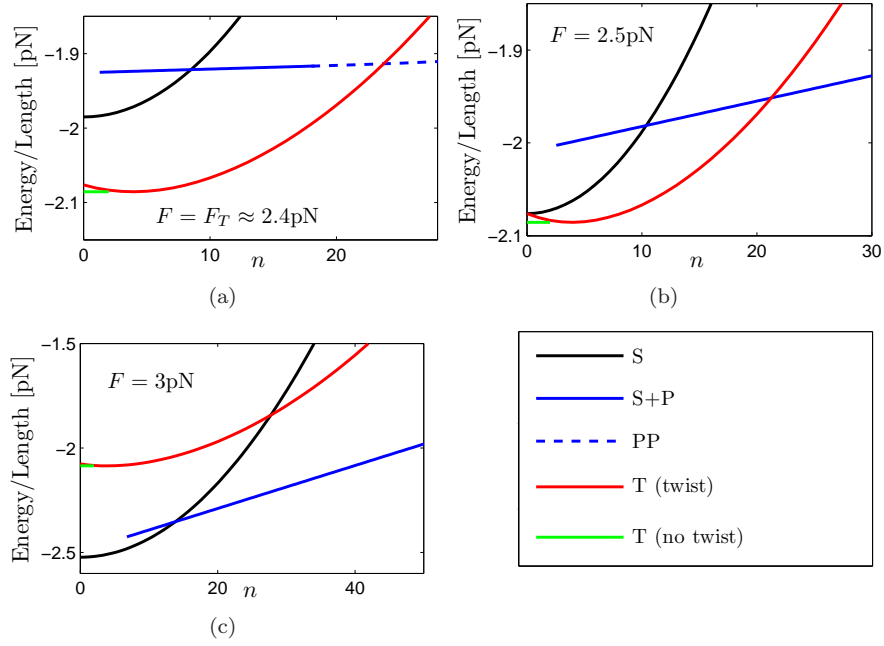


Figure B.4: Energy density of competing structures as a function of the number of turns n for surface tension parameter $\alpha = 0.0003$. The rest of the parameters are the same as in Fig 2 in the main text. Lines represent the same energy densities as in Fig. B.3, but here we do not present the solution for torsionally unconstrained toroids. Changing the value of α changes the preferred state of the DNA molecule as a function of the applied force F . Again we have plot the energy density of the competing structures for three representative forces F : (a) $F = F_T \approx 2.4\text{pN}$, toroids are the preferred state up the intersection point of the blue dashed line and red solid line; (b) as the force increases $F = 2.5\text{pN} < F_T$, both the S and $S + P$ decrease its energy density, and at $n = 0$ the S and $T(\text{twist})$ have the same energy; (c) as the force further increases $F = 3\text{pN} > F_T$, the S configuration becomes the preferred state up to n_{critical} and supercoils are the favorable state for $n \geq n_{\text{critical}}$.

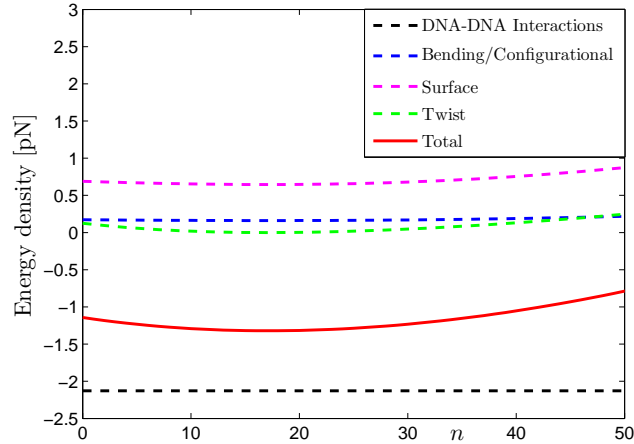


Figure B.5: Energetic contribution in the toroidal configuration including twist. We use the same parameters as in Fig. 2 in the main text.

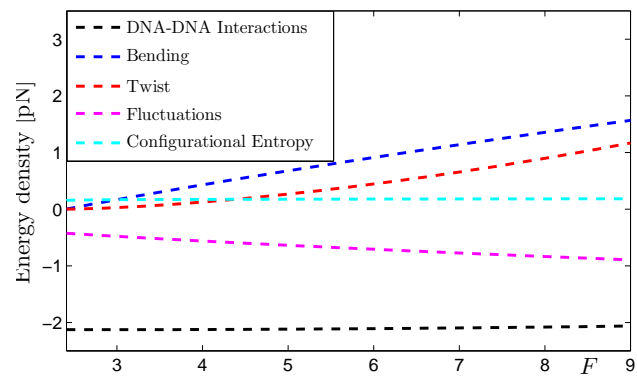


Figure B.6: Energetic contributions in the supercoiled-straight configuration. We use the same parameters as in Fig. 2 in the main text.

Appendix C

Appendix - Chapter 4

C.1 Link, writhe and twist

First we introduce the concept of link Lk (for a detailed description of the topology of link, twist and writhe we refer the reader to [53, 67]). The link is a topological invariant and it can be best seen as the number of turns n put in an initially planar rod before putting its ends together. The link has been shown to consist of two parts [53]:

$$n = Lk = Tw + Wr, \quad (C.1)$$

where the twist Tw is a local quantity in the sense that it can be computed as the single integral of the twist rate and the writhe Wr is a global property of the shape of the rod's center line. Next, we consider the Lk_p in the helices of the plectonemic DNA. The twist is given by:

$$Tw = \frac{1}{2\pi} \int_0^{2l_p} u_3 ds. \quad (C.2)$$

where $K_t u_3 = M_3$ in the helices and $2l_p$ accounts for both helical strands. As shown in [53], Călugăreanu and White introduced the definition of writhe as the following double integral:

$$Wr^{CW}(\gamma) = \frac{1}{4\pi} \int_{\gamma} \int_{\gamma} \frac{[\mathbf{t}(s) \times \mathbf{t}(s^*)] \cdot [\mathbf{r}(s) - \mathbf{r}(s^*)]}{|\mathbf{r}(s) - \mathbf{r}(s^*)|^3} ds ds^*. \quad (C.3)$$

In the above expression $\mathbf{r}(s)$ is the position vector on the curve γ and $\mathbf{t}(s) = \mathbf{d}_3$ is the unit tangent vector. The expression Eq. (C.3) yields the correct value for the writhe Wr of a non-self-intersecting curve. Under the conditions described in [53], Fuller's theorem [113] can be used to compute Wr instead of Eq. (C.3). Fuller's theorem states that the writhe of a curve γ can be computed by considering the writhe of an initial reference curve γ_o and the continuous deformation between both curves:

$$Wr^F(\gamma, \gamma_o) = Wr(\gamma_o) + \frac{1}{2\pi} \int_0^{2l_p} \frac{[\mathbf{t}_{\gamma_o} \times \mathbf{d}_3] \cdot (\mathbf{t}'_{\gamma_o} + \mathbf{d}'_3)}{1 + \mathbf{t}_{\gamma_o} \cdot \mathbf{d}_3} ds, \quad (C.4)$$

where \mathbf{t}_{γ_o} represents the tangent vector of the curve γ_o . By picking γ_o to be a line $\mathbf{t}_{\gamma_o} = \mathbf{e}_3$ then, $Wr(\gamma_o) = 0$ and $\mathbf{t}'_{\gamma_o} = 0$.

As stated in [53], Fuller's formula can not be used directly to compute the writhe of a plectoneme since it does not have a continuous deformation of γ_o into γ that is free of antipodal points. An antipodal point is defined as the point where the direction of \mathbf{d}_3 is opposite to the direction of the tangent vector \mathbf{t}_{γ_o} . Each antipodal event happening along the deformation of γ_o into γ introduces a

shift between $Wr^{CW}(\gamma)$ and $Wr^F(\gamma)$ of $2\pi R$, and it happens once each time as the end rotation $2\pi R$ increases by an amount 2π [53]:

$$2\pi R = \int_0^{l_p} \psi' ds = \psi(l_p) - \psi(0) = \chi \int_0^{l_p} \frac{\sin \theta}{r} ds, \quad (\text{C.5})$$

and the writhe values are related by $Wr^F(\gamma) - Wr^{CW}(\gamma) = N = 2R$

For the analysis in this section we will assume that plectoneme length is very large, $l_p \rightarrow \infty$, and therefore, consider only the helices' contribution to the total writhe. Carrying out the integration in Eq. (C.4) we obtain for helix 1:

$$Wr_1^F = \frac{\chi}{2\pi} \int_0^{l_p} (1 - \cos \theta) \frac{\sin \theta}{r} ds. \quad (\text{C.6})$$

For helix 2, we can create a mirror image using the same coordinate system by a rotation $-\pi$ around the helical axis \mathbf{e}_3 , then the writhe integral is the same as for helix 1. Adding $Wr_1^F + Wr_2^F - N = Wr^{CW}$, we obtain the total writhe of the plectoneme:

$$Wr^{CW} = -\frac{\chi}{\pi} \int_0^{l_p} \frac{\cos \theta \sin \theta}{r} = -\frac{\chi}{\pi} \int_0^{l_p} \frac{\sin 2\theta}{2r} ds. \quad (\text{C.7})$$

Since the problem is symmetric we only need to consider one helix of the plectoneme. The writhe of a single helix is half of the above expression. In the case of θ being a constant, the expression above reduces to the writhe formula used in [45, 73] and [71, 72]. From the symmetry argument we consider only the region $0 < s < l_p$ and recall that $K_t u_3 = M_3$ is a constant, then the final expression for Lk_p is given by:

$$Lk_p = \frac{M_3 l_p}{2\pi K_t} - \chi \int_0^{l_p} \frac{\sin 2\theta}{4\pi r} ds. \quad (\text{C.8})$$

C.2 The Constrained Rod and Variable Curvature

Here we present a brief summary of the model described in detail in van der Heijden [76]. We consider a weightless elastic rod lying on a cylinder and held by an end tension which need not be coaxial with the body axis of the cylinder and a twisting moment applied in the direction of the axis of the cylinder. We assume no friction such that the rod is free to slide along the surface of the cylinder. The unstressed initial configuration corresponds to the rod's center-line, initially straight and without net twist applied. The unit vectors in the orthonormal lab reference frame are \mathbf{e}_i ($i = 1, 2, 3$). This description of the elastic rod corresponds to the purely mechanical problem of one of the two helices in a ply.

The Kirchhoff theory of rods models the center-line as a curve in space $\mathbf{r}(s)$ endowed with mechanical properties which are assumed to be suitable averages over the cross-section of the rod. The configuration of an inextensible, unshereable rod is defined by $\mathbf{r}(s)$ and an associated right-handed orthonormal director frame $\mathbf{d}_i(s)$, $i = 1, 2, 3$, where s is the arc-length. For convenience the vector \mathbf{d}_3 is taken to be tangential to the rod:

$$\mathbf{r}'(s) = \mathbf{d}_3, \quad (\text{C.9})$$

where $()'$ denotes derivative with respect to s . The unshereability and inextensibility are manifested in the constraint given in the center-line equation Eq. (C.9) (note $\mathbf{r}'(s)$ is a unit vector in all configurations). The kinematics of the frame are encapsulated in the director frame equations:

$$\mathbf{d}_i' = \mathbf{u} \times \mathbf{d}_i, \quad (\text{C.10})$$

where the components of $\mathbf{u} = u_i \mathbf{d}_i$ are measures of the strain; u_3 describes the physical twist; u_1 and u_2 are associated with bending such that the square of curvature is given by $\kappa^2 = u_1^2 + u_2^2$. Note that, in general, the physical twist u_3 is different from the geometrical torsion τ [75].

The stresses acting across the cross-section of the rod can be averaged to yield a resultant force $\mathbf{n}(s)$ and moment $\mathbf{m}(s)$. Considering the static case, the balance of linear and angular momentum of an infinitesimal element [75] yield ¹:

$$\mathbf{n}'(s) = \mathbf{f}(s), \quad \mathbf{m}'(s) = \mathbf{n} \times \mathbf{d}_3 + \mathbf{l}(s), \quad (\text{C.11})$$

where $-\mathbf{f}(s)$ is the body force per unit length and $-\mathbf{l}(s)$ is the body couple per unit length. Different functional forms of the body loading can be used to model different effects such as gravity, electrostatics or self-contact. We assume a linear constitutive relation between the stresses and the strains:

$$\mathbf{m} = K_b u_1 \mathbf{d}_1 + K_b u_2 \mathbf{d}_2 + K_t u_3 \mathbf{d}_3. \quad (\text{C.12})$$

Due to the symmetry of the problem, it is convenient to introduce cylindrical coordinates (r, ψ, z) for the position vector:

$$\mathbf{r} = r \cos \psi \mathbf{e}_1 + r \sin \psi \mathbf{e}_2 + z \mathbf{e}_3. \quad (\text{C.13})$$

where $\mathbf{e}_3 = \mathbf{e}_z$ is the axis of the helix that wraps around the cylinder. In particular, as shown by [76], if there are two Euler angles θ and ϕ that describe the director frame \mathbf{d}_i then:

$$\left. \begin{aligned} \mathbf{d}_1 &= \sin \phi \mathbf{e}_r - \cos \phi \cos \theta \mathbf{e}_\psi + \cos \phi \sin \theta \mathbf{e}_z, \\ \mathbf{d}_2 &= \cos \phi \mathbf{e}_r + \sin \phi \cos \theta \mathbf{e}_\psi - \sin \phi \sin \theta \mathbf{e}_z, \\ \mathbf{d}_3 &= \sin \theta \mathbf{e}_\psi + \cos \theta \mathbf{e}_z. \end{aligned} \right\} \quad (\text{C.14})$$

The angle θ measures the deviation from the straight configuration. When the rod deforms into a cylindrical helix, θ is commonly called the helical angle. We will adopt this terminology for all space curves that have a helical shape. The complement $\pi/2 - \theta$ of the helical angle is often referred to as the pitch angle. The angle ϕ measures the internal twist of the rod. The components in cylindrical coordinates are given as: $d_{ir} = \mathbf{d}_i \cdot \mathbf{e}_r$, $d_{i\psi} = \mathbf{d}_i \cdot \mathbf{e}_\psi$ and $d_{iz} = \mathbf{d}_i \cdot \mathbf{e}_z$.

Since $\mathbf{r}(s) = r \mathbf{e}_r + z \mathbf{e}_z$ and $\mathbf{e}'_r = \psi'(s) \mathbf{e}_\psi$, it is clear from the center-line equation Eq. (C.9) and the above expression for \mathbf{d}_3 , that the configuration of the rod is determined by:

$$\psi'(s) = \frac{\sin \theta}{r}, \quad z'(s) = \cos \theta. \quad (\text{C.15})$$

We consider the case of zero distributed body couples ($\mathbf{l} = 0$) and only distributed body forces as done in van der Heijden [76]:

$$\mathbf{n}' = \mathbf{f} = f(s) \mathbf{e}_r, \quad (\text{C.16})$$

where \mathbf{f} is the external reaction force that constrains the rod to lie on a cylinder. For circular cross-section rods, van der Heijden [76] shows that there are *three* constants in the problem. One of these constants is the torque $M_{ext} = M_3$ about the body axis \mathbf{d}_3 :

$$\mathbf{m} \cdot \mathbf{d}_3 = K_t u_3 = M_3. \quad (\text{C.17})$$

¹The derivatives of all quantities are with respect to reference frame \mathbf{e}_i .

C.3 Convexity of κ^2 as a function of θ

Recall that a positive M_3 results in a left-handed helix and a negative M_3 results in a right-handed helix such that $\frac{\chi M_3}{|M_3|} = -1$. Equation Eq. (4.43) can be solved for $(\theta')^2$:

$$(\theta')^2 = a \sin^4 \theta - b \sin 2\theta + cU(r, \theta) + d, \quad (\text{C.18})$$

where the values K_b , χM_{ext} , C_1 and r have been lumped into the constant coefficients $a, b, c \in \mathbb{R}^+$ and $d \in \mathbb{R}$. The functions $U(r, \theta)$ used are convex and consequently:

$$\frac{d^2(\theta')^2}{d\theta^2} > 0 \quad \text{for all } \theta \in J : 0 < \theta \leq \frac{\pi}{4}. \quad (\text{C.19})$$

Furthermore, from Eq. (4.4) it is clear that:

$$\frac{d^2\kappa^2}{d\theta^2} > 0 \quad \text{for all } \theta \in J : 0 < \theta \leq \frac{\pi}{4}. \quad (\text{C.20})$$

It can be easily shown that both, κ^2 and θ'^2 , as functions of θ have a U shape in the interval J .

C.4 Variable end loop length

As mentioned in the main text, as more turns are added to the DNA molecule in figure 4.2 the curvature at point P_o is becoming tighter and tighter, such that the end loop is becoming smaller and smaller. The portion that the end loop is taking from the filament is changing, and consequently s_o is also varying. So, we say that P_o and P_1 are, in general, variable end points. Here let $\Omega(s)$ be the curve where P_o lies and let $\zeta(s)$ be the curve where P_1 lies. In fact $\Omega(s)$ corresponds to the curve that describes the Euler angle θ in the loop solution for $s \leq s_o$. The transversality conditions for variable end points are [127]:

$$\Omega' \frac{\partial I}{\partial \theta'} - \left(\theta' \frac{\partial I}{\partial \theta'} - I \right) \Big|_{s=s_o} = K_b \Omega' \theta' - (C_1 - F + G_{flu}^*) \Big|_{s=s_o} = 0, \quad (\text{C.21})$$

$$\zeta' \frac{\partial I}{\partial \theta'} - \left(\theta' \frac{\partial I}{\partial \theta'} - I \right) \Big|_{s=s_1} = K_b \zeta' \theta' - (C_1 - F + G_{flu}^*) \Big|_{s=s_1} = 0, \quad (\text{C.22})$$

where we have made use of Eq. (4.38) and Eq. (4.43). The subscripts in the above expression denote partial differentiation with respect to that variable.

The curve describing the geometry of the plectonemes has to be continuous and analytic at every point along the arc length s . From the arguments in the main text in section 4.4.5 we obtained:

$$C_1(s_1) = K_b \zeta(s_1)^{\prime 2} + F - G_{flu}^*. \quad (\text{C.23})$$

Next we enforce the continuity of the tangent vector $\mathbf{d}_3 = \mathbf{r}'(s)$ and the moment vector \mathbf{m} at the intersection point s_o . Continuity of $\mathbf{d}_3 = \mathbf{r}'(s)$ implies $\theta(s)$ to be continuous at $s = s_o$ while continuity of the moment implies $\mathbf{m} \cdot \mathbf{e}_r = m_r = -K_b \theta'(s) = -K_b \Omega'(s)$ to be continuous at $s = s_o$. The latter means the first boundary condition in equation Eq. (C.21) can be rewritten as:

$$C_1(s_o) = K_b \Omega(s_o)^{\prime 2} + F - G_{flu}^*. \quad (\text{C.24})$$

Since $C_1(s_1) = C(s_o)$ is a constant independent of the arc-length s , we conclude that $\Omega'^2 \Big|_{s=s_o} = \zeta'^2 \Big|_{s=s_1}$

The same procedure from section 4.4.6 can be carried out replacing l_p for s_1 and s_o as independent variables. The total length of the plectonemic region is $l_p = s_1 - s_o$. The phase transition condition Eq. (4.54) in the main text would become:

$$\int_{s_o^i(n_i)}^{s_1^i(n_i)} [\theta^{*'}(s; n_i)]^2 - [\Theta(n_i)]^2 ds = \int_{s_o(\hat{n})}^{s_1(\hat{n})} [\theta^{*'}(s; \hat{n})]^2 - [\Theta(\hat{n})]^2 ds, \quad (\text{C.25})$$

where $[\Theta(\hat{n})]^2 = \Omega'^2 \Big|_{s=s_o}^{n=\hat{n}} = \zeta'^2 \Big|_{s=s_1}^{n=\hat{n}}$. As in main text in section 4.4.6 where the end loop length is fixed, if M_3 and r are independent of n , we recover as the only possible solution the uniform helix where $C_1 = F - G_{flu}^*$ and $\theta' = 0$ for all values of s , since this is the case that satisfies the boundary conditions together with the phase condition Eq. (C.25).

Appendix D

Appendix - Chapter 5

D.1 Force-extension and torque-link relations in a F, M_3 ensemble.

D.2 Homogeneous chain

For the homogeneous chain, where there is only one set of mechanical properties, it can be shown [49] that:

$$\det \mathbf{A} = \frac{r^{N-1} - d^{N-1}}{r - d}, \quad (\text{D.1})$$

where we have used the following definitions:

$$\kappa = \frac{AN}{2L}, \quad \frac{f}{\kappa} = \frac{FL}{AN^2}, \quad q = f + M_3^2/(16\kappa) \quad (\text{D.2})$$

$$r = \frac{(2\kappa + q) + \sqrt{4\kappa q + q^2}}{2}, \quad d = \frac{(2\kappa + q) - \sqrt{4\kappa q + q^2}}{2}, \quad (\text{D.3})$$

It is clear that from the definitions in Eq. (D.2) that:

$$\frac{\partial(\)}{\partial F} = \frac{L}{2N} \frac{\partial(\)}{\partial q}, \quad (\text{D.4})$$

and from here on in this section we will use notation $(\)' = \partial(\)/\partial F$. So taking derivatives with respect to the controlled variable F we get:

$$\frac{\partial \log [\det \mathbf{A}]}{\partial F} = (N-1) \frac{r'}{r} \cdot \frac{1 - \frac{d'}{r'} \cdot \left(\frac{d}{r}\right)^{N-2}}{1 - \left(\frac{d}{r}\right)^{N-1}} - \frac{r'}{r} \cdot \frac{1 - d'/r'}{1 - d/r} \quad (\text{D.5})$$

$$r' = \frac{L}{2N} \left[\frac{(2\kappa + q) + \sqrt{4\kappa q + q^2}}{2\sqrt{4\kappa q + q^2}} \right], \quad d' = \frac{L}{2N} \left[\frac{-(2\kappa + q) + \sqrt{4\kappa q + q^2}}{2\sqrt{4\kappa q + q^2}} \right], \quad (\text{D.6})$$

We take the limit for $N \rightarrow \infty$ ($L = Nl$ is fixed such that as $N \rightarrow \infty$ then $l \rightarrow 0$) to recover the response function of the chain in the continuous limit:

$$\frac{\partial \log(\det \mathbf{A})}{\partial F} = \frac{1}{2} \frac{L}{Q} \coth \left[L \frac{Q}{A} \right] - \frac{1}{2} \frac{A}{Q^2}, \quad (\text{D.7})$$

where

$$Q = \sqrt{AF - \frac{M_3^2}{4}}. \quad (\text{D.8})$$

For the entire chain (consisting of two single chains with v_x and v_y kinematic variables as specified in equation Eq. (5.23) in the main text), the fluctuating modes are doubled (see [49]), and therefore the expression for the extension is given by:

$$\langle x \rangle = L - k_B T \frac{\partial \log(\det \mathbf{A})}{\partial F} = L - \frac{k_B T}{2} \left(\frac{L}{Q} \coth \left[L \frac{Q}{A} \right] - \frac{A}{Q^2} \right). \quad (\text{D.9})$$

In a similar way the linking number vs. applied torque relation for a homogeneous chain is given by:

$$\begin{aligned} \langle 2\pi n \rangle &= \sum_{i=1}^N \left[\frac{M_3 l}{C_i} + u_i^\circ l \right] - k_B T \frac{\partial \log(\det \mathbf{A})}{\partial M_3} \\ &= \frac{M_3 L}{C} + M_3 \frac{k_B T}{4A} \left(\frac{L}{Q} \coth \left[L \frac{Q}{A} \right] - \frac{A}{Q^2} \right) + u^\circ L. \end{aligned} \quad (\text{D.10})$$

D.2.1 Special heterogeneous chain: 2-phase model

In this section we follow the methods described in [49], but for a heterogeneous chain subjected to partially clamped boundary conditions. Here we model the chain with only one phase boundary. This approximation is valid when the interface energy required to create a boundary is high, as it is the case of highly cooperative transitions [182]. Hence, for the two phase chain we have:

$$\kappa_i = \begin{cases} \kappa_1, & 1 \leq i \leq s \\ \kappa_2, & s+1 \leq i \leq N \end{cases} \quad (\text{D.11})$$

Here the subscripts represent the variables corresponding to each phase, hence l_1 is the length of s links in phase 1, A_1 is the bending modulus in phase 1, l_2 of is the length of $N - s$ links in phase 2 and A_2 is the bending modulus in phase 2. Hence the total length (which is fixed and not a function of the controlled parameters F and M_3) is given by

$$L = sl_1 + (N - s)l_2 = L_1 + L_2, \quad (\text{D.12})$$

and we have the following relations:

$$f_j = \frac{Fl_j}{2}, \quad q_j = f_j + \frac{M_3^2}{16\kappa_j} = f_j + \frac{M_3^2 l_j}{8A_j}, \quad \text{for } j=1,2 \quad (\text{D.13})$$

$$\kappa_1 = \frac{A_1}{2l_1}, \quad Q_1 = \sqrt{A_1 F - \frac{M_3^2}{4}}, \quad \kappa_2 = \frac{A_2}{2l_2}, \quad Q_2 = \sqrt{A_2 F - \frac{M_3^2}{4}}. \quad (\text{D.14})$$

$$\lambda_i = \begin{cases} 2\kappa_1 + q_1 = r_1 + d_1, & i = 1 \\ 2\kappa_1 + q_1 - \frac{r_1 d_1}{\lambda_{i-1}} = r_1 + d_1 - \frac{r_1 d_1}{\lambda_{i-1}}, & 2 \leq i \leq s-1 \\ \kappa_1 + \kappa_2 + q_2 - \frac{r_1 d_1}{\lambda_{i-1}} = r_2 + d_2 + \kappa_1 - \kappa_2 - \frac{r_1 d_1}{\lambda_{i-1}}, & i = s \\ 2\kappa_2 + q_2 - \frac{r_2 d_2}{\lambda_{i-1}} = r_2 + d_2 - \frac{r_2 d_2}{\lambda_{i-1}}, & s+1 \leq i \leq N-2 \end{cases} \quad (\text{D.15})$$

$$\det \mathbf{A} = \prod_{k=1}^{N-2} \lambda_i = \prod_{k=s+1}^{N-2} \lambda_i \cdot \lambda_s \cdot \prod_{k=1}^{s-1} \lambda_i = \prod_{k=s}^{N-2} \lambda_i \cdot \lambda_{s-1} \cdot \prod_{k=1}^{s-2} \lambda_i \quad (\text{D.16})$$

After some manipulation we arrive at the final expression for $\det \mathbf{A}$ for the 2-phase heterogeneous chain subjected to partially clamped boundary conditions which is given by:

$$\det \mathbf{A} = \frac{r_1^s - d_1^s}{(r_1 - d_1)(r_2 - d_2)} [r_2^{N-s} - d_2^{N-s} + (\kappa_1 - \kappa_2)(r_2^{N-s-1} - d_2^{N-s-1})] \quad (\text{D.17})$$

$$- r_1 d_1 \frac{r_1^{s-1} - d_1^{s-1}}{r_1 - d_1} \frac{r_2^{N-s-1} - d_2^{N-s-1}}{r_2 - d_2} \quad (\text{D.18})$$

Next we compute the partial derivative with respect to the applied force F and take the limit as l_1 and $l_2 \rightarrow 0$ while $(N-s)$ and $s \rightarrow \infty$:

$$\frac{\partial \log(\det \mathbf{A})}{\partial F} = \frac{a' + b'}{a + b} - c' - d'. \quad (\text{D.19})$$

In this section of the Appendix ($'$) denotes derivative with respect to F .

$$c' = \frac{r_1' - d_1'}{r_1 - d_1} = \frac{1}{2} \frac{A_1}{Q_1^2}, \quad (\text{D.20})$$

$$d' = \frac{r_2' - d_2'}{r_2 - d_2} = \frac{1}{2} \frac{A_2}{Q_2^2}, \quad (\text{D.21})$$

$$\begin{aligned} \frac{a' + b'}{a + b} &= \frac{\left[\left(L_1 \frac{Q_2}{Q_1} + L_2 \frac{Q_1}{Q_2} \right) \cosh(h_1) \cosh(h_2) + (L_1 + L_2) \sinh(h_1) \sinh(h_2) \right]}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]} \\ &+ \frac{\left(\frac{A_1}{Q_1} \right) \sinh(h_2) \cosh(h_1) + \left(\frac{A_2}{Q_2} \right) \sinh(h_1) \cosh(h_2)}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]}, \end{aligned} \quad (\text{D.22})$$

or using trigonometric identities

$$\begin{aligned} \frac{a' + b'}{a + b} &= \frac{\left[\left(L_1 \frac{Q_2}{Q_1} + L_2 \frac{Q_1}{Q_2} \right) \{ \cosh(h_1 + h_2) + \cosh(h_2 - h_1) \} \right]}{2 [Q_1 \{ \sinh(h_2 + h_1) + \sinh(h_2 - h_1) \} + Q_2 \{ \sinh(h_2 + h_1) - \sinh(h_2 - h_1) \}]} \\ &+ \frac{[(L_1 + L_2) \{ \cosh(h_1 + h_2) - \cosh(h_2 - h_1) \}]}{2 [Q_1 \{ \sinh(h_2 + h_1) + \sinh(h_2 - h_1) \} + Q_2 \{ \sinh(h_2 + h_1) - \sinh(h_2 - h_1) \}]} \\ &+ \frac{\left(\frac{A_1}{Q_1} \right) \{ \sinh(h_2 + h_1) + \sinh(h_2 - h_1) \} + \left(\frac{A_2}{Q_2} \right) \{ \sinh(h_2 + h_1) - \sinh(h_2 - h_1) \}}{2 [Q_1 \{ \sinh(h_2 + h_1) + \sinh(h_2 - h_1) \} + Q_2 \{ \sinh(h_2 + h_1) - \sinh(h_2 - h_1) \}]} \end{aligned}$$

where

$$h_1 = \frac{L_1 Q_1}{A_1}, \quad h_2 = \frac{L_2 Q_2}{A_2}. \quad (\text{D.23})$$

Hence the force-extension relation for the heterogeneous chain as obtained from plugging in Eq. (D.20) and Eq. (D.21) into Eq. (D.19) is:

$$\langle x \rangle = L - k_B T \frac{\partial \log(\det \mathbf{A})}{\partial F} = L - k_B T \left(\frac{a' + b'}{a + b} - \frac{1}{2} \left[\frac{A_1}{Q_1^2} + \frac{A_2}{Q_2^2} \right] \right). \quad (\text{D.24})$$

In order to obtain the linking number vs. applied torque relation for the chain note that since

$$q_1 = \frac{l_1}{2}F - \frac{M_3^2 l_1}{8A_1} \text{ and } q_2 = \frac{l_2}{2}F - \frac{M_3^2 l_2}{8A_2}, \quad (\text{D.25})$$

then we must have:

$$\frac{\partial \Gamma_1(q_1)}{\partial M_3} = \frac{\partial \Gamma(q_1)}{\partial q_1} \frac{\partial q_1}{\partial M_3} = \frac{\partial \Gamma(q_1)}{\partial F} \frac{\partial F}{\partial q_1} \frac{\partial q_1}{\partial M_3} = -\frac{M_3}{2A_1} \frac{\partial \Gamma_1}{\partial F}, \quad (\text{D.26})$$

and similarly

$$\frac{\partial \Gamma_2(q_2)}{\partial M_3} = -\frac{M_3}{2A_2} \frac{\partial \Gamma_2}{\partial F}. \quad (\text{D.27})$$

Therefore,

$$\langle 2\pi n \rangle = M_3 \frac{L_1}{C_1} + M_3 \frac{L_2}{C_2} - k_B T \frac{\partial \log(\det \mathbf{A})}{\partial M_3} \quad (\text{D.28})$$

where

$$\frac{\partial \log(\det \mathbf{A})}{\partial M_3} = \left(\frac{\partial a}{\partial M_3} + \frac{\partial b}{\partial M_3} \right) \frac{1}{a+b} - \frac{\partial c}{\partial M_3} - \frac{\partial d}{\partial M_3} \quad (\text{D.29})$$

$$\frac{\partial c}{\partial M_3} = -\frac{M_3}{4Q_1^2}, \quad \frac{\partial d}{\partial M_3} = -\frac{M_3}{4Q_2^2}, \quad (\text{D.30})$$

$$\begin{aligned} & \frac{1}{a+b} \left(\frac{\partial a}{\partial M_3} + \frac{\partial b}{\partial M_3} \right) = \\ - & \left(\frac{M_3}{4} \right) \frac{\left[\left(\frac{L_1 Q_2}{A_1 Q_1} + \frac{L_2 Q_1}{A_2 Q_2} \right) \cosh(h_1) \cosh(h_2) + \left(\frac{L_1}{A_1} + \frac{L_2}{A_2} \right) \sinh(h_1) \sinh(h_2) \right]}{[Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]} \\ - & \left(\frac{M_3}{4} \right) \frac{(Q_1^{-1}) \sinh(h_2) \cosh(h_1) + (Q_2^{-1}) \sinh(h_1) \cosh(h_2)}{[Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]} \end{aligned} \quad (\text{D.31})$$

Taking the limit as $L_2 \rightarrow 0$ ($L = L_1$) or the limit $A_1 \rightarrow A_2$ and $C_1 \rightarrow C_2$ ($L = L_1 + L_2$) in the expressions $\langle x \rangle$ and $\langle n \rangle$ we recover the homogeneous solutions (equations Eq. (D.9) and Eq. (D.10)).

D.3 Stretch, twist-stretch couple and the (F, n) -ensemble

In this section we present the force-extension and torque-link relations for a heterogeneous (2-phase) chain where the controlled variables are the force F and the linking number $2\pi n$. Here we have included the effects of stretching and the coupling of twist-stretch. As before A is the bending modulus and C is the twist modulus, while in this section S represents the stretch modulus and g the twist-stretch coupling modulus.

D.3.1 Energetics of the chain with stretch and twist-stretch coupling (Cosserat theory of rods):

In the Kirchhoff-Cosserat theory of rods the curvature measures of strain are :

$$\mathbf{u} = k_1 \mathbf{d}_1 + k_2 \mathbf{d}_2 + u_3 \mathbf{d}_3, \quad (\text{D.32})$$

where \mathbf{d}_i is a director frame in the deformed configuration. For convenience we let \mathbf{d}_3 to be parallel to the tangent vector and the transformation between \mathbf{d}_i and the reference frame \mathbf{e}_i is given by:

$$[\mathbf{d}_i] = \mathbf{R}_\phi \mathbf{R}_\theta \mathbf{R}_\psi [\mathbf{e}_i], \quad (\text{D.33})$$

$$\mathbf{R}_\psi = \begin{bmatrix} \sin(\phi) & -\cos(\phi) & 0 \\ \cos(\phi) & \sin(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R}_\theta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix},$$

$$\mathbf{R}_\psi = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where ψ is a rotation about the \mathbf{e}_3 axis, $\theta \in [0, \pi]$ is the angle describing the direction of \mathbf{d}_3 with respect to \mathbf{e}_3 and ϕ is the internal rotation of the filament about \mathbf{d}_3 .

In addition to the \mathbf{u} strains, we now include the possibility of stretch along the rod. The reference undeformed arc length is given by s , while the current deformed configuration's arc length is denoted \hat{s} . The stretch in the rod is hence given by

$$\frac{d\hat{s}}{ds} = (1 + \epsilon), \quad (\text{D.34})$$

such that the tangent vector in the deformed configuration is given by:

$$\mathbf{r}'(s) = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{d\hat{s}} \frac{d\hat{s}}{ds} = (1 + \epsilon) \mathbf{d}_3. \quad (\text{D.35})$$

We now use the more general expression for the stored energy density function given up to quadratic order:

$$\Lambda(\kappa, u_3, \epsilon) = \frac{1}{2} [A(k_1^2 + k_2^2) + Cu_3^2 + 2gu_3\epsilon + S\epsilon^2], \quad (\text{D.36})$$

where $k_1^2 + k_2^2 = \kappa^2$. We assume the rod to be hyperelastic such that the stress is given by:

$$\mathbf{m} = \frac{\partial \Lambda}{\partial \mathbf{u}} = M_i \mathbf{d}_i = Ak_1 \mathbf{d}_1 + Ak_2 \mathbf{d}_2 + M_3 \mathbf{d}_3, \quad (\text{D.37})$$

where $M_3 = \lambda = (Cu_3 + g\epsilon)$ is the external torque about the filament axis for the zero temperature chain. We also take the energetic contribution of the force to be given up to quadratic order in the kinematic variables:

$$F\Delta x = F \int_0^L [1 - \eta_3 \cos \theta] ds \approx F \int_0^L \left[1 - (1 + \epsilon) \left(1 - \frac{\theta^2}{2} \right) \right] ds \approx -F \int_0^L \left[\frac{\theta^2}{2} + \epsilon \right] ds. \quad (\text{D.38})$$

Hence the potential energy of the system is given by:

$$E = \int_0^L \Lambda ds + 2\pi\lambda(n - \Delta T w - W r - \Delta T w_o), \quad (\text{D.39})$$

where the expressions for the excess linking number $n = \Delta L k$ are given in the deformed configuration \hat{s} , but they can be transformed using a change of coordinates to the reference configuration s , such that:

$$2\pi n = \int_0^L [u_3 + u^o + \psi' (1 + \cos \theta)] ds. \quad (\text{D.40})$$

The energy using the small angle approximation is therefore given by (after completing the square

in the variables ψ', u_3 and ϵ):

$$E = \int_0^L \hat{I}_\epsilon ds + 2\pi n\lambda, \quad (\text{D.41})$$

$$\hat{I}_\epsilon = \frac{A}{2} \left[\theta'^2 + \theta^2 \left(\psi' - \frac{\lambda}{2A} \right)^2 \right] + \left[\frac{F}{2} - \frac{\lambda^2}{8A} \right] \theta^2 + 2z, \quad (\text{D.42})$$

$$2z = -\frac{1}{2\tilde{C}} \left(\frac{gF}{S} - \lambda \right)^2 - \frac{F^2}{2S} + \frac{S}{2} (\delta\epsilon)^2 + \frac{\tilde{C}}{2} (\delta u_3)^2 + u^o \lambda. \quad (\text{D.43})$$

In the above expression \tilde{C} is an effective twist modulus

$$\tilde{C} = C - \frac{g^2}{S}. \quad (\text{D.44})$$

Note $\tilde{C}(s)$, $S(s)$, $u_3^{\min}(s)$, $A(s)$, $\epsilon^{\min}(s)$ are not constants for a heterogeneous rod.

Let:

$$\Psi' = \psi' - \frac{\lambda}{2K_b}, \quad (\text{D.45})$$

$$v_x = \theta \sin(\Psi), \quad v_y = \theta \cos(\Psi), \quad (\text{D.46})$$

such that

$$v_x^2 + v_y^2 = \theta^2, \quad (\text{D.47})$$

$$v'_x = \theta' \sin \Psi + \theta \Psi' \cos \Psi, \quad v'_y = \theta' \cos \Psi - \theta \Psi' \sin \Psi, \quad (\text{D.48})$$

$$\kappa^2 = v_x'^2 + v_y'^2 = \theta'^2 + \theta^2 \Psi'^2 = \theta'^2 + \theta^2 \left(\psi' - \frac{\lambda}{2K_b} \right)^2. \quad (\text{D.49})$$

As in the main text the energy of the system is decomposed in two identical rods:

$$\begin{aligned} E &= \int_0^L \left[\frac{A}{2} v_x'^2 + \left(F - \frac{\lambda^2}{4A} \right) \frac{v_x^2}{2} + z \right] ds + \pi n\lambda \\ &+ \int_0^L \left[\frac{A}{2} v_y'^2 + \left(F - \frac{\lambda^2}{4A} \right) \frac{v_y^2}{2} + z \right] ds + \pi n\lambda \end{aligned} \quad (\text{D.50})$$

The discrete version of the energy and link for a single rod are given by:

$$E_F \approx \sum_{i=1}^{N-1} \kappa_i (v_{i+1} - v_i)^2 + \sum_{i=1}^N v_i^2 q_i + \sum_{i=1}^N z_i + \pi n\lambda, \quad (\text{D.51})$$

$$2\pi n = \int_0^L u_3^{\min} + u^o ds \approx \sum [u_{3i}^{\min} + u^o] l = \lambda \sum \frac{l}{C_i} - F \sum \frac{g_i l}{S_i \tilde{C}_i} + \sum u_i^o l, \quad (\text{D.52})$$

where

$$u_{3i}^{\min} \tilde{C}_i + \frac{g_i F}{S_i} = \lambda = \frac{(2\pi n + F\Phi - U^o)}{\varpi}, \quad (\text{D.53})$$

$$\Phi = \sum \frac{g_i l}{S_i \tilde{C}_i}, \quad \varpi = \sum \frac{l}{C_i}, \quad U^o = \sum u_i^o l, \quad (\text{D.54})$$

$$f = \frac{Fl}{2}, \quad \kappa_i = \frac{A_i}{2l}, \quad q_i = f - \frac{\lambda^2}{16\kappa_i}, \quad (\text{D.55})$$

$$2z_i = \left(-\frac{1}{2\tilde{C}_i} \left(\frac{g_i F}{S_i} - \lambda \right)^2 - \frac{F^2}{2S_i} \right) l + 2s_i \Delta \epsilon_i + 2\tilde{c}_i \Delta_i + u_i^o \lambda, \quad (\text{D.56})$$

$$\tilde{c}_i = \frac{\tilde{C}_i}{4l}, \quad \Delta_i = \delta_i u_3 l, \quad s_i = \frac{S_i l}{4}, \quad \Delta \epsilon_i = \delta_i \epsilon \quad (\text{D.57})$$

As in the main text we impose the constraint $v_0 = v_N = 0$ due to the BCs, hence the energy of one rod is:

$$\begin{aligned} \beta E_F &\approx \beta \left[\sum_{i=2}^{N-1} \{(\kappa_{i-1} + \kappa_i + q_i) v_i^2\} - 2 \sum_{i=1}^{N-1} \kappa_i v_i v_{i+1} + \sum_{i=1}^N \tilde{c}_i (\Delta_i)^2 + \sum_{i=1}^N s_i (\Delta \epsilon_i)^2 \right] + \beta \frac{E_{min}}{2} \\ &= \tilde{\Theta} \mathbf{M} \tilde{\Theta}^T + \tilde{\Lambda} \mathbf{T} \tilde{\Lambda}^T + \tilde{\Upsilon} \mathbf{E} \tilde{\Upsilon}^T + \beta \frac{E_{min}}{2}, \end{aligned} \quad (\text{D.58})$$

where $\beta^{-1} = k_B T$, $\tilde{\Theta} = [v_2, v_2, \dots, v_{N-1}]$, $\tilde{\Lambda} = [\Delta_1, \Delta_2, \dots, \Delta_N]$, $\tilde{\Upsilon} = [\Delta \epsilon_1, \Delta \epsilon_2, \dots, \Delta \epsilon_N]$, \mathbf{T} and \mathbf{E} are an N -dimensional diagonal matrices:

$$T_{ij} = \beta c_i \delta_{ij}, \quad E_{ij} = \beta s_i \delta_{ij}, \quad (\text{D.59})$$

and \mathbf{M} is the $N - 1$ -dimensional tridiagonal matrix which can be written in compact form as:

$$[\mathbf{K}]_{ij} = \left\{ \beta (\kappa_{i-1} + \kappa_i + q_i) \delta_{ij} - \beta \kappa_t \cdot \delta_{(|i-1|, 1)} \right\}, \quad (\text{D.60})$$

where $t = \min(i, j)$ and δ is the Kronecker delta and the zero temperature minimum energy is:

$$E_{min} = \sum_{i=1}^N \left\{ -\frac{1}{2\tilde{C}_i} \left(\frac{g_i F}{S_i} - \frac{2\pi n - U^o}{\varpi} - F \frac{\Phi}{\varpi} \right)^2 - \frac{F^2}{2S_i} \right\} l + [2\pi n + U^o] \left[\frac{2\pi n - U^o}{\varpi} + F \frac{\Phi}{\varpi} \right]. \quad (\text{D.61})$$

The partition function is the integral:

$$Z = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp[-\beta E_F] d\tilde{\Theta} d\tilde{\Lambda} d\tilde{\Upsilon}, \quad (\text{D.62})$$

where $d\tilde{\Theta} = [dv_2, dv_2, \dots, dv_{N-1}]$, $d\tilde{\Lambda} = [d\Delta_1, d\Delta_2, \dots, d\Delta_N]$ and $d\tilde{\Upsilon} = [d\epsilon_1, d\epsilon_2, \dots, d\epsilon_N]$, and the free energy is given by:

$$G = -2k_B T \ln Z, \quad (\text{D.63})$$

where the factor 2 accounts for the two identical chains making up the fluctuating rod (see Eq. (D.50)).

D.3.2 General force-extension and link-torque relations for a two phase heterogeneous chain.

Here the subscripts represent the value of the variable for each phase such that:

$$f = \frac{Fl}{2}, \quad q_j = f - \frac{\lambda^2}{16\kappa_j} \quad \text{for } j=1,2, \quad (\text{D.64})$$

where

$$\lambda = \tilde{C}_1 u_{3,1}^{min} + \frac{g_1}{S_1} F = \tilde{C}_2 u_{3,2}^{min} + \frac{g_2}{S_2} F = \left(\frac{2\pi n + F\Phi - U^o}{\varpi} \right), \quad (\text{D.65})$$

$$\varpi = \frac{L_1}{\tilde{C}_1} + \frac{L_2}{\tilde{C}_2}, \quad \Phi = \frac{g_1 L_1}{S_1 \tilde{C}_1} + \frac{g_2 L_2}{S_2 \tilde{C}_2}, \quad U^o = u_1^o L_1 + u_2^o L_2, \quad (\text{D.66})$$

and the effective twist modulus is given by:

$$\tilde{C}_j = C_j - \frac{g_j^2}{S_j} \quad \text{for } j = 1, 2, \quad (\text{D.67})$$

$$\kappa_j = \frac{A_j}{2l_j}, \quad Q_j = \sqrt{A_j F - \frac{1}{4}\lambda^2}. \quad (\text{D.68})$$

The force-extension relation for the heterogeneous chain subjected to tension, twist, stretch and partially clamped boundary conditions is given by:

$$\begin{aligned} \langle x \rangle &= L - \frac{\partial E_{min}}{\partial F} - k_B T \frac{\partial \log(\det \mathbf{A})}{\partial F} \\ &= L - \frac{\partial E_{min}}{\partial F} - k_B T \left(\frac{a' + b'}{a + b} - \frac{1}{2} \left[\nabla_1 \frac{A_1}{Q_1^2} + \nabla_2 \frac{A_2}{Q_2^2} \right] \right). \end{aligned} \quad (\text{D.69})$$

where we have made used of the fact that

$$\frac{\partial \langle x \rangle}{\partial F} = \nabla_j \frac{l_j}{2} \frac{\partial \langle x \rangle}{\partial q_j}, \quad (\text{D.70})$$

and

$$\nabla_j = \left[1 - \frac{\Phi}{2A_j} \left(\frac{2\pi n - U^o + F\Phi}{\varpi^2} \right) \right], \quad (\text{D.71})$$

$$\begin{aligned} -\frac{\partial E_{min}}{\partial F} &= \frac{L_1}{C_1} \left[\left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) F - \frac{2\pi n - U^o}{\varpi} \right] \left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) \\ &+ \frac{L_2}{C_2} \left[\left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) F - \frac{2\pi n - U^o}{\varpi} \right] \left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) \\ &+ \left(F \frac{L_1}{S_1} + F \frac{L_2}{S_2} \right) - 2\pi n \frac{\Phi}{\varpi}, \end{aligned} \quad (\text{D.72})$$

$$\begin{aligned} \frac{a' + b'}{a + b} &= \frac{\left[\left(L_1 \nabla_1 \frac{Q_2}{Q_1} + L_2 \nabla_2 \frac{Q_1}{Q_2} \right) \cosh(h_1) \cosh(h_2) + (L_1 \nabla_1 + L_2 \nabla_2) \sinh(h_1) \sinh(h_2) \right]}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]} \\ &+ \frac{\left(\nabla_1 \frac{A_1}{Q_1} \right) \sinh(h_2) \cosh(h_1) + \left(\nabla_2 \frac{A_2}{Q_2} \right) \sinh(h_1) \cosh(h_2)}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]}. \end{aligned} \quad (\text{D.73})$$

For an ensemble where n is the controlled variable, the average torque is then given by:

$$\langle M_3 \rangle = \frac{1}{2\pi} \left[\frac{\partial E_{min}}{\partial n} + k_B T \frac{\partial \log(\det \mathbf{A})}{\partial n} \right], \quad (\text{D.74})$$

where

$$\frac{1}{2\pi} \frac{\partial E_{min}}{\partial n} = \frac{L_1}{\varpi \tilde{C}_1} \left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) F + \frac{L_2}{\varpi \tilde{C}_2} \left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) F + \frac{2\pi n}{\varpi} + F \frac{\Phi}{\varpi} - \frac{U^o}{\varpi}, \quad (\text{D.75})$$

$$\frac{\partial \log(\det \mathbf{A})}{\partial n} = \left(\frac{\partial a}{\partial n} + \frac{\partial b}{\partial n} \right) \frac{1}{a + b} - \frac{\partial c}{\partial n} - \frac{\partial d}{\partial n}, \quad (\text{D.76})$$

$$\frac{1}{2\pi} \frac{\partial c}{\partial n} = -\tilde{\Lambda}_1 \frac{A_1}{2Q_1^2}, \quad \frac{1}{2\pi} \frac{\partial d}{\partial n} = -\tilde{\Lambda}_2 \frac{A_2}{2Q_2^2}, \quad (\text{D.77})$$

$$\begin{aligned}
& \frac{1}{2\pi} \left[\left(\frac{\partial a}{\partial Lk} + \frac{\partial b}{\partial Lk} \right) \frac{1}{a+b} \right] = \\
& \frac{\left[\left(\tilde{\Lambda}_1 \frac{L_1 Q_2}{Q_1} + \tilde{\Lambda}_2 \frac{L_2 Q_1}{Q_2} \right) \cosh(h_1) \cosh(h_2) + \left(\tilde{\Lambda}_1 L_1 + \tilde{\Lambda}_2 L_2 \right) \sinh(h_1) \sinh(h_2) \right]}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]}, \\
& \frac{\left(\frac{\tilde{\Lambda}_1 A_1}{Q_1} \right) \sinh(h_2) \cosh(h_1) + \left(\frac{\tilde{\Lambda}_2 A_2}{Q_2} \right) \sinh(h_1) \cosh(h_2)}{2 [Q_1 \sinh(h_2) \cosh(h_1) + Q_2 \sinh(h_1) \cosh(h_2)]}, \tag{D.78}
\end{aligned}$$

where we have made use of the fact that:

$$\frac{1}{2\pi} \frac{\partial(\cdot)}{\partial n} = \tilde{\Lambda}_j \frac{l_j}{2} \frac{\partial(\cdot)}{\partial q_j}, \tag{D.79}$$

and

$$\tilde{\Lambda}_j = \frac{1}{2A_j} \left(\frac{2\pi n + F\Phi - U^o}{\varpi^2} \right) \quad \text{for } j=1,2. \tag{D.80}$$

D.3.3 Results for Specific cases

Homogeneous Chain

For a single phase and setting $U^o = 0$:

$$Q_n = \sqrt{AF - \frac{1}{4} \cdot \left(\frac{2\pi\tilde{C}n}{L} + \frac{Fg}{S} \right)^2}, \tag{D.81}$$

$$\nabla = \left(1 - \frac{g}{2AS} \cdot \left(\frac{2\pi\tilde{C}n}{L} + \frac{Fg}{S} \right) \right), \tag{D.82}$$

$$\tilde{\Lambda} = \frac{\tilde{C}}{2AL} \left(\frac{2\pi\tilde{C}n}{L} + \frac{gF}{S} \right), \tag{D.83}$$

$$\langle x \rangle = L \left(1 + \frac{F}{S} - \frac{g}{LS} 2\pi n \right) - \nabla \frac{k_B T}{2} \left(\frac{L}{Q_n} \coth \left[L \frac{Q_n}{A} \right] - \frac{A}{Q_n^2} \right), \tag{D.84}$$

$$\langle M_3 \rangle = \frac{2\pi\tilde{C}n}{L} + \frac{gF}{S} - \tilde{\Lambda} \cdot \frac{k_B T}{2} \left(\frac{L}{Q_n} \coth \left[L \frac{Q_n}{A} \right] - \frac{A}{Q_n^2} \right), \tag{D.85}$$

which for large F (equivalently large Q) becomes:

$$\langle x \rangle = L \left(1 + \frac{F}{S} - \frac{g}{LS} (2\pi n) \right) - \nabla \frac{k_B T}{2} \frac{L}{Q_n}, \tag{D.86}$$

$$\langle M_3 \rangle = \frac{2\pi\tilde{C}n}{L} + \frac{gF}{S} - \tilde{\Lambda} \cdot \frac{k_B T}{2} \frac{L}{Q_n}. \tag{D.87}$$

The extension Eq. (D.86) is the same as the one obtained in [134] using a different method.

A very long heterogenous chain

Most of the experimental work has been performed on long molecules. In the case of DNA, the number of base pairs in experiments range from $N \approx 1\text{ kbp}$ to $\sim 100\text{ kbps}$. Therefore we think

it is useful to present the solutions corresponding to this special scenario. We analyze the case when the number of monomers or base pairs in the system is so large that one can assume that the leading order terms $O(N^x)$ are much larger than $O(N^{x-1})$. Since $N \rightarrow \infty$ then $\cosh(h_j) \approx 1$ and $\sinh(h_j) \approx 0$. The extension per monomer and external torque per monomer are therefore given by:

$$\begin{aligned} \frac{\langle x \rangle}{N} &= \left\{ 1 + \frac{F}{S_1} + \frac{1}{C_1} \left[\left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) F - \frac{2\pi n - U^o}{\varpi} \right] \left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) - \nabla_1 \frac{k_B T}{2Q_1} \right\} x_1 l_1 \\ &+ \left\{ 1 + \frac{F}{S_2} + \frac{1}{C_2} \left[\left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) F - \frac{2\pi n - U^o}{\varpi} \right] \left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) - \nabla_2 \frac{k_B T}{2Q_2} \right\} x_2 l_2 \\ &- 2\pi \frac{n}{N} \frac{\Phi}{\varpi}, \end{aligned} \quad (\text{D.88})$$

$$\begin{aligned} \frac{\langle M_3 \rangle}{N} &= \left\{ \frac{F}{\varpi \tilde{C}_1} \left(\frac{g_1}{S_1} - \frac{\Phi}{\varpi} \right) - \tilde{\Lambda}_1 \frac{k_B T}{2Q_1} \right\} x_1 l_1 \\ &+ \left\{ \frac{F}{\varpi \tilde{C}_2} \left(\frac{g_2}{S_2} - \frac{\Phi}{\varpi} \right) - \tilde{\Lambda}_2 \frac{k_B T}{2Q_2} \right\} x_2 l_2 \\ &+ \frac{2\pi n + F\Phi - U^o}{\varpi N}. \end{aligned} \quad (\text{D.89})$$

where $x_1 + x_2 = 1$ are the fractions in each state. The expressions Eq. (D.88) and Eq. (D.89) show that in the limit of a long chain the contributions of two independent chains can simply be added up. This allows us to recover an analytical approximation for the free energy per monomer:

$$\begin{aligned} H_N &= \left\{ -\frac{1}{2\tilde{C}_1} \left(\frac{g_1 F}{S_1} - \lambda \right)^2 - \frac{F^2}{2S_1} + \frac{k_B T Q_1}{A_1} \right\} x_1 l_1 \\ &+ \left\{ -\frac{1}{2\tilde{C}_2} \left(\frac{g_2 F}{S_2} - \lambda \right)^2 - \frac{F^2}{2S_2} + \frac{k_B T Q_2}{A_2} \right\} x_2 l_2 \\ &+ \left[2\pi \frac{n}{N} + u_1^o x_1 l_1 + u_2^o x_2 l_2 \right] \lambda. \end{aligned} \quad (\text{D.90})$$

where we have used λ given in expression Eq. (D.65). Notice that simplifying the expression above in the case of a homogeneous chain recovers the Legendre transform of the free energy given by [56] in the controlled torque ensemble.

D.4 Link Lk and supercoiling parameter σ relations

$$Lk = \Delta Lk + Lk_o, \quad Lk_o = \frac{N}{10.5} = \frac{w_o L_o}{2\pi}, \quad w_o = \frac{2\pi}{3.54} \approx 1.77. \quad (\text{D.91})$$

Lk_o is the natural link of unperturbed B-DNA due to internal twist of the right-handed double helix. The number of turns applied to the system (which is initially in a B-DNA state) is equal to the excess link ΔLk stored in the filament, which is composed of the stored twist $\Delta T w$, the writhe of the curve describing the filament axis $W r$ and any changes to the natural-internal twist of the filament $\Delta T w_o$ (i.e unwinding of the B-DNA double helix when there is a phase transition):

$$\Delta Lk = n = \Delta T w + W r + \Delta T w_o = \Delta T w^{min} + \Delta T w_o, \quad (\text{D.92})$$

where

$$2\pi\Delta T w^{min} = \int_0^L u_3^{min}(s) ds, \quad 2\pi\Delta T w_o = \int_0^L u^o(s) ds, \quad (D.93)$$

and its discrete version is:

$$2\pi\Delta T w^{min} = \sum_{i=0}^N (u_{3,i}^{min} l_i), \quad 2\pi\Delta T w_o = \sum_{i=0}^N (u_{3,i}^o l_i). \quad (D.94)$$

The supercoiling parameter σ is given as a fraction of Lk_o [134]:

$$\sigma = \frac{\Delta Lk}{Lk_o} = \frac{Lk}{Lk_o} - 1. \quad (D.95)$$

D.4.1 Linking number and phase transitions

At a phase transition we have the constraint that the total number of base pairs remains fixed $N = N_1 + N_2$ such that the fraction of base pairs in state 1 is $x_1 = N_1/N$ and in state $x_2 = N_2/N$. The excess link can also be expressed as a fraction of the supercoiling value at the beginning σ_1 (phase 1) and end of the transition σ_2 (phase 2):

$$\frac{2\pi\Delta Lk}{w_o L_o} = \sigma = x_1 \sigma_1 + x_2 \sigma_2 = \frac{2\pi n_1}{w_o L_o} + \frac{2\pi n_2}{w_o L_o}. \quad (D.96)$$

For the case of a two-phase model using the definition of twist in its discrete form (note we set $u_1^o = 0$ as the reference natural twist) we have the following relations:

$$\frac{2\pi\Delta T w^{min}}{N} = u_{3,1}^{min} l_1 x_1 + u_{3,2}^{min} l_2 x_2, \quad \frac{2\pi\Delta T w_o}{N} = u_2^o l_2 x_2, \quad (D.97)$$

$$2\pi \frac{\Delta Lk}{N} = u_1^{min} l_1 x_1 + (u_2^{min} + u_2^o) l_2 x_2. \quad (D.98)$$

When there is a phase transition the new unperturbed internal link is given by $\hat{L}k_o = \Delta T w_o + Lk_o = N/\gamma$, where γ is the helical repeat (number of base pairs per turn) in the new phase.

$$\sigma_2^o = \frac{\Delta T w_o}{Lk_o} = \frac{\hat{L}k_o}{Lk_o} - 1, \quad (D.99)$$

and

$$\gamma = \frac{10.5}{\sigma_2^o + 1}. \quad (D.100)$$

Finally, from equations Eq. (D.96), Eq. (D.97), Eq. (D.98) and Eq. (D.95) it is clear that:

$$u_1^{min} = w_o \sigma_1, \quad u_2^{min} = w_o (\sigma_2 - \sigma_2^o) \frac{l_1}{l_2}, \quad \sigma_2^o = u_2^o \frac{l_2}{w_o l_1}. \quad (D.101)$$

In the literature it is common to express the change in internal link using the difference in helicity defined as

$$\Delta t w = \frac{2\pi}{10.5} \sigma_2^o. \quad (D.102)$$

Next we compute the critical number of turns σ_1 at the beginning and σ_2 at the end of the transition in single molecule experiments where the addition of negative turns leads to a phase transition from B-DNA into L-DNA [29, 57].

D.4.2 B-DNA to L-DNA transition where L-DNA is an effective single state

For notation purposes we let the fraction in state 1 (B-DNA) to be $x_1 = \alpha$ and state 2 (L-DNA) $x_2 = 1 - \alpha$. The variation of the free energy per monomer of a system under imposed F and n is given by:

$$dH_N = \frac{\partial H_N}{\partial F} dF + \frac{\partial H_N}{\partial n} dn + \frac{\partial H_N}{\partial \alpha} \frac{\partial \alpha}{\partial F} dF + \frac{\partial H_N}{\partial \alpha} \frac{\partial \alpha}{\partial n} dn. \quad (\text{D.103})$$

For a system at constant force, during a phase transition we have the requirement that the Gibbs free energy per unit monomer $G_{gibbs} = H_N - (2\pi n/N)M_3$ must be equal ($dG_{gibbs} = 0$), such that:

$$dG_{gibbs} = \left(\frac{\partial G_{gibbs}}{\partial M_3} \right) \partial M_3 = 2\pi \left(\frac{n}{N} \right) \partial M_3 = 0, \quad (\text{D.104})$$

which is the well known result that the torque M_3 is a constant during the phase transition [55, 85]. Using Eq. (D.103) we can express this condition as

$$dG_{gibbs} = \frac{\partial H_N}{\partial \alpha} \frac{\partial \alpha}{\partial n} dn, \quad (\text{D.105})$$

and since $\alpha(n)$ is not zero during the transition the requirement of two states coexisting is:

$$\partial H_N / \partial \alpha = 0. \quad (\text{D.106})$$

To obtain the value of the torque $\langle M_3 \rangle$ during the transition we use the free energy Eq. (D.90) and condition Eq. (D.106):

$$\begin{aligned} & \left\{ -\frac{1}{2\tilde{C}_1} \left(\frac{g_1 F}{S_1} - \lambda \right)^2 - \frac{F^2}{2S_1} + \frac{k_B T Q_1}{A_1} \right\} l_1 \\ & - \left\{ -\frac{1}{2\tilde{C}_2} \left(\frac{g_2 F}{S_2} - \lambda \right)^2 - \frac{F^2}{2S_2} + \frac{k_B T Q_2}{A_2} + u_2^o \lambda \right\} l_2 + \Delta G^o = 0. \end{aligned} \quad (\text{D.107})$$

We have added ΔG^o in the expression to account for the energy offset between the B-DNA and L-DNA ground states. Using $\lambda = u_{3,1}^{min} \tilde{C}_1 + g_1 F / S_1 = u_{3,2}^{min} \tilde{C}_2 + g_2 F / S_2$, we can find the $u_{3,1}^{min}$ and $u_{3,2}^{min}$ as a function of ΔG^o . From Eq. (D.101) we obtain the critical number of turns at which the transition starts $\sigma_{start} = \sigma_1$ and the number of turns at which the transition ends $\sigma_{end} = \sigma_2$. While the transition torque $\langle M_3 \rangle$ can be approximated using σ_1 or σ_2 in the single state expression Eq. (D.87).

The properties for B-DNA and L-DNA are listed in tables 5.1 and 5.2 in the main text, while for ΔG^o we use $2.5k_B T/\text{bp}$. This value for the energy offset is smaller than the one fitted by [29] $\Delta G^o \sim 3k_B T/\text{bp}$ and $\sim 3.3k_B T/\text{bp}$ (energy required for denaturation [191]). The values of the torque $\langle M_3 \rangle$, σ_1 and σ_2 vary slightly as a function of the force. For forces F between 1.6 to 36 pN, $\langle M_3 \rangle$ ranges from -10 to -11.2pNnm, σ_1 from -0.016 to -0.013 and σ_2 from -1.77 to -1.71. These estimates agree well with the reported measurements of the transition variables in [31] where $\sigma_2 = 0.02$ and $\langle M_3 \rangle = -10.1 \pm 0.1 \text{pNnm}$ at $F = 0.9 \text{pN}$ while $\langle M_3 \rangle = -10.9 \pm 0.3 \text{pNnm}$ at $F = 10 \text{pN}$.

When using $\Delta G^o \approx 2k_B T$ there is no significant change in σ_1 and σ_2 , but $\langle M_3 \rangle \approx -9 \text{pNnm}$ at large forces $F > 10 \text{pN}$ which is closer to the value reported at high salt concentrations in [31]. When using $\Delta G^o \approx 3k_B T$ again there is no significant change in σ_1 and σ_2 , but $\langle M_3 \rangle \approx -13 \text{pNnm}$ at large forces $F > 10 \text{pN}$ which is closer to the value reported at low salt concentrations in [31].

D.4.3 B-DNA to L-DNA transition where L-DNA = (Z+S)-DNA

Here we still let the fraction in state 1 (B-DNA) to be $x_1 = \alpha$ and state 2 (L-DNA) $x_L = (1 - \alpha)$. But since L-DNA is itself a two state system then $x_L = (x_{L,Z} + x_{L,S})(1 - \alpha)$, where $x_{L,Z} = 0.77$ is the fraction of L-DNA in Z-DNA form and $x_{L,S} = 0.23$ is the fraction of L-DNA in S-DNA form (as fitted to experiment in the main text). We remind the reader that $x_{L,Z}$ and $x_{L,S}$ are not functions of n , while α is the fraction describing the phase transition driven by n . Since $x_{L,Z}$ and $x_{L,S}$ are constant through the transition we can express the energy of state 2 as a weighted average and we obtain an analogous expression to Eq. (D.107) :

$$\begin{aligned}
0 &= \left\{ -\frac{1}{2\tilde{C}_1} \left(\frac{g_1 F}{S_1} - \lambda \right)^2 - \frac{F^2}{2S_1} + \frac{k_B T Q_1}{A_1} \right\} l_1 \\
&- \left[\left\{ -\frac{1}{2\tilde{C}_Z} \left(\frac{g_Z F}{S_Z} - \lambda \right)^2 - \frac{F^2}{2S_Z} + \frac{k_B T Q_Z}{A_Z} + u_Z^o \lambda \right\} l_Z + \Delta G_Z^o \right] x_{L,Z} \\
&- \left[\left\{ -\frac{1}{2\tilde{C}_S} \left(\frac{g_S F}{S_S} - \lambda \right)^2 - \frac{F^2}{2S_S} + \frac{k_B T Q_S}{A_S} + u_S^o \lambda \right\} l_S + \Delta G_S^o \right] x_{L,S}. \quad (\text{D.108})
\end{aligned}$$

The properties for Z-DNA and S-DNA are listed in table 5.2 in the main text. ΔG_Z^o is the energy offset between the B-DNA and Z-DNA ground states. The value of ΔG_Z^o that we have used came from doing a similar phase transition analysis between the B-DNA and Z-DNA. For the B-Z transition we use $\Delta G_Z^o = 1.5k_B T = 0.9\text{kcal/mol}$ per bp which is within the range of values listed in the literature [31, 192]. At $F = 1.4\text{pN}$, we obtain for the B-Z transition torque $\langle M_3^{B-Z} \rangle \approx 5\text{pNm}$ and the number of turns at which the transition starts $\sigma_{start}^{B-Z} \approx 0.007$, which compare well with the experimental data of the B-Z transition reported by [145] at the same force. For the energy offset between the B-DNA and S-DNA ground states we use $\Delta G_S^o = 5k_B T$. Using this value for the energy offset and for F between 1.6 to 36 pN, the solution to the B-L transition yields: $\langle M_3 \rangle \approx -10\text{pNm}$, $\sigma_1 \approx -0.13$ and $\sigma_2 \approx -1.7$.

D.5 Free energy expressions of a Discrete Rod: bending, twisting and stretching

In this section we provide the free energy expressions for the discrete rod used in chapter 5. Working with u_3 and ϵ as done in the main text in section D does provide accurate results for the response functions (first derivatives of the free energy), but it should not be used in the calculation for the free energy itself (or entropy). The procedure using u_3 and ϵ as generalized variable misses a constant term in the partition function (which is independent of the external controlled parameters F, M_3, n). When doing the statistical mechanics analysis and treating u_3^i (or ϵ_i) for $i \in [1, N]$ as the independent variables results in computing independent strains for each link, and the total strain is the sum of piecing together the strain segments for the entire rod. So although the total strain is the summation of the strains in single segments (that is why the response function calculations are correct), this procedure neglects the fact that at the nodes of the links the kinematic variables (displacements/position) have to be continuous. Taking into account the displacement compatibility condition by working with displacements/position variables provides the missing term in the free energy calculation. As before, we consider the potential energy of the rod given by:

$$E = \int_0^L \Lambda(s) ds \quad (\text{D.109})$$

where Λ is the free energy per unit length:

$$\Lambda(s) = \frac{A}{2} \left| \frac{d\mathbf{t}}{ds} \right|^2 + \frac{C}{2} u_3^2 + \frac{S}{2} \epsilon^2 + g u_3 \epsilon, \quad (\text{D.110})$$

$$\mathbf{t} = -\sin \theta \sin \psi \mathbf{e}_1 + \sin \theta \cos \psi \mathbf{e}_2 + \cos \theta \mathbf{e}_3, \quad \left| \frac{d\mathbf{t}}{ds} \right|^2 = \kappa^2 = \theta'^2 + \psi'^2 \sin^2 \theta, \quad (\text{D.111})$$

where \mathbf{t} is the tangent vector to the filament and $\left| \frac{d\mathbf{t}}{ds} \right|^2$ is the curvature of the rod. Next we define the arc-length in the deformed configuration to be $\hat{s} = s + \rho$ such that the stretch ϵ is :

$$\epsilon = \frac{d\hat{s}}{ds} - 1 = \rho'. \quad (\text{D.112})$$

where $()'$ denotes differentiation with respect to s . Similarly the local twist of a rod can always be with respect to a twist free configuration (Bishop frame) [193] such that:

$$u_3 = \omega', \quad (\text{D.113})$$

where ω is the twisting angle of the director frame \mathbf{d}_i with respect to the Bishop Frame. Therefore, we can rewrite the free energy density as:

$$\Lambda(s) = \frac{A}{2} \left| \frac{d\mathbf{t}}{ds} \right|^2 + \frac{C}{2} \omega'^2 + \frac{S}{2} \rho'^2 + g(\rho\omega)'. \quad (\text{D.114})$$

Completing the square of the energy in ω' and ρ' , we can solve for the minimum energy configuration. Next, we express the energy as quadratic form in fluctuations away from the minima, where $\delta(\rho') = \rho'_{min} - \rho' = (\rho_{min} - \rho)' = (\delta\rho)'$ expresses the fluctuation of ρ' . Similarly we express the energy in

terms of the fluctuation variable ω' . Then the energy of the system becomes (for two identical rods):

$$\begin{aligned}
E &= \int_0^L \left[\frac{A}{2} v_x'^2 + \left(F - \frac{\lambda^2}{4A} \right) \frac{v_x^2}{2} + \frac{S}{4} (\delta\rho)^2 + \frac{\hat{C}}{4} (\delta\omega)^2 \right] ds + E_{min}/2 \\
&+ \int_0^L \left[\frac{A}{2} v_y'^2 + \left(F - \frac{\lambda^2}{4A} \right) \frac{v_y^2}{2} + \frac{S}{4} (\delta\rho)^2 + \frac{\hat{C}}{4} (\delta\omega)^2 \right] ds + E_{min}/2,
\end{aligned} \tag{D.115}$$

where $\hat{C} = C - g^2/S$ is the effective twist modulus of the rod and

$$E_{min} = \int_0^L \left[-\frac{1}{2\hat{C}} \left(\frac{gF}{S} - \lambda \right) - \frac{F^2}{2S} + u^o \lambda \right] ds + 2\pi n \lambda. \tag{D.116}$$

In the following section we will calculate the partition function of the rod. Hence we remind the reader that for a random field-variable Φ (continuously-varying) and a Hamiltonian (Action) $H[\Phi]$ that is infinite dimensional $\Phi \in [-\infty, \infty]$, the partition function in classical mechanics is defined as the Wiener functional integral:

$$Z = \text{tr} [\exp(-\beta H)] = \int_{-\infty}^{-\infty} \mathcal{D}\Phi \exp(-\beta H), \tag{D.117}$$

where $\mathcal{D}\Phi$ indicates a sum over all possible configurations on the field satisfying the boundary conditions of the system. For convenience we will focus on the energy contribution to the action H of one of the two identical rods given in expression Eq. (D.115):

$$H = E_{single} = E_v + E_\rho + E_\omega + E_{min}/2.$$

In what follows, we have separated the calculations into the contribution of ω (twist), ρ (stretch) and the generalized bending variable ν .

D.5.1 E_ω : Twist Hamiltonian

Here we calculate the *conditional* Wiener functional integral Z_ω for the continuum Hamiltonian (Action) defined as [194]:

$$H_\omega = E_\omega = \int_0^L \frac{\hat{C}}{4} (\delta\omega)^2 ds, \tag{D.118}$$

$$Z_\omega = \int_{-\infty}^{-\infty} \mathcal{D}\omega(s) \exp(-\beta H_\omega). \tag{D.119}$$

Following a standard discretization procedure we divide the integral into N equal links such that the twist of the discrete rod is given by $N + 1$ variables (ω_i at each node):

$$H_{\omega,N} = \sum_{i=1}^N \frac{\hat{C}}{4} \left[\frac{\omega_i - \omega_{i-1}}{l} \right]^2 l, \quad \omega_0 = 0 \text{ and } \omega_N = \text{fixed}, \tag{D.120}$$

and we have assumed the rod to be uniform and homogeneous such that $\hat{C}_i = \hat{C}$. The functional integral Z_ω is defined by a limiting procedure from the finite-dimensional integral $Z_{\omega,N}$, which is obtained by replacing the infinite measure $[\mathcal{D}\omega(s)]$ by the conditional finite-dimensional measure

[194]:

$$\left[\frac{\beta \hat{C}}{4\pi l} \right]^{N/2} \prod_{i=1}^{N-1} d\omega_i$$

such that

$$Z_{\omega,N} = \left[\frac{\beta \hat{C}}{4\pi l} \right]^{N/2} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_{\omega,N}) \prod_{i=1}^{N-1} d\omega_i. \quad (\text{D.121})$$

We define $J = \beta \hat{C}/(4l)$ and we rewrite the Hamiltonian and partition function as:

$$\beta H_{\omega,N} = \sum_{i=1}^N J [\omega_i - \omega_{i-1}]^2, \quad \text{with } \omega_0 = 0 \text{ and } \omega_N = \text{fixed} \quad (\text{D.122})$$

$$\beta H_{\omega,N} = \mathbf{X}^T \mathbf{J}_{N-1} \mathbf{X} - \mathbf{C}^T \mathbf{X} + J\omega_N^2, \quad \text{with } \omega_0 = 0 \text{ and } \omega_N = \text{fixed} \quad (\text{D.123})$$

$$Z_{\omega,N} = \left[\frac{J}{\pi} \right]^{N/2} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_{\omega,N}), \prod_{i=1}^{N-1} d\omega_i \quad (\text{D.124})$$

where $\mathbf{X}^T = [\omega_1, \omega_2, \dots, \omega_{N-1}]$ is a vector containing the $N - 1$ kinematic variables,

$$\mathbf{C}^T = [0, \dots, 0, 2J\omega_N]$$

is a $N - 1$ vector and \mathbf{J}_{N-1} is $N - 1$ -dimensional tridiagonal matrix:

$$\mathbf{J}_{N-1} = \begin{bmatrix} 2J & -J & 0 & \dots & 0 \\ -J & 2J & -J & \dots & 0 \\ 0 & -J & \dots & \dots & 0 \\ 0 & 0 & \dots & 2J & -J \\ 0 & 0 & 0 & -J & 2J \end{bmatrix} \quad (\text{D.125})$$

with

$$\det \mathbf{J}_{N-1} = \prod_{i=1}^{N-1} \lambda_i, \quad (\text{D.126})$$

$$\lambda_1 = 2J \text{ and } \lambda_i = 2J - \frac{J^2}{\lambda_{i-1}} \text{ for } i > 2, \quad (\text{D.127})$$

$$\det \mathbf{J}_{N-1} = NJ^{N-1}. \quad (\text{D.128})$$

Next we compute the integrals:

$$\int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp[-\mathbf{X}^T \mathbf{J}_{N-1} \mathbf{X} + \mathbf{C}^T \mathbf{X}] \prod_{i=1}^{N-1} d\omega_i = \left[\frac{\pi^{N-1}}{\det \mathbf{J}_{N-1}} \right]^{1/2} \exp[\mathbf{C}^T \mathbf{J}_{N-1}^{-1} \mathbf{C}]. \quad (\text{D.129})$$

Using the recurrence relation for the tridiagonal matrix [195]:

$$\mathbf{J}_{N-1}^{-1} = \frac{\det \mathbf{J}_{N-2}}{\det \mathbf{J}_{N-1}} = \frac{N-1}{N} \frac{J^{N-2}}{J^{N-1}} = J^{-1} \left(1 - \frac{1}{N} \right), \quad (\text{D.130})$$

then

$$\exp[\mathbf{C}^T \mathbf{J}_{N-1}^{-1} \mathbf{C}] = \exp \left[J\omega_N^2 \left(1 - \frac{1}{N} \right) \right], \quad (\text{D.131})$$

and combining Eq. (D.124) and Eq. (D.129), we get to the final result :

$$Z_{\omega,N} = \frac{1}{N^{1/2}} \left[\frac{J}{\pi} \right]^{N/2} \left[\frac{\pi}{J} \right]^{\frac{N-1}{2}} \exp \left[J\omega_N^2 \left(1 - \frac{1}{N} \right) \right] \exp \left[-J\omega_N^2 \right], \quad (\text{D.132})$$

$$Z_{\omega,N} = \left[\frac{J}{N\pi} \right]^{1/2} \exp \left[-\frac{J\omega_N^2}{N} \right] = \left[\frac{\beta\hat{C}}{4\pi L} \right]^{1/2} \exp \left[-\frac{\beta\hat{C}\omega_N^2}{4L} \right]. \quad (\text{D.133})$$

D.5.2 E_ρ : Stretch Hamiltonian

Next we calculate the *unconditional* Wiener functional integral Z_ρ for the continuum Hamiltonian (Action) defined as [194]:

$$H_\rho = E_\rho = \int_0^L \frac{S}{4} (\delta\rho)^2 ds, \quad (\text{D.134})$$

$$Z_\rho = \int_{-\infty}^{-\infty} \mathcal{D}\rho(s) \exp(-\beta H_\rho). \quad (\text{D.135})$$

Following a standard discretization procedure we divide the integral into N equal links such that the stretch of the discrete rod is given by $N + 1$ variables (ρ_i at each node):

$$H_{\rho,N} = \sum_{i=1}^N \frac{S}{4} \left[\frac{\rho_i - \rho_{i-1}}{l} \right]^2 l, \quad \rho_0 = 0. \quad (\text{D.136})$$

Where we have assumed the rod to be uniform and homogeneous such that $S_i = S$. The functional integral Z_ρ is defined by a limiting procedure from the finite-dimensional integral $Z_{\rho,N}$, which is obtained by replacing the infinite measure $[\mathcal{D}\rho(s)]$ by the unconditional finite-dimensional measure [194]:

$$\left[\frac{\beta S}{4\pi l} \right]^{N/2} \prod_{i=1}^N d\rho_i,$$

such that

$$Z_{\rho,N} = \left[\frac{\beta S}{4\pi l} \right]^{N/2} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_{\rho,N}) \prod_{i=1}^N d\rho_i. \quad (\text{D.137})$$

We define $K = \beta S/(4l)$ and rewrite the Hamiltonian and partition function as:

$$\beta H_{\rho,N} = \sum_{i=1}^N K [\rho_i - \rho_{i-1}]^2 = \mathbf{X}^T \mathbf{K}_N \mathbf{X}, \quad \text{with } \rho_0 = 0, \quad (\text{D.138})$$

$$Z_{\rho,N} = \left[\frac{K}{\pi} \right]^{N/2} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_{\rho,N}) \prod_{i=1}^N d\rho_i, \quad (\text{D.139})$$

where $\mathbf{X}^T = [\rho_1, \rho_2, \dots, \rho_N]$ is a vector containing the N kinematic variables and \mathbf{K}_N is N -dimensional tridiagonal matrix:

$$\mathbf{K}_N = \begin{bmatrix} 2K & -K & 0 & \dots & 0 \\ -K & 2K & -K & \dots & 0 \\ 0 & -K & \dots & \dots & 0 \\ 0 & 0 & \dots & 2K & -K \\ 0 & 0 & 0 & -K & K \end{bmatrix} \quad (\text{D.140})$$

Next we can expand the $\det \mathbf{K}_N$ about the last column :

$$\det \mathbf{K}_N = K \det \mathbf{K}_{N-1} - K^2 \det \mathbf{K}_{N-2}, \quad (\text{D.141})$$

where \mathbf{K}_{N-1} is a $N - 1$ -dimensional tridiagonal matrix

$$\mathbf{K}_{N-1} = \begin{bmatrix} 2K & -K & 0 & \dots & 0 \\ -K & 2K & -K & \dots & 0 \\ 0 & -K & \dots & \dots & 0 \\ 0 & 0 & \dots & 2K & -K \\ 0 & 0 & 0 & -K & 2K \end{bmatrix} \quad (\text{D.142})$$

$$\det \mathbf{K}_{N-1} = \prod_{i=1}^{N-1} \lambda_i \quad (\text{D.143})$$

$$\lambda_1 = 2K \text{ and } \lambda_i = 2K - \frac{K^2}{\lambda_{i-1}} \text{ for } i > 2 \quad (\text{D.144})$$

hence

$$\det \mathbf{K}_{N-1} = NK^{N-1}, \quad (\text{D.145})$$

and similarly

$$\det \mathbf{K}_{N-2} = (N - 1)K^{N-2}, \quad (\text{D.146})$$

Next we can compute the Gaussian integral:

$$\begin{aligned} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_{\rho,N}) \prod_{i=1}^N d\rho_i &= \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp[-\mathbf{X}^T \mathbf{K}_N \mathbf{X}] \prod_{i=1}^N d\rho_i \\ &= \left[\frac{\pi^N}{\det \mathbf{K}_N} \right]^{1/2}, \end{aligned} \quad (\text{D.147})$$

Combining Eq. (D.141), Eq. (D.145) and Eq. (D.146):

$$\det \mathbf{K}_N = (K)^N,$$

and finally from Eq. (D.139) and Eq. (D.147) we get to the final result:

$$Z_{\rho,N} = 1 \quad (\text{D.148})$$

D.5.3 E_ν : Bending Hamiltonian - Partially clamped

Here we consider a filament subjected to tension under partially-clamped boundary conditions as defined in Su and Purohit [49]. We are going to calculate the *conditional* Wiener functional integral Z_ν for the continuum Hamiltonian (Action) defined as:

$$H_\nu = E_\nu = \int_0^L \frac{A}{2} \nu'^2 + P\nu^2 ds, \quad (\text{D.149})$$

$$Z_\nu = \int_{-\infty}^{-\infty} \mathcal{D}\nu(s) \exp(-\beta H_\nu). \quad (\text{D.150})$$

Following a standard discretization procedure we divide the integral into N equal links such that

the stretch of the discrete rod is given by N variables (ν_i at each node):

$$H_{\nu, N-1} = \sum_{i=1}^{N-1} \frac{A}{2} \left[\frac{\nu_i - \nu_{i-1}}{l} \right]^2 l + P\nu_i^2 l, \quad \nu_0 = \nu_{N-1} = 0. \quad (\text{D.151})$$

Here ν_i angle is the angle that each link makes with the horizontal \mathbf{e}_3 axis, therefore for N links there is N angles instead of $N + 1$ angles. We have assumed the rod to be uniform and homogeneous such that $A_i = A$. The functional integral Z_ν is defined by a limiting procedure from the finite-dimensional integral $Z_{\nu, N}$, which is obtained by replacing the infinite measure $[\mathcal{D}\nu(s)]$ by the conditional finite-dimensional measure (since $\nu_0 = \nu_{N-1} = 0$) are constrained:

$$\left[\frac{\beta A}{2\pi l} \right]^{\frac{N-1}{2}} \prod_{i=1}^{N-2} d\nu_i,$$

such that

$$Z_{\nu, N-1} = \left[\frac{\beta A}{2\pi l} \right]^{\frac{N-1}{2}} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_{\nu, N-1}) \prod_{i=1}^{N-2} d\nu_i. \quad (\text{D.152})$$

We define $\kappa = \beta A/(2l)$ and $q = \beta Pl$ and we rewrite the Hamiltonian and partition, function as:

$$\beta H_{\nu, N-1} = \sum_{i=1}^{N-2} \kappa [\nu_i - \nu_{i-1}]^2 + q\nu_i^2 = \vec{\Theta} \mathbf{M} \vec{\Theta}^T \quad \nu_0 = \nu_{N-1} = 0, \quad (\text{D.153})$$

$$Z_{\nu, N-1} = \left[\frac{\kappa}{\pi} \right]^{\frac{N-1}{2}} \left[\frac{\pi^{N-2}}{\det \mathbf{M}} \right]^{1/2} = \left[\frac{\kappa^{N-1}}{\pi \det \mathbf{M}} \right]^{1/2}, \quad (\text{D.154})$$

where $\vec{\Theta} = [\nu_2, \dots, \nu_{N-1}]$ and \mathbf{M} is $N - 2$ -dimensional tridiagonal matrix:

$$\mathbf{M} = \begin{bmatrix} 2\kappa + q & -\kappa & 0 & \dots & 0 \\ -\kappa & 2\kappa + q & -\kappa & \dots & 0 \\ 0 & -\kappa & \dots & \dots & 0 \\ 0 & 0 & \dots & 2\kappa + q & -\kappa \\ 0 & 0 & 0 & -\kappa & 2\kappa + q \end{bmatrix} \quad (\text{D.155})$$

$$\det \mathbf{M} = \prod_{i=1}^{N-2} \lambda_i, \quad (\text{D.156})$$

$$\lambda_1 = 2\kappa + q \text{ and } \lambda_i = 2\kappa + q - \frac{\kappa^2}{\lambda_{i-1}} \text{ for } i > 2. \quad (\text{D.157})$$

From the results in [49] we have:

$$\det \mathbf{M} = \frac{r^{N-1} - d^{N-1}}{r - d} = \frac{r^{N-1}}{r - d} \left(1 - \frac{d^{N-1}}{r^{N-1}} \right), \quad (\text{D.158})$$

where

$$r = \frac{2\kappa + q + \sqrt{4\kappa q + q^2}}{2} \quad \text{and} \quad d = \frac{2\kappa + q - \sqrt{4\kappa q + q^2}}{2}. \quad (\text{D.159})$$

Therefore the partition function becomes

$$Z_{\nu, N-1} = \left[\left(\frac{\pi}{r-d} \right) \left(\frac{r^{N-1}}{\kappa^{N-1}} \right) \left(1 - \frac{d^{N-1}}{r^{N-1}} \right) \right]^{-1/2}. \quad (\text{D.160})$$

Next we take the limit as $N \rightarrow \infty$ and $l \rightarrow 0$ with $Nl = L$ fixed:

$$r - d = \sqrt{4\kappa q + q^2} = \beta \sqrt{2[AP] + [Pl]^2} = \beta \sqrt{2[AP]} = \beta Q, \quad (\text{D.161})$$

$$\left(1 - \frac{d^{N-1}}{r^{N-1}} \right) \approx 1 - \exp \left[-2L \sqrt{\frac{2P}{A}} \right] = 1 - \exp \left[-2L \frac{Q}{A} \right], \quad (\text{D.162})$$

$$\frac{r^{N-1}}{\kappa^{N-1}} \approx \left(1 + \frac{\beta Q}{2\kappa} \right)^{N-1} = \left(1 + \frac{QL}{AN} \right)^{N-1} \approx \exp \left[L \frac{Q}{A} \right]. \quad (\text{D.163})$$

Such that Eq. (D.160) becomes:

$$Z_{\nu, \infty} = \left[\left(\frac{\pi}{\beta Q} \right) \left(\exp \left[L \frac{Q}{A} \right] \right) \left(1 - \exp \left[-2L \frac{Q}{A} \right] \right) \right]^{-1/2}. \quad (\text{D.164})$$

Then free energy of a partially clamped rod with only bending fluctuations is given by:

$$\beta G = -\ln Z_{\nu, \infty} = \frac{1}{2} \ln \left[\left(\frac{\pi}{\beta Q} \right) \left(\exp \left[L \frac{Q}{A} \right] \right) \left(1 - \exp \left[-2L \frac{Q}{A} \right] \right) \right], \quad (\text{D.165})$$

$$\beta G = \frac{1}{2} \ln \left[\exp \left[L \frac{Q}{A} \right] \left(1 - \exp \left[-2L \frac{Q}{A} \right] \right) \right] + \frac{1}{2} \ln \frac{\pi}{\beta Q}. \quad (\text{D.166})$$

By letting $P = F/2$ (only force) and performing $\langle x \rangle = -\partial G / \partial F$:

$$\langle x \rangle = \frac{1}{4\beta} \frac{L}{\sqrt{AF}} \frac{1 + \exp \left(\frac{2L\sqrt{F/A}}{A} \right)}{1 - \exp \left(\frac{2L\sqrt{F/A}}{A} \right)} + \frac{1}{4F\beta} = -\frac{1}{4\beta} \frac{L}{\sqrt{AF}} \coth \left[L \sqrt{\frac{F}{A}} \right] + \frac{1}{4F\beta} \quad (\text{D.167})$$

which is the well-known force-extension relation for a two dimensional homogeneous fluctuation rod.

D.5.4 Free energy of a homogeneous chain subjected to tension and twist

In this section we piece together the results from sections D.5.1, D.5.2 and D.5.3, such that we obtain the total partition function of the filament used in the main text in section 4. We start with the definition for the partition function of the entire system:

$$Z = \int_{-\infty}^{-\infty} \mathcal{D}\Phi \exp(-\beta H) = Z_{E_{min}, \infty} (Z_{\omega, \infty} Z_{\rho, \infty} Z_{\nu, \infty})^2, \quad (\text{D.168})$$

where the square term comes from considering the entire system (two-fluctuating chains) and

$$Z_{E_{min}} = \exp[-\beta E_{min}]. \quad (\text{D.169})$$

Therefore combining the previous results :

$$Z = \left[\left(\frac{\pi}{\beta Q} \right) \left(\exp \left[L \frac{Q}{A} \right] \right) \left(1 - \exp \left[-2L \frac{Q}{A} \right] \right) \right]^{-1} \left[\frac{\beta \hat{C}}{4\pi L} \right] \exp \left[-\frac{\beta \hat{C} \omega_N^2}{2L} \right] \exp [-\beta E_{min}], \quad (\text{D.170})$$

and after some simplifications we arrive to the continuous version of the partition function of a chain under imposed tension and torsion:

$$Z = \left[\frac{\beta^2 \hat{C} Q}{4\pi^2 L} \right] \exp \left[-\frac{\beta \hat{C} \omega_N^2}{2L} - \beta E_{min} \right] \left[\left(\exp \left[L \frac{Q}{A} \right] \right) \left(1 - \exp \left[-2L \frac{Q}{A} \right] \right) \right]^{-1}. \quad (\text{D.171})$$

The continuous version of the free energy is readily computed to be:

$$\begin{aligned} \beta G = -\ln Z &= \left[\beta E_{min} + \frac{\beta \hat{C} \omega_N^2}{2L} \right] + \ln \left[\frac{4\pi^2 L}{\beta^2 \hat{C} Q} \right] \\ &+ \ln \left\{ \left[\left(\exp \left[L \frac{Q}{A} \right] \right) \left(1 - \exp \left[-2L \frac{Q}{A} \right] \right) \right] \right\} \end{aligned} \quad (\text{D.172})$$

D.5.5 E_ν : Bending Hamiltonian - (Hinged-Hinged)

In this section we compute the free energy for a hinged-hinged rod subject to tension as described in Su and Purohit [49]. Although this boundary conditions are the not applicable when the rod is also subjected to torsion, we find it pertinent to document this result here. We calculate a *special* case of the Wiener functional integral Z_ν for the continuum Hamiltonian (Action) defined as:

$$H_\nu = E_\nu = \int_0^L \frac{A}{2} \nu'^2 + P \nu^2 ds \quad (\text{D.173})$$

$$Z_\nu = \int_{-\infty}^{\infty} \mathcal{D}\nu(s) \exp(-\beta H_\nu) \quad (\text{D.174})$$

but with special boundary conditions (initial conditions of the diffusion process which correspond to the hinged-hinged rod).

Following a standard discretization procedure we divide the integral into N equal links such that the stretch of the discrete rod is given by N variables (ν_i at each node):

$$H_{\nu,N} = \sum_{i=1}^{N-1} \frac{A}{2} \left[\frac{\nu_i - \nu_{i-1}}{l} \right]^2 l + P \nu_i^2 l, \quad (\text{D.175})$$

Since the rod is hinged-hinged here ν_0 and ν_N are free to fluctuate. The functional integral Z_ν is defined by a limiting procedure from the finite-dimensional integral $Z_{\nu,N}$, which is obtained by replacing the infinite measure $[\mathcal{D}\nu(s)]$ by the finite-dimensional measure (since $\nu_0 = \nu_{N-1}$ are free to fluctuate):

$$\frac{1}{2\pi} \left[\frac{\beta A}{2\pi l} \right]^{\frac{N-1}{2}} \prod_{i=0}^{N-1} d\nu_i \quad (\text{D.176})$$

Expression Eq. (D.176) is obtained by solving a diffusion equation (Weiner Process) where the initial condition (initial distribution) $w(\nu, s)$ is a uniform distribution:

$$w(\nu_0, s)|_{s=0} = \frac{1}{b-a} = \frac{1}{2\pi},$$

such that

$$Z_{\nu, N} = \frac{1}{2\pi} \left[\frac{\beta A}{2\pi l} \right]^{\frac{N-1}{2}} \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} \dots \int_{-\infty}^{-\infty} \exp(-\beta H_\nu + I k g(\nu_i)) dk \prod_{i=0}^{N-1} d\nu_i, \quad (\text{D.177})$$

where $I^2 = -1$ and $g(\nu)$ is the position constraint as shown in [49]:

$$g(\nu) = \int_0^L \sin \nu(s) ds \approx \sum_{i=0}^{N-1} \nu_i \frac{L}{N} = \sum_{i=0}^{N-1} \nu_i l. \quad (\text{D.178})$$

In [49] the constraint g has been express as $g(\nu)/l$ to make it dimensionless.

We define $\kappa = \beta A/(2l)$ and $q = \beta Pl$ and we follow [49] such that the Hamiltonian, constraint and partition function are given by:

$$\beta H_{\nu, N} - I k g = \vec{\Theta} \mathbf{M} \vec{\Theta}^T, \quad (\text{D.179})$$

$$Z_{\nu, N} = \frac{1}{4\pi^2} \left[\frac{\kappa}{\pi} \right]^{\frac{N-1}{2}} \left[\frac{\pi^{N+1}}{\det \mathbf{M}} \right]^{1/2} = \frac{1}{4\pi} \left[\frac{\kappa^{N-1}}{\det \mathbf{M}} \right]^{1/2}, \quad (\text{D.180})$$

where $\vec{\Theta} = [v_0, \dots, v_N, k]$ and \mathbf{M} is $N + 1$ -dimensional matrix:

$$\mathbf{M} = \begin{bmatrix} \kappa + q & -\kappa & 0 & \dots & -Il/2 \\ -\kappa & 2\kappa + q & -\kappa & \dots & -Il/2 \\ 0 & -\kappa & \dots & \dots & -Il/2 \\ 0 & 0 & \dots & \kappa + q & -Il/2 \\ -Il/2 & -Il/2 & -Il/2 & -Il/2 & 0 \end{bmatrix} \quad (\text{D.181})$$

$$\det \mathbf{M} = \frac{1}{4N} \prod_{i=1}^{N-1} \lambda_i, \quad (\text{D.182})$$

$$\lambda_1 = 2\kappa + q \text{ and } \lambda_i = 2\kappa + q - \frac{\kappa^2}{\lambda_{i-1}} \text{ for } i > 2. \quad (\text{D.183})$$

From [49]:

$$\prod_{i=1}^{N-1} \lambda_i = \frac{r^N - d^N}{r - d} = \frac{r^N}{r - d} \left(1 - \frac{d^N}{r^N} \right) \quad (\text{D.184})$$

where

$$r = \frac{2\kappa + q + \sqrt{4\kappa q + q^2}}{2} \quad \text{and} \quad d = \frac{2\kappa + q - \sqrt{4\kappa q + q^2}}{2}. \quad (\text{D.185})$$

Therefore the partition function becomes

$$Z_{\nu, N} = \frac{1}{4\pi} \left[\frac{\kappa^{N-1}}{\det \mathbf{M}} \right]^{1/2} = \frac{1}{2\pi} \left[\left(\frac{1}{r-d} \right) \left(\frac{r^N}{N \kappa^{N-1}} \right) \left(1 - \frac{d^N}{r^N} \right) \right]^{-1/2} \quad (\text{D.186})$$

Next we take the limit as $N \rightarrow \infty$ and $l \rightarrow 0$, then:

$$r - d = \sqrt{4\kappa q + q^2} = \beta \sqrt{2[AP] + [Pl]^2} = \beta \sqrt{2[AP]} = \beta Q, \quad (\text{D.187})$$

$$\left(1 - \frac{d^N}{r^N}\right) \approx 1 - \exp\left[-2L\sqrt{\frac{2P}{A}}\right] = 1 - \exp\left[-2L\frac{Q}{A}\right], \quad (\text{D.188})$$

$$\frac{r^N}{\kappa^N} \approx \left(1 + \frac{\beta Q}{2\kappa}\right)^N = \left(1 + \frac{QL}{AN}\right)^N \approx \exp\left[L\frac{Q}{A}\right], \quad (\text{D.189})$$

and

$$\frac{\kappa}{N} = \frac{\beta A}{2L}. \quad (\text{D.190})$$

Therefore, in the continuous limit Eq. (D.186) becomes:

$$Z_{\nu,\infty} = \left[\frac{2\pi^2 A}{QL} \left(\exp\left[L\frac{Q}{A}\right]\right) \left(1 - \exp\left[-2L\frac{Q}{A}\right]\right)\right]^{-1/2} \quad (\text{D.191})$$

Then the free energy is given by:

$$\beta G = -\ln Z_{\nu,\infty} = \frac{1}{2} \ln \left[\left(\frac{2\pi^2 A}{QL}\right) \left(\exp\left[L\frac{Q}{A}\right]\right) \left(1 - \exp\left[-2L\frac{Q}{A}\right]\right) \right] \quad (\text{D.192})$$

$$\beta G = \frac{1}{2} \ln \left[\exp\left[L\frac{Q}{A}\right] \left(1 - \exp\left[-2L\frac{Q}{A}\right]\right) \right] + \frac{1}{2} \ln \frac{2\pi^2 A}{QL} \quad (\text{D.193})$$

by letting $P = F/2$ (only force) and performing $\langle x \rangle = -\partial G / \partial F$:

$$\langle x \rangle = \frac{1}{4\beta} \frac{L}{\sqrt{AF}} \frac{1 + \exp\left(2L\sqrt{F/A}\right)}{1 - \exp\left(2L\sqrt{F/A}\right)} + \frac{1}{4F\beta} = -\frac{1}{4\beta} \frac{L}{\sqrt{AF}} \coth\left[L\sqrt{\frac{F}{A}}\right] + \frac{1}{4F\beta} \quad (\text{D.194})$$

which is the well-known force-extension relation for a two dimensional homogeneous fluctuation rod.

D.6 More Results

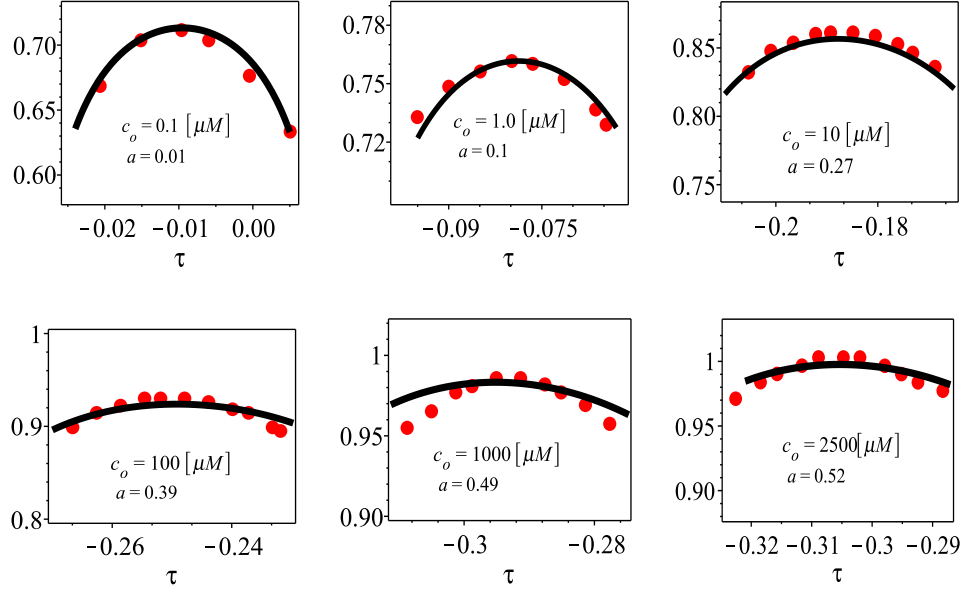


Figure D.1: Normalized extension $\langle x \rangle / L$ as a function of the normalized twist τ . We compare the results of our 2-state chain model (black solid line) with the experiments in [27] (red circles). The experiments were performed using $N = 20.6\text{kp}$ and $F = 0.25\text{pN}$. The fraction a of the DNA where EtBr has bound is presented in each figure as a function of c_o . The values of the properties of B-DNA are presented in table 5.1. For EtBr we use the same properties as in the main text but replace $S_2 = 1220\text{pN}$ and $g_2 = 385\text{pNm}$.

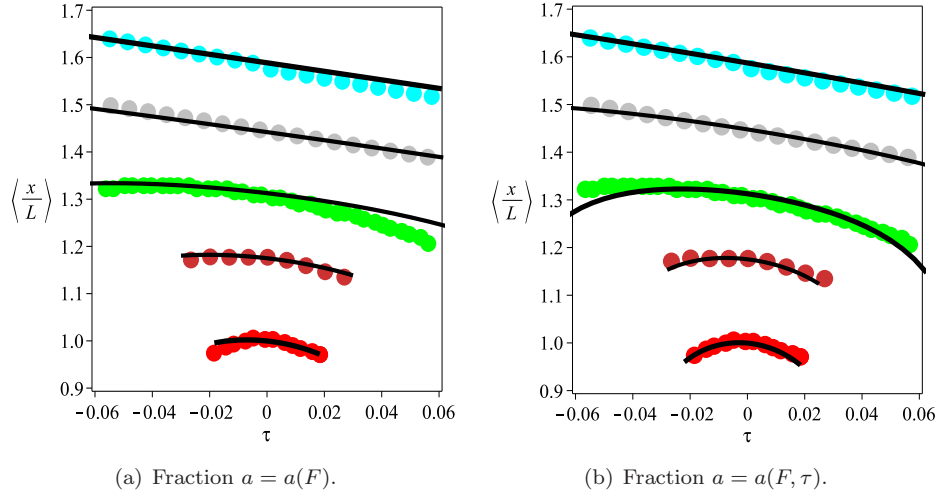


Figure D.2: (a) $F = [0.25, 0.5, 1, 3, 10]$ pN and $a_o(F) = [0.52, 0.55, 0.57, 0.60, 0.67]$. (b) Same values but assume linear relationship between applied applied turns (torque) and fraction $a(F, \tau) = a_o(F) - 1.3\tau$. This linear approximation has been also used in [28]. Both in (a) and (b) for each force the unwinding shift Δz_{EtBr} per intercalation event is different. We have used the same properties as in Fig. D.1.

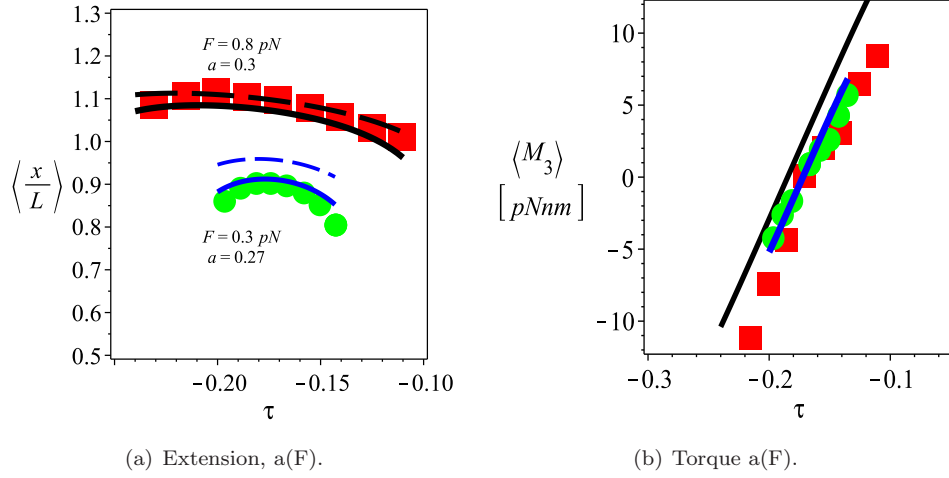


Figure D.3: Blue lines correspond to $F = 0.3\text{pN}$ and black lines to $F = 0.8\text{pN}$. Green circles and red squares are the data series from [28] for $F = 0.3\text{pN}$ and $F = 0.8\text{pN}$ respectively. For the solid lines we have used the same properties as in Fig. D.1. Dashed line calculated using $A_2 = 54k_B T$ which is the value obtained for $c_o = 10\mu\text{M}$ fitted by [28] using single state WLC model.

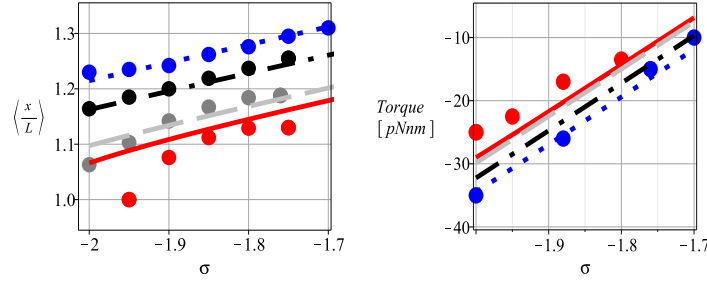


Figure D.4: L-DNA as two states: S-DNA + Z-DNA. Circles correspond to the experimental data in [29]. Color code is the same as in Fig. 5.8. Red-solid line corresponds to $F = 8.5\text{pN}$, gray-dash to $F = 12\text{pN}$, black-dash to $F = 24\text{pN}$ and blue-dot to $F = 36\text{pN}$. Here we use $A_2 = 11.6k_B T$ for the bending modulus of S-DNA as measured by [30], the rest of the properties used for S-DNA and Z-DNA are presented in table 5.2.

Appendix E

Appendix - Chapter 6

E.1 More Results

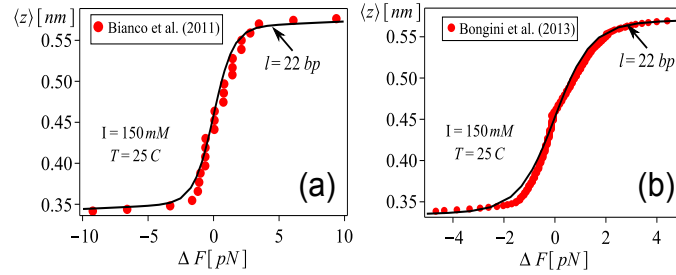


Figure E.1: Comparison of our theoretical predictions (solid lines) of the force-extension curves with experimental data in Bianco et al. [35] and Bongini et al [34]. Since experimental conditions are very similar to those in Fig 3(a), we use the same parameters $A_2 = 105.6\text{pN}$ and $A_4 = 500\text{pN}$. The predictions with $l = 22\text{bp}$ are in good agreement with experiment. Here we have used $F_c \approx 65.5\text{pN}$ as reported in [34, 35].

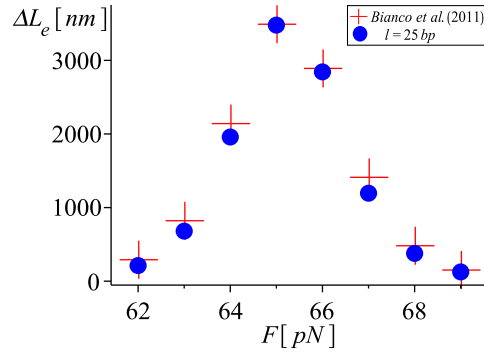


Figure E.2: Change in extension ΔL_e at equilibrium (steady state) corresponding to 1pN force-step protocol. Conditions are the same as in Bianco et al. [35] experiments at $T = 25\text{C}$ (See Fig. E.1) . Red crosses correspond to Bianco et al. [35] Monte Carlo simulations using a segmented chain composed of 1936 two-state units, each of length 25bp ($L = 48.4\text{kbp}$). Blue circles are the predictions of our analytical model using the same $l = 25\text{bp}$. Besides the value of l , here we have used the same parameters as in Fig. E.1.

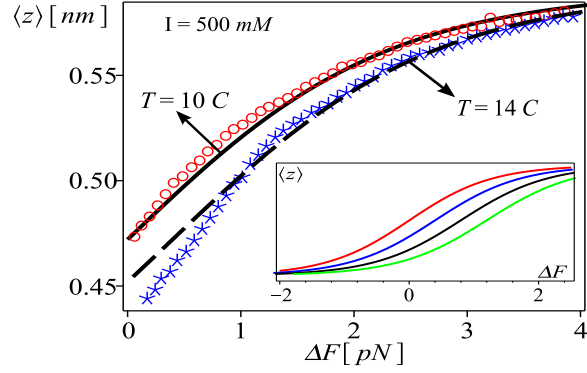


Figure E.3: Temperature effects on the force-extension curve at $I = 500\text{mM}$. Red circles ($T = 10\text{C}$) and blue markers $T = 14\text{C}$ correspond to [30] experimental data. Lines are our theoretical predictions using $l = 15\text{bp}$, $A_2 = 120\text{pN}$, $A_4 = 500\text{pN}$ (same as in Fig. 4 in the main text), where we have used the phenomenological relation Eq. (E.1) with $\eta \approx 0.1[\text{pN/K}]$ and $F^\circ \approx 50\text{pN}$ [30]. In the inset we use increments of 2C , going from 10C (red) up to 16C (green) using same parameters as in main figure.

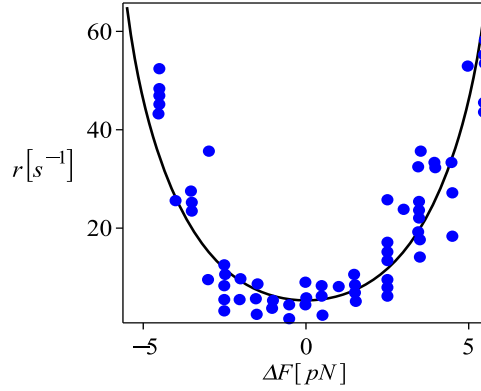


Figure E.4: Relaxation rate $r = k_L + k_R$ in dsDNA overstretching experiments at $I = 150\text{mM}$. Blue markers corresponds to *square-wave protocol* data in [34], where $F_c \approx 66\text{pN}$. Black solid line correspond to our theoretical predictions using $l = 22\text{bp}$. The *square-wave protocol* consists of applying a large force step $\sim 30\text{pN}$ during loading, waiting until system reaches equilibrium, and subsequently unloading the molecule with a force-step of the same magnitude [34]. As in the main text, where we fitted r to 2pN force-step experiments, we have used $\gamma \sim 200$.

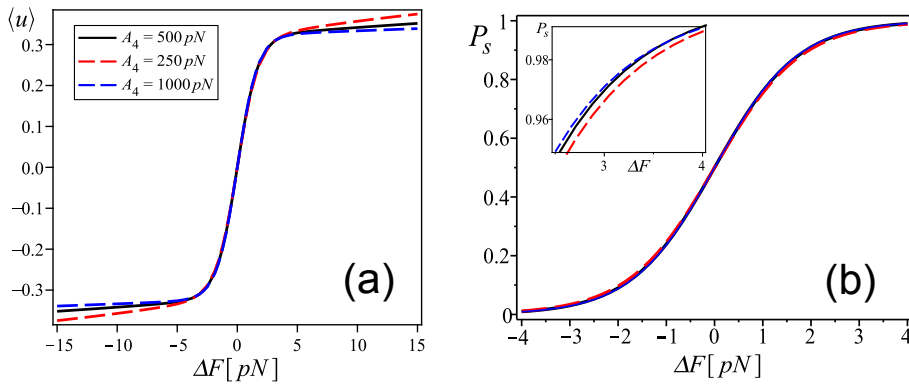


Figure E.5: Effect of A_4 on the average stretch $\langle u \rangle$ and on the probability of segments in the overstretched state (P_s). Here we have used parameters corresponding to the overstretching transition in Fig. E.1, where $A_2/A_4 = 0.2$, $l = 22$ bp and $T = 25$ C. Both in (a) and (b) the black solid line corresponds to $A_4 = 500$ pN (as used throughout the main text), blue dashed line corresponds to $A_4 = 1000$ pN and red-dashed line corresponds to $A_4 = 250$ pN. In (a) we present the solutions for $\langle u(F) \rangle$. As seen from the graph there is almost no difference between the three solutions in the range of forces δF where the transition takes place and δF remains approximately constant. For $|\Delta F| = |F - F_c| \gg 0$ the three solutions have different slopes. In (b) we present the solutions for $P_s(F)$, where it is evident that both, P_s and the width of the transition, are unaffected by A_4 . The inset shows a magnification of P_s near $F \in [3, 4]$ pN.

E.2 Empirical expressions for the critical force as a function of temperature and ionic concentration.

Fixed I and variable T For fixed salt concentration, Zhang et al. [30, 32] found a linear empirical relation between the critical force F_c^T and temperature T at which the non-hysteretic B-to-S transition takes place:

$$F_c^T(T) = F^o + \eta T. \quad (\text{E.1})$$

The exact phenomenological values of the ordinate F^o and slope η depend on the salt concentration I used in the experiments. At $I = 0.5\text{M}$ the values are $F^o \approx 50\text{pN}$ and $\eta \approx 0.1[\text{pN/K}]$. In the case of the B-to-M (close-end set up) transition at lower salt concentration $I \approx 20\text{mM}$ and higher temperatures $T > 25\text{C}$ [32], there is also a linear relationship between F_c^T and temperature T , but in this case the slope η is negative. Once again the exact value of the slope and ordinate are a function of the salt concentration I [30, 32]. For $I \approx 20\text{mM}$, the corresponding slope is $\eta \approx -0.9[\text{pN/K}]$ and the ordinate is $F^o \approx 90\text{pN}$.

Using λ DNA at $I = 150\text{mM}$, Zhang et al. [30] found that F_c^T increases from approximately 65pN at $T = 10\text{C}$ to approximately 67pN at $T = 20\text{C}$. While for the same experimental conditions, Bongini et al. [34] found that the critical force decreases from $F_c^T \approx 71\text{pN}$ at $T = 10\text{C}$ to approximately 65.5pN at $T = 25\text{C}$.

Fixed T and variable I

Similarly, for fixed T , Zhang et al. [30] found an empirical relation between the critical force and the critical concentration of salt at which the non-hysteretic transition B to S DNA takes place:

$$F_c^I(I) = F^s + k_B T l_B \eta_I \ln \left[\frac{I}{I_o} \right]. \quad (\text{E.2})$$

where $I_o = 1\text{M}$ is the reference salt concentration, k_B is the Boltzmann constant, T is Kelvin and $l_B \sim 0.71\text{nm}$ is the Bjerrum length. For $T = 284\text{K}$ the ordinate $F^s \sim 67[\text{pN}]$ and the slope $\eta_I \sim 0.53[\text{pN}]$, but in general the phenomenological constants are functions of T [30].

E.3 Chain Kinetics

When the interface boundary is sharp the phase transition can be approximated using a mean-field model where spatial fluctuations are not present. Given the potential H_j , at $C = 0$ the energetic cost required for a segment of length l to undergo a spatially homogeneous phase transition is given by [196]:

$$E_l = \int_0^l [H_j(u_b) - H_j(u_o)] ds = \frac{A_2^2 l}{4A_4} = (u_o)^4 A_4 l. \quad (\text{E.3})$$

For the values used in the overstretching analysis in section 6.3 $2u_o \approx L - \bar{L} \in [0.5 - 0.7]\text{pN}$, $A_4 \approx 500\text{pN}$ and $l \in [22 - 60]\text{bp}$ yields $E_l \gg k_B T$. Therefore there is separation of time scales in the activation process, with the mean escape time Θ representing a slow process (quasi-stationary) in comparison with the other relevant fast timescales [174]. Similarly, for $C \neq 0$, but near the transition point (there are two wells), we will denote E_l^L as the left energy barrier, E_l^R as the right energy barrier and still use $\bar{u}_b = u_b$ for the saddle potential (See Fig. E.6). For small values of C the saddle point can be approximated to be (see section E.3.1) $\bar{u}_b = (2A_2)^{-1}C$. In the DNA overstretching experiments $|C| = |\Delta F| \approx 2\text{pN}$, while A_2 is of the order 10^2pN , hence we let the saddle point be $\bar{u}_b \approx u_b = 0$. The methods we follow will remain valid as long as $(2A_2)^{-1}C \approx 0$, and $[E_l^L, E_l^K] \gg k_B T$.

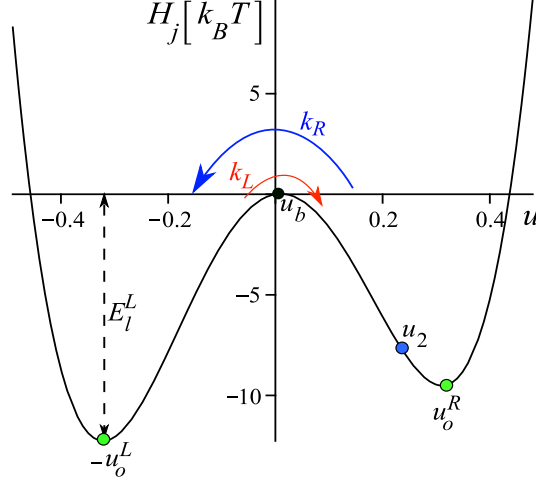


Figure E.6: Potential H_j landscape. We have used $A_2 = 100\text{pN}$, $A_4 = 500\text{pN}$ and $l \approx 25\text{bp}$. These values correspond to the overstretching transition fits found in section 4 of the main text at room temperature. At $C = 2\text{pN}$ the left well is deeper and $k_R > k_L$. From the graph it is clear that the energy barrier $E_l^L \gg k_B T$ and $\bar{u}_n \approx u_b = 0$.

Next, we calculate the thermally activated escape rate k_L from B-to-overstretched. In the overdamped case, the non-equilibrium probability density function $\rho(u, t)$ obeys the Smoluchowski equation [174]:

$$\frac{\partial \rho(u, t)}{\partial t} = \gamma \left[\frac{\partial \rho(u, t)}{\partial u} \frac{\partial^2 H_j}{\partial u^2} + k_B T \frac{\partial^2 \rho(u, t)}{\partial u^2} \right]. \quad (\text{E.4})$$

Under strong friction conditions (overdamped) [174, 197] the deviations from the equilibrium Boltzmann probability $\hat{\rho}(u)$:

$$\hat{\rho}(u) = Z_\rho^{-1} \exp -\beta H_j, \quad (\text{E.5})$$

can be safely neglected inside the well near the point $-u_o^L$ (Z_ρ is the normalization factor), such that thermal equilibrium in the initial (left) well is maintained over time. But near the saddle u_b there would be friction-induced deviations from thermal equilibrium, hence one must consider diffusion over the barrier described by the non-equilibrium probability $\rho(u, t)$, with $\rho \rightarrow 0$ for $u > u_b$ [197]. Next, to find k_L we follow the flux over population method [174]. In steady state $\partial_t \rho(u, t) = -\partial_u j(u, t) = 0$, the non-equilibrium probability $\rho(u)$ generates constant flux j_o :

$$-j_o = \gamma \left[\frac{\partial H_j}{\partial u} + k_B T \frac{\partial}{\partial u} \right] \rho(u). \quad (\text{E.6})$$

Following Kramer's method the non-equilibrium probability ρ can be evaluated imagining a source point near the bottom of the left well ($u_1 < -u_o^L$) (particle is injected) and an sink $u_2 > u_b$ (particle is absorbed) as depicted in Fig E.6. Following Kramer's initial ideas, the non-equilibrium probability ρ can be expressed as [174]:

$$\rho(u) = f(u) \hat{\rho}(u), \quad (\text{E.7})$$

where $f(u)$ satisfy the thermal equilibrium assumption $f(u_1) \approx 1$ and $f(u_2) \approx 0$ [197].

Combining Eq. (E.6) and Eq. (E.7) we get:

$$-j_o = \gamma f(u) \left[\hat{\rho} \frac{\partial H_j}{\partial u} + k_B T \frac{\partial \hat{\rho}(u)}{\partial u} \right] + \gamma \hat{\rho} k_B T \frac{\partial f(u)}{\partial u}, \quad (\text{E.8})$$

but the first term in brackets is zero, since $\hat{\rho}$ is the equilibrium thermal distribution (zero flux), and therefore:

$$j_o = -\gamma \hat{\rho} k_B T \frac{\partial f(u)}{\partial u}. \quad (\text{E.9})$$

Using chain rule of integration (Leibniz Formula) from a point u on the left well to u_2 , where the probability $f(u_2)$ vanishes, we obtain:

$$f(u) = \frac{j_o}{\gamma k_B T} \int_u^{u_2} \hat{\rho}(x) dx. \quad (\text{E.10})$$

$$\rho(u) = \frac{j_o}{\gamma k_B T} \exp(-\beta H_j) I_b, \quad (\text{E.11})$$

where

$$I_b = \int_u^{u_2} \exp(\beta H_j(x)) dx \approx \frac{\sqrt{2k_B T \pi}}{w_b}. \quad (\text{E.12})$$

For $E_l^L \gg k_B T$, significant contributions to the integral Eq. (E.12) will come mainly from the neighborhood of the global maximum point $u_b \in [u, u_2]$, and therefore the integral can be estimated following Laplace's Method (steepest descent) [174].

$$\begin{aligned} I_b &\approx \int_{-\infty}^{\infty} \exp\left(-\beta \frac{w_b^2}{2}(u - u_b^2)\right) du \\ &\approx \frac{\sqrt{2k_B T \pi}}{w_b}, \end{aligned} \quad (\text{E.13})$$

where

$$w_b^2 = \left| \frac{\partial^2 H_j}{\partial u^2} \right|_{u=u_b}. \quad (\text{E.14})$$

Finally the rate is given by:

$$k_L = \frac{j_o}{n_o}, \quad (\text{E.15})$$

where n_o denotes the (non-equilibrium) population density inside the initial well [174]:

$$n_o = \int_{-\infty}^{u_b} \rho(u) du = \frac{j_o I_b}{\gamma k_B T} \int_{-\infty}^{u_b} \exp(-\beta H_j) du, \quad (\text{E.16})$$

hence,

$$k_L = \frac{\gamma k_B T}{I_b Z_L}, \quad Z_L = \int_{-\infty}^{u_b} \exp(-\beta H_j) du. \quad (\text{E.17})$$

Similarly, the rate k_R (from the right well to left well) is :

$$k_R = \frac{\gamma k_B T}{I_b Z_R}, \quad Z_R = \int_{u_b}^{\infty} \exp(-\beta H_j) du. \quad (\text{E.18})$$

Next to evaluate integrals Z_L and Z_R , instead of using steepest-descent-method about u_o^L (as it is common within Kramer's theory [174]), we follow methods similar to the calculation of the partition function Z_m in section 3 of the main text. In what follows we present the calculations leading to the final expressions for k_L and k_R :

With $u_b \approx 0$, and the definitions $a = \beta A_4 l$, $b = -\beta A_2 l$, $c = \beta Cl$, the integrals given above are:

$$Z_R = \int_0^{+\infty} \exp[-au^4 - bu^2 - cu] du \quad (\text{E.19})$$

and

$$\begin{aligned} Z_L &= \int_{-\infty}^0 \exp[-au^4 - bu^2 - cu] du \\ &= - \int_0^{-\infty} \exp[-au^4 - bu^2 - cu] du \\ &= \int_0^{+\infty} \exp[-ay^4 - by^2 + cy] dy \end{aligned} \quad (\text{E.20})$$

Next we do a Taylor expansion in the linear term,

$$Z_R = \sum_{n=0}^{\infty} \frac{(-c)^n}{n!} \int_0^{+\infty} u^n \exp[-au^4 - bu^2] du, \quad (\text{E.21})$$

and separate the sum in odd and even terms,

$$\begin{aligned} Z_R &= \sum_{m=0}^{\infty} \frac{(-c)^{2m}}{(2m)!} \int_0^{+\infty} [u^{2m} \exp(-au^4 - bu^2)] du \\ &+ \sum_{\hat{m}=0}^{\infty} \frac{(-c)^{2\hat{m}+1}}{(2\hat{m}+1)!} \int_0^{+\infty} [u^{2\hat{m}+1} \exp(-au^4 - bu^2)] du. \end{aligned} \quad (\text{E.22})$$

Similarly for Z_L we obtain:

$$\begin{aligned} Z_L &= \sum_{m=0}^{\infty} \frac{(c)^{2m}}{(2m)!} \int_0^{+\infty} [u^{2m} \exp(-au^4 - bu^2)] du \\ &+ \sum_{\hat{m}=0}^{\infty} \frac{(c)^{2\hat{m}+1}}{(2\hat{m}+1)!} \int_0^{+\infty} [u^{2\hat{m}+1} \exp(-au^4 - bu^2)] du. \end{aligned} \quad (\text{E.23})$$

We define:

$$Z_1 = \sum_{m=0}^{\infty} \frac{(c)^{2m}}{(2m)!} \int_0^{+\infty} [u^{2m} \exp(-au^4 - bu^2)] du, \quad (\text{E.24})$$

and

$$Z_2 = \sum_{\hat{m}=0}^{\infty} \frac{(c)^{2\hat{m}+1}}{(2\hat{m}+1)!} \int_0^{+\infty} [u^{2\hat{m}+1} \exp(-au^4 - bu^2)] du, \quad (\text{E.25})$$

such that:

$$Z_R = Z_1 - Z_2, \quad Z_L = Z_1 + Z_2 \quad (\text{E.26})$$

The analytical solution for the integrals in in the form given in equations Eq. (E.24) and Eq. (E.25) is given in Ryzhik and Gradshteyn [178]:

$$\begin{aligned}
\Phi(x) &= 2 \int_0^{+\infty} u^{2x} \exp[-au^4 - bu^2] du & (E.27) \\
&= 2 \int_0^{+\infty} y^x \exp[-ay^2 - by] \left(\frac{y^{-1/2}}{2}\right) dy \\
&= \frac{\Gamma(x+1/2)}{(2a)^{\frac{2x+1}{4}}} \exp\left(\frac{q^2}{4}\right) U(x, q),
\end{aligned}$$

where $U(x, q)$ is the parabolic cylinder function [179] and $q^2 = b^2/2a$, such that:

$$Z_1 = \sum_{m=0}^{\infty} \frac{c^{2m}}{(2m)!} \frac{\Phi(m)}{2}, \quad Z_2 = \sum_{\hat{m}=0}^{\infty} \frac{c^{2\hat{m}+1}}{(2\hat{m}+1)!} \frac{\Phi(\hat{m} + \frac{1}{2})}{2}. \quad (E.28)$$

The final expressions for the transition rates are :

$$k_L = \gamma \sqrt{\frac{k_B T}{2\pi}} w_b Z_L^{-1}, \quad k_R = \gamma \sqrt{\frac{k_B T}{2\pi}} w_b Z_R^{-1}, \quad (E.29)$$

with Z_L and Z_R as defined above.

E.3.1 Appendix: Approximation of the equilibrium value of u for small C .

For small values of C , equation Eq. (6.2) can be linearized around the zero-field solution u^* , which yields for small deviations α :

$$g(\alpha) = 4A_4(u^* + \alpha)^3 - 2A_2(u^* + \alpha) + C = 0 \quad (E.30)$$

Next do a Taylor expansion around $\alpha = 0$:

$$g \approx g|_{\alpha=0} + \left. \frac{\partial g}{\partial \alpha} \right|_{\alpha=0} \alpha, \quad (E.31)$$

$$12A_4 u^* \alpha - 2A_2 \alpha + C = 0. \quad (E.32)$$

Hence for $u^* = u_o$ and small values of C we can approximate:

$$\alpha = -\frac{C}{4A_2} \quad \text{and} \quad \bar{u}_o = u_o - \frac{C}{4A_2}, \quad (E.33)$$

and for $u^* = u_b = 0$:

$$\alpha = \frac{C}{2A_2} \quad \text{and} \quad \bar{u}_b = \frac{C}{2A_2}. \quad (E.34)$$

E.4 Asymmetric Potentials

Here we present a sample of how our methods can be extended to use other potentials. We briefly discuss the methods using two asymmetric potentials. For convenience, we will write the expressions in terms of $a = \beta A_4 l$, $b = -\beta A_2 l$ and $c = \beta C l$. With this notation, the symmetric quartic potential used in the main text is:

$$\beta H_j = au^4 + b^2 + cu. \quad (\text{E.35})$$

E.4.1 Appendix: Hamiltonian including the cubic term

Now we introduce the contribution of a cubic term wu^3 to the potential H_j given by Eq. (E.35), such that the partition function is:

$$Z_c = \int_{-\infty}^{+\infty} \exp[-au^4 - wu^3 - bu^2 - cu] du. \quad (\text{E.36})$$

Next we proceed to do a Taylor expansion both in the linear and cubic terms:

$$Z_c = \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} \frac{(-c)^n}{n!} \frac{(-w)^r}{r!} \int_{-\infty}^{+\infty} u^{n+3r} \exp[-au^4 - bu^2] du \quad (\text{E.37})$$

Now the partition function Eq. (E.37) is equal to zero if $(n + 3r)$ is odd. But $Z \neq 0$ for both n and r even ($n = 2m$ and $r = 2k$) or n and r odd ($n = 2\hat{m} + 1$ and $r = 2\hat{k} + 1$), such that the net result $(n + 3r)$ is even. So the partition function $Z = Z_s$ for $(n + 3r)$ even can be expressed as:

$$\begin{aligned} Z_c &= \sum_{\{m,k\}=0}^{\infty} \frac{(-c)^{2m}}{(2m)!} \frac{(-w)^{2k}}{(2k)!} \int_{-\infty}^{+\infty} [u^{2s} \exp(-au^4 - bu^2)] du \\ &+ \sum_{\{\hat{m},\hat{k}\}=0}^{\infty} \frac{(-c)^{2\hat{m}+1}}{(2\hat{m}+1)!} \frac{(-w)^{(2\hat{k}+1)}}{(2\hat{k}+1)!} \int_{-\infty}^{+\infty} [u^{2\hat{s}} \exp(-au^4 - bu^2)] du, \end{aligned}$$

where $s = m + 3k$ and $\hat{s} = \hat{m} + 3\hat{k} + 2$. Next, making use of the result [178]:

$$\begin{aligned} \Phi(x) &= \int_{-\infty}^{+\infty} u^{2x} \exp[-au^4 - bu^2] du \\ &= 2 \int_0^{+\infty} y^x \exp[-ay^2 - by] \left(\frac{y^{-1/2}}{2}\right) dy \\ &= \frac{\Gamma(x + 1/2)}{(2a)^{\frac{2x+1}{4}}} \exp\left(\frac{q^2}{4}\right) U(x, q), \end{aligned} \quad (\text{E.38})$$

where $q = b/(\sqrt{2a})$. Therefore, the final expression is:

$$\begin{aligned} Z_c &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{c^{2m}}{(2m)!} \frac{w^{2k}}{(2k)!} \Phi(s) \\ &+ \sum_{\hat{m}=0}^{\infty} \sum_{\hat{k}=0}^{\infty} \frac{c^{2\hat{m}+1}}{(2\hat{m}+1)!} \frac{w^{(2\hat{k}+1)}}{(2\hat{k}+1)!} \Phi(\hat{s}). \end{aligned} \quad (\text{E.39})$$

E.4.2 Piece-wise Potential: Two quartics

Here we consider a piecewise potential, where each well is described by two distinct quartic potentials. Each of this quartics is of the form given by Eq. (E.35), then the resulting potential is:

$$\begin{aligned}
H_p(u) &= \{au^4 + bu^2 + cu\} \text{ for } u \leq u_p \\
&= \{\hat{a}u^4 + \hat{b}u^2 + cu\} \text{ for } u \geq u_p
\end{aligned}
\tag{E.40}$$

where $[a, \hat{a}] > 0$ to ensure stability and $[b, \hat{b} < 0]$. We impose the condition that the junction of the piece-wise potential u_p to be at the saddle point of both potentials ($u_p = u_b = \bar{u}_b$). Here u_b is the non-stable equilibrium solution to H_p in the left well and \hat{u}_b is the non-stable equilibrium solution to H_p on the right well.

Near the transition, for small c (see section E.3.1), the saddle points can be approximated to be:

$$\hat{u}_b = \frac{c}{2\hat{b}} \quad \text{and} \quad u_b = \frac{c}{2b}.
\tag{E.41}$$

Since these points must coincide we require that $\hat{b} = b$. The values of a , \hat{a} and b are fitted in a similar way to the procedure described in the main text (section 6.1.4). We require that the change in extension between compact and extended states near the midpoint of the transition is given by:

$$\frac{\delta z}{\text{bp}} = \frac{\bar{L} - L}{L} = u_o(b) + \hat{u}_o(\hat{b}) = \sqrt{\frac{-b}{2a}} + \sqrt{\frac{-b}{2\hat{a}}}.
\tag{E.42}$$

Next we fit a to the the slope of the force-extension curve for $C \ll 0$ (B-form) and \hat{a} to the the slope of the force-extension curve for $C \gg 0$ (overstretched form).

Then the partition function Z_p of one segment of length l is:

$$\begin{aligned}
Z_p &= \int_{-\infty}^{\infty} \exp H_p(u) du \\
&= \int_{-\infty}^{\bar{u}_b} \exp [(au^4 + bu^2 + cu)] du \\
&+ \int_{\bar{u}_b}^{\infty} \exp [(\hat{a}u^4 + bu^2 + cu)] du
\end{aligned}
\tag{E.43}$$

Under the approximation that $|b| \gg |c|$, which is consistent with the fitted values to DNA over stretching experiments, then the partition function is equal to:

$$Z_p = Z_L(a) + Z_R(\hat{a}) = Z_1(a) + Z_1(\hat{a}) + (Z_2(a) - Z_2(\hat{a})),
\tag{E.44}$$

where Z_1 and Z_2 are given by expressions Eq. (E.28). Here $Z_1(\cdot)$ and $Z_2(\cdot)$ are evaluated with the corresponding parameter a or \hat{a} . For $a = \hat{a}$, the value of $Z_p = 2Z_1$ equal the partition function Z used in the main text.

E.5 Kinetics of the chain with finite size interfaces: Domain Walls

In the main text in section 6.2 we presented a model for a chain of identical independent subsystems of length l , where the division of the chain into segments was founded on the idea that l represents the cooperative unit. Based on that model, in section 6.4 we presented the kinetics of the chain under the assumption that there are no spatial fluctuations of the order parameter and therefore no physical propagation of interface boundaries. In a more realistic scenario, for a phase transition between two states to physically take place over time, we need the existence and propagation of a domain wall (interface). Still keeping in mind that each subsystem of length l independently undergoes the phase transition, there must be one nucleation of a domain wall per subsystem. Kink or domain wall solutions are present when in addition to the potential $V(u) = A_4 u^4 - A_2 u^2$ there is an energetic contribution through the gradient (du/ds) . This new term accounts for the possible spatial fluctuations of the order parameter $u(s)$ and it allows for inhomogeneous solutions evident by the presence of domain walls separating the two homogeneous solutions [198, 199]. Under such a scenario, the energetics at the midpoint of the transition are given by the Ginzburg-Landau potential:

$$H_d = \int_0^L \left[\frac{K}{2} \left(\frac{du}{ds} \right)^2 + V(u) \right] ds. \quad (\text{E.45})$$

As in section 6.4, we assume that the system is highly over-damped meaning that the inertial term is negligible, such that ultimately we consider the dynamics of a quartic field on $u(s)$ that obeys the stochastic Ginzburg-Landau equation as done in Stein [200]¹:

$$\partial_t u(t, s) = \nu [K \partial_{ss} u(t, s) - \partial_u V(u)] + \sqrt{2\nu k_B T} \xi(s, t), \quad (\text{E.46})$$

where $\xi(s, t)$ is defined analogously to $\tilde{\xi}(t)$:

$$\langle \xi(s_1, t_1) \xi(s_2, t_2) \rangle = \delta(s_1 - s_2) \delta(t_1 - t_2). \quad (\text{E.47})$$

Next we follow the methods Stein and coworkers [196, 200–202]. Once again, in the weak noise limit, the classical activation rate of transition out of a stable well is given the Arrhenius type relation:

$$\Gamma \sim \Gamma_o \exp \left[-\hat{E}_{DP} / k_B T \right], \quad (\text{E.48})$$

where Γ_o is the Kramers transition rate prefactor and \hat{E}_{DP} is the activation barrier between stable states (energetic cost of creating interfaces). For the overdamped system obeying Eq. (E.46) driven by stochastic noise, the formula for Γ_o can be derived analytically [200].

In subsection E.5.1 we summarize the results of the kinetics analysis with propagating interfaces. In the remainder of subsections E.5.2 - E.5.7, we present the different methods and calculation required to elaborate subsection E.5.1.

¹Note that the kinetic coefficient ν appearing in Eq. (E.46) and γ appearing in Eq. (6.28) are not the same and they have different units.

E.5.1 Transition Rates Γ and Mean lifetime Θ of the overstretching transition

At equilibrium, the mean escape time over the energy barrier Θ (time that it takes to transition each segment of B-DNA to its overstretched form) is related to transition rate by [174]:

$$\Theta = \Gamma^{-1}. \quad (\text{E.49})$$

Assuming that the kinetic coefficient is a constant independent of the cooperativity of the system, then we can study how Γ_o varies as function l by computing the ratio $\Gamma_o^f = \Gamma_o(l_i)/\Gamma_o(l_j)$, where l_i and l_j are represent two systems with different cooperativity. Furthermore making use of expression Eq. (E.48) and Eq. (E.49) we can obtain the desired relation between the mean lifetime Θ and l :

$$\Theta^f = \frac{\Theta(l_i)}{\Theta(l_j)} \approx \exp \left[-\beta \left(\hat{E}_{DP}(l_j) - \hat{E}_{DP}(l_i) \right) \right] \left[\frac{\Gamma_o(l_j)}{\Gamma_o(l_i)} \right]. \quad (\text{E.50})$$

From the analysis in section 6.2 we have the values of A_2 , A_4 and l , but before we proceed to compute the transition rate factor and kink lifetime, we need to determine the value of the parameter K . Following the treatment in Krumhansl and Schrieffer [198], we make use of the relationship between the cost of creating a domain wall E_{DP} and the density of domain walls $\langle n_w \rangle$ for a one-dimensional system of independent, particles of width $\Delta = \sqrt{K/A_2}$ (See section E.5.4):

$$l \approx \sqrt{\frac{K}{A_2}} \exp \left[\frac{2}{3k_B T} \left[\frac{A_2^3 K}{A_4^2} \right]^{1/2} \right], \quad E_{DP} = \frac{2}{3} \left[\frac{A_2^3 K}{A_4^2} \right]^{1/2}. \quad (\text{E.51})$$

Here we have made use of the fact that there is one domain wall per segment of the chain ($\langle n_w \rangle = L/l$).

The formation and evolution of the domain walls along the system is dictated by the boundary conditions on $u(s)$ of each segment n making up the entire chain. The specific set of boundary conditions for each segments can not be easily determined from physics, but a set of boundary conditions that is analytically tractable that resembles the physical evolution of the domain walls is given by periodic boundary conditions (see Fig. E.8). The choice of periodic B.Cs is less restrictive than assuming either Neumann or Dirichlet boundary conditions for each segment of the chain, since they require specific knowledge a priori on either $u(s)$ or its space derivative. Certainly knowledge of the specific value of $u(s)$ for each segment along the DNA filament is a very strong assumption that rules out the case of Dirichlet B.Cs. On the other hand, studying the kinetics of propagating interfaces with Neumann B.Cs might describe closer a physical system where each segment l is completely independent of each other (as done in sections 6.2 and 6.4). The method using Neumann B.Cs is thoroughly described in Maier and Stein [196], while the results using periodic B.Cs are described in Stein [201] and Berglund and Gentz [203]. An outline of the analytical derivations using Periodic and Neumann B.Cs are presented in subsection E.5.5-E.5.7, and the final expressions for the kinetic rate prefactor Γ_o are given by Eq. (E.92) and Eq. (E.99) respectively. In Fig. E.7 we present a comparison of the mean-lifetime ratio obtained using Eq. (E.50) under Neumann (N) B.Cs, Periodic (P) B.Cs and the solution following methods from section 6.4 for sharp interfaces ($K=0$). The result shown in Fig. E.7 further validates the approximation of sharp interfaces, since the full solutions including propagating interfaces (N,P) share the same qualitative characteristics of the approximation ($K = 0$).

E.5.2 Calculation of the non-uniform solution in the infinite length case

The Hamiltonian including the possibility of having a domain wall (interface) is given by:

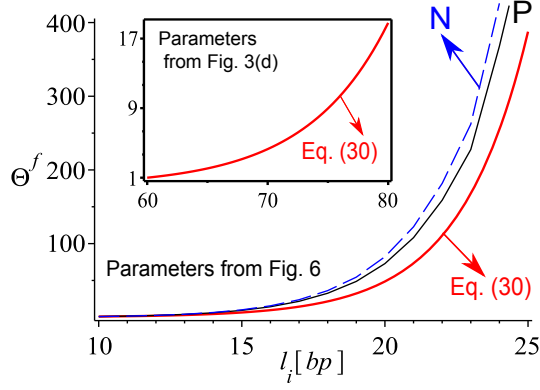


Figure E.7: Ratio Θ^f normalized by Θ evaluated at $l_j = 10\text{bp}$. Here we use $T = 25\text{C}$ and the same parameters as in Fig. 6.6. Blue-dashed line correspond to solution with Neumann B.Cs (N) using Eq. (E.99), black-solid line is the solution with Periodic B.Cs (P) using Eq. (E.92) and red solid line is the solution using expression Eq. (6.30) in the main text. For the red-solid line we use $\Gamma = k_L = k_R$ (since $C = 0$ here). Inset shows the ratio Θ^f computed from expression Eq. (6.30) in the main text at $T = 25\text{C}$ and using the same parameters as in Fig.6.3(d). In the inset we have normalized by Θ evaluated at $l_j = 60\text{bp}$. It is clear that as the cooperativity l increases, the time it takes to reach steady state $\propto r^{-1}$ increases rapidly.

$$H_d = \int_0^L \left[\frac{K}{2} \left(\frac{\partial u}{\partial s} \right)^2 + V(u) \right] ds, \quad V(u) = A_4 u^4 - A_2 u^2. \quad (\text{E.52})$$

For a one component (scalar) field $u(s)$ there are two possible value for the stretch in the ordered phase ($A_2 > 0$) [204]. Although the two possible states have the same energy, it is not possible to continuously deform one into the other and consequently the two states are separated by sharp domain walls[204]. A domain wall can be introduced by forcing the two sides of the system to be in different states. Assuming that L is very large we can then impose the boundary conditions

$$u(s = -\infty) = -u_o \quad \text{and} \quad u(s = +\infty) = u_o. \quad (\text{E.53})$$

In between, the most probable configuration is given by the minimum of the free energy functional which is given by:

$$\frac{\delta h(u)}{\delta u} = \frac{d}{ds} \frac{\partial h(u)}{\partial u'} - \frac{\partial h(u)}{\partial u} = 0 \quad (\text{E.54})$$

which yields

$$K \frac{d^2 u(s)}{ds^2} = \frac{dV(u)}{du} = -2A_2 u(s) + 4A_4 u(s)^3. \quad (\text{E.55})$$

After multiplying each side by (du/ds) and integrating once Eq. (E.55) from any point $s = s_w$ along the rod to $s = \infty$ we get:

$$\left[\left(\frac{du}{ds} \right)^2 \right]_{s_w}^{\infty} = \left[\frac{2}{K} V(u) \right]_{s_w}^{\infty}, \quad (\text{E.56})$$

At $s \rightarrow \pm\infty$ the derivative term (du/ds) must vanish and $u \rightarrow u_o$ and to simplify notation we let $u(s_w) = u_w$, hence :

$$\left(\frac{du_w}{ds}\right)^2 = \frac{2}{K} [V(u_w) - V(u)] \quad (\text{E.57})$$

$$\left(\frac{du_w}{ds}\right)^2 = 2 \left[\frac{A_2}{K} (-u_w^2 + u_o^2) + \frac{A_4}{K} (u_w^4 - u_o^4) \right]. \quad (\text{E.58})$$

But from $2A_4u_o^2 = A_2$ we obtain:

$$\left(\frac{du_w}{ds}\right)^2 = 2\frac{A_4}{K} [u_o^4 + u_w^4 - 2u_o^2u_w^2] = 2\frac{A_4}{K} [u_o^2 - u_w^2]^2, \quad (\text{E.59})$$

$$\begin{aligned} s - s_o &= \left(\frac{K}{2A_4}\right)^{1/2} \int_0^{u_o} \frac{du_w}{u_o^2 - u_w^2} \\ &= \left(\frac{K}{2A_4}\right)^{1/2} u_o^{-1} \tanh^{-1}\left(\frac{u_w}{u_o}\right) \\ &= \left(\frac{K}{A_2}\right)^{1/2} \tanh^{-1}\left(\frac{u_w}{u_o}\right), \end{aligned} \quad (\text{E.60})$$

where the domain wall $u_w = 0$ is located at s_o . Inverting relation Eq. (E.60) for the order parameter we get the final expression:

$$u_w(s) = u_o \tanh\left(\frac{s - s_o}{\Delta}\right), \quad \Delta = \sqrt{K/A_2}, \quad (\text{E.61})$$

where Δ is the width of a domain wall [204].

Energetic cost of creating a domain wall for the infinite rod E_{DP} the free energy cost of creating a domain wall is given by the energetic difference between the non-uniform solution and the stable solutions of the system[198, 205]:

$$E_{DP} = H_d - H_{min} = \int_{-\infty}^{\infty} \left[\frac{K}{2} \left(\frac{du_w}{ds}\right)^2 + V(u_w) - V(u_o) \right] ds. \quad (\text{E.62})$$

but from the relation in Eq. (E.57):

$$E_{DP} = K \int_{-\infty}^{\infty} \left[\left(\frac{du_w}{ds}\right)^2 \right] = 2 \int_{-\infty}^{\infty} [V(u_w) - V(u_o)] ds, \quad (\text{E.63})$$

where $V(\cdot)$ is given expression Eq. (E.52), u_o is given by expression Eq. (6.3) and the non-uniform solution u_w is given by Eq. (E.61). For the infinite limit case, after some algebraic manipulation we have:

$$E_{DP} = A_2 u_o \int_{-\infty}^{\infty} \left[\cosh^{-4}\left(\frac{s - s_o}{\Delta}\right) \right] ds = \frac{4}{3} u_o^2 A_2 \Delta = \frac{2}{3} \left[\frac{A_2^3 K}{A_4^2} \right]^{1/2}. \quad (\text{E.64})$$

E.5.3 Thermodynamics equations of motion: infinite length rod

Building on the phenomenological quartic model to describe the overstretching transition we now focus on the propagation of the domain walls following the treatment in [205]. The form of free

energy that we have been using through this work is:

$$H_d = \int \left[\frac{K}{2} (\partial_s u(t, s))^2 + V(u) + Cu \right] ds, \quad (\text{E.65})$$

where $V(u)$ is given by Eq. (E.52) and C plays the role of the external field (see Eqs. Eq. (6.1) and Eq. (6.4) in main text). Therefore, the local rate of propagation of the order parameter during the phase transition should be proportional to the thermodynamic driving force and obey the following linear equation of motion [205]:

$$\frac{\partial u(s, t)}{\partial t} = -\nu \frac{\delta H_d}{\delta u} = \nu [K \partial_{ss} u(t, s) - \partial_u V(u) - C], \quad (\text{E.66})$$

where s is space coordinate along the rod, t is the time variable, and ν is the kinetic coefficient which we assume to be a constant independent of external parameters. Equation Eq. (E.66) is just a statement that the order parameter evolves in time toward the local free energy minimum. If $C = 0$ the free energy density of the two phases is equal, but as we apply a external change in force $C = -\delta F$ the system departs from equilibrium and the interface must start moving [205]. Assuming that the range of the force over which the transition takes place $|2C| = |2\delta F|$ is small, it is possible to describe analytically the evolution of the domain wall. We constrain our analysis to a constant velocity (ρ) solution such that $u(s, t) = \hat{u}(s - \rho t)$, then Eq. (E.66) can be rewritten as [205]:

$$-\frac{\rho}{\nu} \frac{d\hat{u}}{ds} = K \frac{d^2 \hat{u}}{ds^2} - \frac{dV(\hat{u})}{d\hat{u}} - C, \quad (\text{E.67})$$

where the last term plays the role of an effective dissipative force. We still require that as $s \rightarrow \pm\infty$ such that $\hat{u} \approx \pm u_o$. Next multiply both sides of equation Eq. (E.67) by $d\hat{u}/ds$ and integrate once over the entire rod:

$$-\frac{\rho}{\nu} \int_{-\infty}^{\infty} \left(\frac{d\hat{u}}{ds} \right)^2 = \left[\frac{K}{2} \left(\frac{d\hat{u}}{ds} \right)^2 - V(\hat{u}) - C\hat{u} \right]_{-u_o}^{u_o} = 2Cu_o. \quad (\text{E.68})$$

But from expression Eq. (E.63) we know the first term is related to the energetic cost of creating a domain wall E_{DP} (at equilibrium) and therefore we get the desired relationship between the force range $2C$ and the velocity of propagation [205]:

$$-\frac{\rho}{\nu} \frac{E_{DP}}{K} = 2u_o C. \quad (\text{E.69})$$

Finally, to study the nucleation of domain walls, it is imperative to add the fluctuations to the system [205]. To see this, consider the total time derivate of the energy H [205]:

$$\frac{dH_d}{dt} = \int \frac{\delta H_d}{\delta u} \frac{du}{dt} ds = -\nu \int \left(\frac{\delta H}{\delta u} \right)^2 ds \leq 0, \quad (\text{E.70})$$

which means that energy of the system is always decreasing with time, and therefore it would exclude any activated process. In order to describe nucleation it is necessary to add a noise contribution[205], as presented in the analysis using the stochastic Ginzburgh-Landau equation Eq. (E.46) at the beginning of section E.5.

E.5.4 Appendix: Statistical Mechanics of Domain Walls: evaluating K

The partition function for a system of identical weakly interacting domain walls (non-interacting particles) is given by[198]:

$$Z_{DP} = \sum_{n_w=0}^{n_s} \frac{n_s!}{(n_s - n_w)!n_w!} \exp[-\beta n_w E_{DP}] = (1 + \exp[-\beta E_{DP}])^{n_s}, \quad (\text{E.71})$$

where the binomial prefactor is due to the presence of degenerate states in the system. Here n_w is the number of domain walls and n_s is the maximum number of locations where a domain wall can nucleate in the filament. The energetic cost of creating one domain wall E_{DP} for a system with Hamiltonian H_d Eq. (E.45) is given in Appendix E.5.2.

For the case where the energetic cost is high $\exp[-\beta E_{DP}] \ll 1$, the average number of domain walls $\langle n_w \rangle$ can be approximated by[198]:

$$\langle n_w \rangle \approx n_s \exp[-\beta E_{DP}]. \quad (\text{E.72})$$

Assuming that the partition function is dominated by the most probable value of n_w , which under the above approximations turns out to be the average number of domain walls $\langle n_w \rangle$ [198], the partition function Eq. (E.71) is given by:

$$Z_{DP} = \left(1 + \frac{\langle n_w \rangle}{n_s}\right)^{n_s} \approx \exp(\langle n_w \rangle). \quad (\text{E.73})$$

For completion we note that the partition function Z_{total} of a chain with identical non-interacting segments as described in the main text in section 6.2, where each segment is assumed to contain one domain wall (such that $n = \langle n_w \rangle$) and under the approximations described in this Appendix regarding domain walls, is given by:

$$Z_{total} \approx Z^{\langle n_w \rangle} \exp(\langle n_w \rangle). \quad (\text{E.74})$$

The value of n_s in Eq. (E.72) can be approximated by [198]:

$$n_s \sim L/\Delta, \quad (\text{E.75})$$

where Δ is the width of the domain wall given in Eq. (E.61). Making use of Eq. (E.75) yields an analytic expression of the number of domain walls $\langle n_w \rangle$ [198]²:

$$\frac{\langle n_w \rangle}{L} = l^{-1} = (\Delta)^{-1} \exp[-\beta E_{DP}]. \quad (\text{E.76})$$

E.5.5 Nondimensionalization of the stochastic Ginzburgh-Landau

It is convenient to express the stochastic evolution equation Eq. (E.46) presented in the main text in dimensionless form. Next we define the following dimensionless variables:

$$\begin{aligned} \epsilon &= u/u_c, & \tau &= t/t_c, & x &= s/s_c, \\ \hat{T} &= k_b T/T_c, & \text{and} & & \hat{\xi}(xs_c, \tau t_c) &= \xi(s, t). \end{aligned} \quad (\text{E.77})$$

Where the last expression above follows the definition of the $\hat{\xi}(x, \tau)$ function:

²Since the solution given in Eq. (E.76) for the average number of domain walls is based on assumption Eq. (E.75), in Currie et al. [206], Habib and Grant [207] the result for the density of domain walls is presented as a proportionality equation rather than equality. Therefore, we point out that in general a constant α can be added to right hand side of Eq. (E.76).

$$\begin{aligned}
\langle \xi(x_1 s_c, \tau_1 t_c) \xi(x_2 s_c, \tau_2 t_c) \rangle &= \frac{\delta(x_1 - x_2) \delta(\tau_1 - \tau_2)}{s_c t_c} \\
&= \frac{\langle \xi(x_1, \tau_1) \xi(x_2, \tau_2) \rangle}{s_c t_c}.
\end{aligned} \tag{E.78}$$

Therefore expression Eq. (E.46) becomes:

$$\begin{aligned}
\left(\frac{1}{4A_4 t_c \nu u_c^2} \right) \partial_\tau \epsilon(\tau, x) &= \frac{K}{4A_4 s_c^2 u_c^2} \partial_{xx} \epsilon(\tau, x) \\
&+ \frac{A_2}{2A_4 u_c^2} \epsilon - \epsilon^3 \\
&+ \hat{\xi}(x, \tau) \sqrt{\frac{\hat{T} T_c}{8A_4^2 u_c^6 s_c t_c \nu}},
\end{aligned} \tag{E.79}$$

where the field strength unit u_c , length unit s_c , time unit t_c and energy unit T_c are defined by:

$$u_c = \sqrt{\frac{A_2}{2A_4}}, \quad s_c = \sqrt{\frac{K}{2A_2}}, \quad t_c = \frac{1}{2\nu A_2}, \quad T_c = \sqrt{\frac{A_2^3 K}{2A_4^2}}. \tag{E.80}$$

Making use of the factors defined in Eq. (E.80), the final dimensionless expression for the stochastic Ginzburg-Landau equation is:

$$\partial_\tau \epsilon(\tau, x) = \frac{\delta \hat{h}}{\delta \epsilon} + \sqrt{2\hat{T}} \xi(\tau, x) = \partial_{xx} \epsilon - \epsilon^3 + \epsilon + \sqrt{2\hat{T}} \xi(\tau, x). \tag{E.81}$$

E.5.6 Appendix: Chain segments with Periodic B.Cs

We consider a system where the nucleation of domain walls takes place as kinks and antikinks along a system (See Fig. E.8), such that each block contains two segments of length l and two domain walls. Therefore each block satisfies periodic boundary conditions. Next we establish the behavior of the system in the absence of stochastic effects to find the stationary solutions $\hat{T} = 0$:

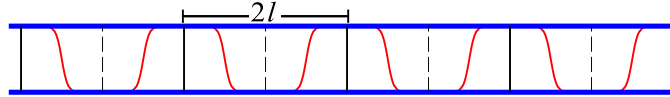


Figure E.8: Cartoon showing periodic boundary conditions. Depending on the boundary conditions (B.Cs), the domain walls can have a positive slope (kinks) or negative slope (anti-kink).

$$\partial_{xx} \epsilon = \epsilon^3 - \epsilon \tag{E.82}$$

There are three constant solutions to Eq. (E.82) given by $\partial_{xx} \epsilon = 0$. The two stable solutions are given by $\epsilon_o = \pm 1$ and the unstable solution is $\epsilon_u = 0$. However it is easy to check that the non-uniform solution periodic solution for Eq. (E.82) is [196]:

$$\epsilon_p(x) = \sqrt{\frac{2m}{m+1}} \operatorname{sn} \left(\frac{x}{\sqrt{m+1}} \middle| m \right) \tag{E.83}$$

where $\text{sn}(y|m) = \text{sn}(y, \sqrt{m})$ is the Jacobi elliptic function with parameter m (modulus $k = \sqrt{m}$) [179, 196]. The $\text{sn}(y|m) = \text{sn}(y + 4K(m)|m)$ is periodic with the quarter-period equal to $K(m)$, the complete integral of the first kind [179].

The problem at hand has periodic boundary conditions:

$$\epsilon(x_i) = \epsilon(x_f) \text{ and } \epsilon_x(x_i) = \epsilon_x(x_f), \quad (\text{E.84})$$

where $2\hat{l} = x_f - x_i$ is the dimensionless length measure of a domain with two domain walls (one kink and anti-kink) respectively (see Fig E.8). Note that because of the normalization procedure $\hat{l} = l/s_c$, where l is the physical length between domain walls. For the non-constant solution Eq. (E.83) to satisfy the periodic boundary conditions Eq. (E.84), $2\hat{l}/\sqrt{m+1}$ must be an integer multiple of a full period [196]:

$$2\hat{l} = 4\sqrt{m_p + 1}K(m_p), \quad (\text{E.85})$$

where m_p is the value of the parameter that satisfy the periodic B.Cs of the problem. The non-uniform solution ϵ_p is only present for $\hat{l} \geq \pi$, because in the limit of $m = 0$ we obtain $K(m) = \pi/2$. In the other limit as $m \rightarrow 1$, $K(m)$ diverges to infinity and hence $\hat{l} \rightarrow \infty$ is the only allowed solution. In this last scenario Eq. (E.83) degenerates in the infinite length case given by equation Eq. (E.61). In the following sections we use the non-uniform solution Eq. (E.83) since the values of the cooperative segment l and the phenomenological constants A_2 and K describing the S-DNA and M-DNA transition yield values of $\hat{l} > 2\pi$.

The activation barrier between stable states at the ends of each segment $x_f - x_i = 2\hat{l} \geq 2\pi$ is [196]:

$$\begin{aligned} \frac{\hat{E}_{DP}}{T_c} &= \int_0^{2\hat{l}} \left[\frac{1}{2} \left(\frac{d\epsilon_p}{dx} \right)^2 + \frac{(\epsilon_p^4 - \epsilon_o^4)}{4} - \frac{(\epsilon_p^2 - \epsilon_o^2)}{2} \right] dx \\ &= \frac{1}{3\sqrt{1+m_p}} \left[8E(m_p) - \frac{(1-m_p)(5+3m_p)}{1+m_p} K(m_p) \right]. \end{aligned} \quad (\text{E.86})$$

In the case where $\hat{l} \rightarrow \infty$ then $m_p \rightarrow 1$, the complete elliptic integral of the second kind $E(m_p) \rightarrow 1$, while $K(m_p)$ diverges to infinity in a logarithmic fashion [208]

$$(1 - m \rightarrow 16 \exp(-2K(m_p))).$$

Therefore in the $\hat{l} \rightarrow \infty$ limit:

$$\hat{E}_{DP} = \frac{4\sqrt{2}}{3} T_c = \frac{4}{3} \left[\frac{A_2^3 K}{A_4^2} \right]^{1/2}, \quad (\text{E.87})$$

which is twice the energy of required to create single domain wall Eq. (E.51). In this case the energy doubles since the non-constant periodic solution makes two swings between $\epsilon_o = \pm 1$ as x varies along the segment $2\hat{l}$ in order to satisfy the periodic B.Cs. For $\hat{l} > \sim 10$ in the periodic case, \hat{E}_{DP} can be approximated using expression Eq. (E.87).

The procedure to calculate rate prefactor Γ_o appearing in Eq. (E.48) can be found in [196, 200–203, 209], where the authors discuss the solutions associated with different boundary conditions. Next we give a brief summary of the procedure. To calculate the rate prefactor Γ_o appearing in Eq. (E.48), it is necessary to study the fluctuation about the stable ϵ_o and unstable ϵ_p configurations. The full details of the procedure can be found in [196, 200–203, 209], where the authors discuss the solutions associated with different boundary conditions. Next we give a summary of the method. Consider a small perturbation η about the stable solution $\epsilon_S = \epsilon_o = \pm 1$ and a perturbation η about the periodic

non-constant solution $\epsilon_U = \epsilon_p$, such that $\epsilon = \epsilon_b + \eta$, where $b = \{S, U\}$. The next step is to linearize the zero-noise evolution operators $\partial_\tau \epsilon = -\delta \hat{h} / \delta \epsilon$ (see Eq. (E.81)) and to leading order:

$$\partial_\tau \eta = -\mathbf{\Lambda}_b \eta, \quad (\text{E.88})$$

where $\mathbf{\Lambda}_b$ is the Hessian $\delta^2 \hat{h} / \delta \epsilon^2$ evaluated at $\epsilon = \epsilon_b$. Next, expression Eq. (E.88) can be diagonalized by decomposing fluctuation about the stable and unstable states into normal modes:

$$\partial_\tau \eta_i = -\mathbf{\Lambda}_b \eta_i = -\lambda_i^b \eta_i, \quad (\text{E.89})$$

where λ_i are the eigenvalues and have units of $[\tau^{-1}]$. Then the rate factor Γ_o^τ in $[\tau^{-1}]$ units is [200]:

$$\Gamma_o^\tau = \frac{1}{2\pi} \sqrt{\left| \frac{\det \mathbf{\Lambda}_s}{\det \mathbf{\Lambda}_u} \right|} |\lambda_o^u|, \quad (\text{E.90})$$

where λ_o^u is the only negative eigenvalue of the operator $\mathbf{\Lambda}_u$ [196]. Introducing the physical time unit $t_c \tau = t$ then Eq. (E.90) in units $[t]^{-1}$ is:

$$\Gamma_o = \frac{\Gamma_o^\tau}{t_c} = 2\nu A_2 \Gamma_o^\tau, \quad (\text{E.91})$$

For Neumann or Dirichlet boundary conditions the nucleation of the domain walls begin at the boundaries of each segment (two degeneracy), but for periodic kinks the nucleation can take place anywhere along the segment leading to the appearance of a *soft collective mode* with zero eigenvalue [196, 202]. The removal of the zero eigenvalue do to the infinite degeneracy can be done following the McKane-Tarlie regularization procedure [208, 210].

In what follows we let $m_p = p$ to simplify notation, where m_p is defined by expression Eq. (E.85). The rate factor Γ_o per unit length in the presence of periodic boundary conditions with $2\hat{l} > 2\pi$ is given by [201, 203]:

$$\begin{aligned} \frac{\Gamma_o}{2\hat{l}} &\sim 2\nu A_2 \left| 1 - \frac{2}{p+1} \sqrt{p^2 - p + 1} \right| \dots \\ &\dots \sqrt{\frac{2p(1-p) \sinh^2(\hat{l}\sqrt{2})}{(1+p)^{5/2} \left| K(p) - \frac{1+p}{1-p} E(p) \right|} \left[\frac{\hat{T}^{-1/2}}{(2\pi)^{3/2}} \right]}, \end{aligned} \quad (\text{E.92})$$

and it has units of t^{-1} since we account for the physical units of the system. The dimensionless variables $\hat{l} = l/s_c$ and $\hat{T} = k_b T / T_c$ are defined trough the relations provided in Eq. (E.80).

The parameters A_2, A_4 and l for this system can be obtained through the methods described in section 6.2. Also K and $m_p = p$ are known once l is defined through expressions Eq. (E.51) and Eq. (E.85) respectively. The only unknown parameter in Eq. (E.92) is the kinetic coefficient ν .

E.5.7 Appendix: Chain segments with Neumann B.Cs

Following the work in Maier and Stein [196], next we present the results for a chain in which each segment of length \hat{l} is subjected to Neumann B.Cs :

$$\left. \frac{d\epsilon}{dx} \right|_{x=0} = \left. \frac{d\epsilon}{dx} \right|_{x=l} = 0, \quad (\text{E.93})$$

Note that in here the boundary conditions are specific for each segment \hat{l} , rather than two segments as done in the main text in the case of periodic B.C.s

For $\hat{l} > \pi$, the non-uniform solution to Eq. (E.82) is given by [196]:

$$\epsilon_N(x) = \sqrt{\frac{2m}{m+1}} \operatorname{sn} \left(\frac{x}{\sqrt{m+1}} + K(m) \middle| m \right) \quad (\text{E.94})$$

Similarly, for the non-constant solution Eq. (E.94) to satisfy the boundary conditions, $\hat{l}/\sqrt{m+1}$ must be an integer multiple of a full period [196]:

$$\hat{l} = 2\sqrt{m_N + 1}K(m_N). \quad (\text{E.95})$$

The activation barrier for $\hat{l} > \pi$ is [196]:

$$\frac{\hat{E}_{DP}}{T_c} = \frac{1}{3(1+m_N)^{3/2}} [4(1+m_N)E(m_N) \dots \quad (\text{E.96})$$

$$\dots - 2^{-1}(1-m_N)(5+3m_N)K_N(m_N)]. \quad (\text{E.97})$$

In the case where $\hat{l} \rightarrow \infty$ then $m_N \rightarrow 1$, $E(m_p) \rightarrow 1$, while $K(m_p)$ diverges to infinity in a logarithmic fashion [208]:

$$(1-m \rightarrow 16 \exp(-2K(m_p))).$$

Therefore in the $\hat{l} \rightarrow \infty$ limit:

$$\hat{E}_{DP} = \frac{2}{3} \left[\frac{A_2^3 K}{A_4^2} \right]^{1/2}, \quad (\text{E.98})$$

which as expected is the energy required to create a single domain wall Eq. (E.64) with boundary conditions at infinity. For $\hat{l} > \sim 5$ the \hat{E}_{DP} can be approximated using expression Eq. (E.98) [196]. Finally the rate factor for $\hat{l} > \pi$ in units of t^{-1} , is given by [196]:

$$\Gamma_o \sim \frac{2A_2\nu}{\pi} \left| 1 - \frac{2}{m_N+1} \sqrt{m_N^2 - m_N + 1} \right| \dots \quad (\text{E.99})$$

$$\dots \sqrt{\frac{\sinh(l\sqrt{2}/s_c)}{\sqrt{2} |(1-m_N)K(m_N) - (1+m_N)E(m_N)|}},$$

where ν is the kinetic coefficient and s_c is given in Eq. (E.80) in the main text.

E.6 Short dsDNA constructs

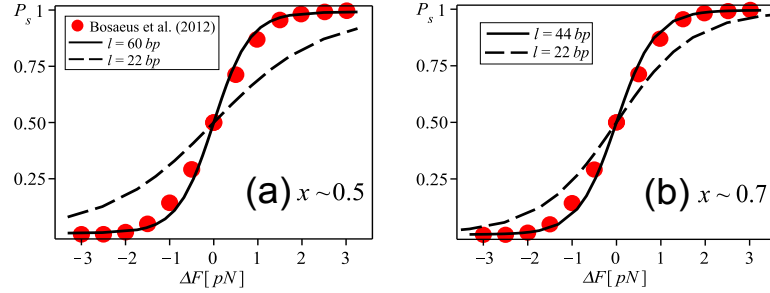


Figure E.9: Comparison of our theoretical predictions (lines) for the fraction of segments in the overstretched state P_s and Bosaeus et al. [36] experimental results (red markers). (a) Theoretical fit using $x \approx 0.5$ ($A_2 = 60\text{pN}$). (b) Theoretical fit using $x \approx 0.7$ ($A_2 = 105\text{pN}$).

Bosaeus et al. [36, 211] measured the mechanical stretching of 60 bp dsDNA in very high salt concentrations $I = 1\text{M NaCl}$ at $T = 22\text{C}$ using different fractions of GC vs AT tracts. Although they found that for high GC content (OLIGOGC1 construct with 60% GC content) and high ionic strength the molecule undergoes a non-hysteretic B-to-S DNA transition, their findings regarding the elongation of the molecule do not agree with the elongation measurement of the B-to-S DNA using larger DNA molecules (length $\sim \mu\text{m}$ [30, 32, 34, 35]). The experiments in [36] were performed by attaching the short dsDNA ~ 60 bp pieces to single-stranded DNA handles, and the construct was pulled controlling the distance and measuring the force. At a critical force region the OLIGOGC1 construct underwent a fast dynamic transition measured by a sudden decrease in the force (see Fig. 2 in [36]). The initial and final elongation (linear regime between force and distance) of the OLIGOGC1 construct is due to the extension of both the dsDNA and mainly the ssDNA handles. Using statistical tools, the authors measured the conversion dynamics between the B-DNA and overstretched form where they saw a Gaussian bimodal distribution, which supports the use of two-state models. The authors presented the probability $P_s = P_{stretched}$ of finding the molecule in the overstretched state as a function of the force, and they fitted their data using the formula of a maximum cooperativity two-state model $P_s = (1 + \exp(\beta(F_c - F)\delta z))^{-1}$. From their methods they found that the midpoint of the transition is at $F_c \approx 63.75\text{pN}$ and a change in extension $\Delta z = \delta z N \approx 10.4 \pm 0.46$ nm. Based on this result and assuming that the entire piece $N = 60\text{bp}$ underwent the phase change the authors reported $x = 0.5$ times elongation per basepair in the B-to-S transition ³:

$$x \approx \frac{\bar{L} - L}{L} = \frac{\Delta z}{0.34N}. \quad (\text{E.100})$$

Using equation Eq. (6.39) in the main text we can compute the fraction P_s as a function of $x = |2u_o|$ where u_o is the stable minima of the system at $F = F_c$ (see Eq. (6.3) in the main text). As described in the methods of section 6.1.4, $A_4 \approx 500\text{pN}$ provides a good fit to all experimental curves analyzed, hence x is basically a function of A_2 . Next, we fit the value of l to the experimental data for the 60 bp OLIGOGC1 construct from [36]. In Fig E.9(a) we show the fit when using $x \approx 0.5$ as reported in [36] and $l = 60\text{bp}$, which agrees very well with experiment. This result suggests that it is in fact possible that the entire dsDNA construct transitions from B-to-S in one single passage, and our model is able to capture such phenomena.

³The extension during the transition can be approximate by $z = z_B(1 + xP_s)$, where $z_B \approx 0.34\text{nm}$ is the B-DNA basepair length, $(x = z_S/z_B - 1)$ is the elongation per basepair and z_S is the basepair length in the overstretched state.

Under similar experimental conditions at very high salt concentration ($I \in [0.5, 1.5]\text{M}$), where one expects to see only the B-to-S transition [30, 34], the experiments measured $x \approx 0.7$. Next we try to explain this apparent disagreement. Therefore, we decided to fit Bosaeus et al. [36] data using $x \approx 0.7$ and the result is shown in Fig E.9(b) where the fit yields $l = 44\text{bp}$. Using $x = 0.7$ and assuming that only a fraction of the basepairs underwent the transition ($N = 44\text{bp}$), the predicted value of $\Delta z \approx 10\text{nm}$ is in agreement with Bosaeus et al. [36] measurements.

Although there is a possibility that a combination of factors such as basepair sequence content, composition of the preparation, temperature, ionic strength and mechanical pulling protocols can lead to a difference in the S-DNA length per basepair (x), we consider that it is more likely that in fact $x \approx 0.7$ and $N = 44\text{bp}$ in Bosaeus et al. [36] OLIGO GC experiments based on the following rationale:

(i) The phase transition seen as the dynamic force jump is a consequence of the interface between ssDNA and B-DNA of the entire molecule used in the experiment. The phase transition between two possible states (Configuration 1 and Configuration 2) takes place when the energy of the system in Configuration 2 is equal/lower to the energy of Configuration 1. In this case Configuration 1 corresponds to the system including the ssDNA handles and the energetic cost of their interfaces, not only the double stranded B-DNA. Configuration 2 corresponds to the system including the ssDNA handles, the overstretched DNA and the energetic cost of their interfaces. Therefore, the onset of the transition does not take place when the energy per basepair of the B-form equals the energy per basepair of the S-form, but at a force $F > F_c$ where it is likely that the energy per base-pair of the S-form is already lower than the energy of the B-form, leading to an abrupt decrease in F and an instantaneous transition of $\sim 44\text{bp}$.

(ii) There is a difference in the Gibbs free energy $\Delta G_F = z_F \Lambda F \approx 100[\text{pNnm}]$ between the initial Configuration 1 and the final Configuration 2 because of the presence of the instantaneous jump of the force $F_{initial} - F_{final} = \Lambda F \approx 1.9[\text{pN}]$ at the fixed position $z_f \approx 52[\text{nm}]$ (see Fig 2a in [36]) where the transition takes place. At the transition, this excess mechanical energy ΔG_F is likely transformed into the energetic cost of the instantaneous nucleation of S-DNA. This instantaneous transition of B-to-S can be understood as if N basepairs will undergo a spatially homogeneous transition at the same time. Under such an scenario the energetic cost using the spatially homogeneous model described in section 6.1.1 (in the main text) is just given by :

$$E_{Transition} = \frac{A_2^2}{4A_4}(N * 0.34) = (u_o)^2 \frac{A_2}{2}(N * 0.34). \quad (\text{E.101})$$

Using the fitted variables in overstretching experiments from section 6.3 in the main text, we found that $E_{transition} = \Delta G_F \approx 100[\text{pNnm}]$ when using $x = 0.68$ and $N = 44$ which is equal to the one measured in [36], while $E_{transition} \approx 40[\text{pNnm}]$ when $x = 0.5$ and $N = 60$. Furthermore, since $E_{transition} = \Delta G_F$ is the energy required for nucleation in extension controlled experiments, as the entire length of the construct (ssDNA + dsDNA) increases the value of z_f will increase and the dynamic force jump ΓF will decrease. This will explain why when using longer DNA constructs the force jump in extension controlled experiments can not be seen.

(iii) The location of the 2 boundaries between the ssDNA and the B-form are imperfections and therefore could serve as nucleation sites for the phase change of the B-form into S-form. At each nucleation site it is likely the number of basepairs ($N/2$) that will undergo the instantaneous phase change as a single block is at most equal to the cooperativity length of the B-to-S transition ($l \approx 22\text{bp}$).

(iv) Bosaeus [211] also performed experiments using ATGC 120bp and GCGC dimers, where two 60bp constructs corresponding to GC-rich sequences or AT-rich sequences were covalently bonded. For the ATGC dimer at 1M NaCl, there are two transition jumps (see Fig 34 in [211]). As pointed out by the author, the first jump corresponds to transition of the AT-rich sequence (non-hysteretic melting), and the second transition to the GC-rich sequence (S-DNA transition). The first dynamic

jump is non-hysteretic but its behavior shows clear signs of melting (interaction with glycolax) [211]. The second transition of the ATGC corresponds to the GC-rich sequence and it had the same characteristics as when the single OLIGOGC sequence was used. The GCGC dimer in 1M NaCl shows a reversible overstretching transition with signs of at least one intermediate state [211], where two very close consecutive dynamic jumps take place (see fig 35 in [211]). For the ATGC dimer it is clear why two dynamic jumps are visible, as each segment is essentially a different material. But for the GCGC dimer this is not the case and an analysis based solely on basepair sequencing would have to yield a single jump. We think that the covalent bond between the GC-sequences serves as another nucleation site (used by both GC-rich sequences independently), which leads to 2 consecutive jumps of approximately $N \sim 45$ bp each. Therefore, the presence of two jumps for the GCGC dimer reinforces point (ii) explained above.

(v) Finally, in Bosaeus et al. [36, 211] extension controlled experiments of GC-rich sequences at 1 M NaCl, the system is brought to a specific point where the force is jumping dynamically at much faster time scale (see Fig 2D in [36]) than what it takes an S-DNA domain to propagate (~ 1 s [35, 171]), therefore the system can not reach equilibrium. As the molecule is further extended past the point of the abrupt jump, it is possible that the S-DNA domains propagate into the remaining 16bp accompanied with a change in force of a few pN as it is observed in the non-hysteretic melting transition of the 3'5'AT in [211]. Using an 3'5'AT-rich sequence of 60 bp with closed ends (which prevented peeling), Bosaeus [211] found that the 3'5'AT undergoes a melting transition that propagates with force.

Further experiments could in fact help verify if in fact $N=44$ bp underwent the dynamic jump for the OLIGOGC sequence. According to our reasoning, under the same experimental conditions, using a single shorter (not a dimer) GC-rich sequence 45 – 50bp or slightly longer 70 – 80bp, should yield a similar dynamic behavior with the same $\Delta z \approx 10$ nm jump.

References

1. Stigter, D., 1975. The charged colloidal cylinder with a Gouy double layer. *J. Colloid Interface Sci.* 53:296.
2. Stigter, D., 1977. Interactions of Highly Charged Colloidal Cylinders with Applications to Double-Stranded DNA. *Biopolymers.* 16:1435–1448.
3. Schellman, J., and D. Stigter, 1977. Electrical Double Layer, Zeta Potential, and Electrophoretic Charge of Double-Stranded DNA. *Biopolymers.* 16:1415–1434.
4. Neukirch, S., and J. Marko, 2011. Analytical Description of Extension, Torque, and Supercoiling Radius of a Stretched Twisted DNA. *Phys. Rev. Lett.* 106(13):138104.
5. Maffeo, C., S. Robert, H. Brutzer, S. René, A. Aleksei, W. Gero, and R. Seidel, 2010. DNA-DNA Interactions in Tight Supercoils Are Described by a Small Effective Charge Density. *Phys. Rev. Lett.* 105(15):158101.
6. Vladescu, I., M. McCauley, M. Nunez, I. Rouzina, and M. Williams, 2007. Quantifying force-dependent and zero-force DNA intercalation by single molecule stretching. *Nature* 4:517–522.
7. Rouzina, I., and V. Bloomfield, 2001. Force-Induced Melting of the DNA Double Helix 1. Thermodynamic Analysis. *Biophys. J.* 80:882–893.
8. Williams, M., I. Rouzina, and V. Bloomfield, 2002. Thermodynamics of DNA Interactions from Single Molecule Stretching Experiments. *Acc. Chem. Res.* 106:159–166.
9. Nelson, D., and M. Cox, 2010. Lehninger Principles of Biochemistry, W.H. Freeman and Company, New York, chapter 24.2: DNA Supercoiling, 930–938. 4th edition.
10. Brutzer, H., N. Luzzietti, D. Klaue, and R. Seidel, 2010. Energetics at the DNA Supercoiling Transition. *Biophys. J.* 98(7):1267–1276.
11. Mosconi, F., J. Allemand, D. Bensimon, and V. Croquette, 2009. Measurement of the Torque on a Single Stretched and Twisted DNA using Magnetic Tweezers. *Physical Review Letter* 102(7):078301.
12. Forth, S., C. Deufel, M. Y. Sheinin, B. Daniels, J. P. Sethna, and D. Wang, 2008. Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. *Phys. Rev. Lett.* 100(14):148301.
13. Todd, B., V. Parsegian, A. Shirahata, T. Thomas, and D. Rau, 2008. Attractive Forces between Cation Condensed DNA Double Helices. *Biophys. J.* 95:4775–4782.
14. Raspaud, E., D. Durand, and F. Livolant, 2005. Interhelical Spacing in Liquid Crystalline Spermine and Spermidine-DNA Precipitates. *Biophys. J.* 88:392–403.

15. Besteman, K., S. Hage, N. Dekker, and S. Lemay, 2007. Role of Tension and Twist in Single-Molecule DNA Condensation. *Phys. Rev. Lett.* 98(5):058103–.
16. Besteman, K., 2006. Charge inversion and DNA condensation by multivalent ions. Ph.D. thesis, Technische Universiteit Delft.
17. Charvin, G., A. Vologodskii, D. Bensimon, and V. Croquette, 2005. Braiding DNA: Experiments, Simulations, and Models. *Biophysical Journal* 88:4124–4136.
18. Noom, M., B. van den Broek, J. van Mameren, and G. Wuite, 2007. Visualizing single DNA-bound proteins using DNA as a scanning probe. *Nature Methods* 4:1031–.
19. Stasiak, A., and J. Maddocks, 2000. Mathematics. Best packing in proteins and DNA. *Nature* 406(6793):251–253.
20. Ubbink, J., and T. Odijk, 1999. Electrostatic-undulatory theory of plectonemically supercoiled DNA. *Biophys. J.* 76(5):2502–2519.
21. Argudo, D., and P. K. Purohit, 2012. The dependence of DNA supercoiling on solution electrostatics. *Acta Biomater.* 8:2133–2143.
22. Argudo, D., and P. K. Purohit, 2012. Competition between supercoils and toroids in single molecule DNA condensation. *Biophys. J.* 103(1):118–128.
23. Abels, J., F. Moreno-Herrero, C. van der Heijden, and N. Dekker, 2005. Single-Molecule Measurements of the Persistence Length of Double-Stranded RNA. *Biophys. J.* 88:2737–2744.
24. Herrero-Galan, E., M. Fuentes-Perez, C. Carrasco, J. Valpuesta, J. Carrascosa, F. Moreno-Herrero, and J. Arias-Gonzalez, 2013. Mechanical identities of RNA and DNA double helices unveiled at the single-molecule level. *J. Am. Chem. Soc.* 135:122–131.
25. Faustino, I., A. Perez, and M. Orozco, 2010. Toward a consensus view of duplex RNA flexibility. *Biophys. J.* 99:1876–1885.
26. Sheinin, M. Y., and D. Wang, 2009. Twist-stretch coupling and phase transition during DNA supercoiling. *Phys. Chem. Chem. Phys.* 11:4800–4803.
27. Lipfert, J., S. Klijnhout, and N. Dekker, 2010. Torsional sensing of small-molecule binding using magnetic tweezers. *Nucleic Acids Res.* 38(20):7122–7132.
28. Celedon, A., D. Wirtz, and S. Sun, 2010. Torsional Mechanics of DNA are regulated by small-molecule intercalation. *J. Phys. Chem.* 114:16929–16935.
29. Sheinin, M., S. Forth, J. Marko, and M. Wang, 2011. Underwound DNA under Tension: structure, elasticity and sequence-dependent behaviors. *Phys. Rev. Lett.* 107:108102–.
30. Zhang, X., H. Chen, H. Fu, P. Doyle, and J. Yan, 2012. Two distinct overstretched DNA structures revealed by single-molecule thermodynamics measurements. *Proc. Nat. Acad. Sci.* 109:8103–8108.
31. Oberstrass, F., L. Fernandes, and Z. Bryant, 2012. Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. *Proc. Natl. Acad. Sci.* 109:6106–6111.
32. Zhang, X., H. Chen, S. Le, I. Rouzina, P. Doyle, and J. Yan, 2013. Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry. *Proc. Nat. Acad. Sci.* 110:3865–3870.

33. King, G., P. Gross, U. Bockelmann, M. Modesti, G. Wuite, and E. Peterman, 2013. Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy. *Proc. Nat. Acad. Sci.* 110:3865–3870.
34. Bongini, L., L. Melli, V. Lombardi, and P. Bianco, 2013. Transient kinetics measured with force steps discriminate between double-stranded DNA elongation and melting and define the reaction energetics. *Nucl. Acids. Res.* gkt1297.
35. Bianco, P., L. Bongini, L. Melli, M. Dolfi, and V. Lombardi, 2011. PicoNewton-Millisecond Force Steps Reveal the Transition Kinetics and Mechanism of the Double-Stranded DNA Elongation. *Biophys. J.* 101:866–874.
36. Bosaeus, N., A. El-Sagheer, T. Brown, S. Smith, B. Akerman, C. Bustamante, and B. Norden, 2012. Tension induces a base-paired overstretched DNA conformation. *Proc. Nat. Acad. Sci.* 109:15179–15184.
37. van den Broek, B., M. Noom, J. van Mameren, C. Battle, F. MacKintosh, and G. Wuite, 2010. Visualizing the Formation and Collapse of DNA Toroids. *Biophys. J.* 98:1902–1910.
38. Williams, M., I. Rouzina, and M. McCauley, 2009. Peeling back the mystery of DNA overstretching. *Proc. Nat. Acad. Sci.* 106:18047–18048.
39. Nelson, P., 2004. Biological Physics. W. H. Freeman and Company, New York.
40. Moroz, J., and P. Nelson, 1997. Torsional directed walks, entropic elasticity, and DNA twist stiffness. *Proc. Natl. Acad. Sci.* 94(26):14418–14422.
41. Strick, T., J. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, 1998. Behavior of Supercoiled DNA. *Biophys. J.* 74:2016–2028.
42. Swigon, D., 2009. The Mathematics of DNA Structure, Mechanics and Dynamics. *IMA Volumes in Mathematics and Its Applications* 150:293–320.
43. Bustamante, C., Z. Bryant, and S. Smith, 2003. Ten years of tension: single-molecule DNA mechanics. *Nature* 421:423–427.
44. Strick, T., J. Allemand, V. Croquette, and D. Bensimon, 2000. Twisting and stretching single DNA molecules. *Prog. Biophys. Mol. Biol.* 74(1-2):115–140.
45. Clauvelin, N., B. Audoly, and S. Neukirch, 2008. Mechanical Response of Plectonemic DNA: An analytical solution. *Macromolecules* 41(12):4479–4483.
46. Shao, Q., S. Goyal, L. Finzi, and D. Dunlap, 2012. Physiological level of salt and Polyamines favor writhe and limit twist in DNA. *Macromolecules* 45:3188–3196.
47. Chouaieb, N., and J. Maddocks, 2004. Kirchhoff’s problem of helical equilibria of uniform rods. *J. Elasticity* 77:221–247.
48. Marko, J., and E. Siggia, 1995. Stretching DNA. *Macromolecules* 28:8759–8770.
49. Su, T., and P. K. Purohit, 2010. Thermodynamics of a heterogeneous fluctuating chain. *J. Mech. Phys. Solids.* 58:164–186.
50. Smith, S., L. Finzi, and C. Bustamante, 1992. Direct mechanical measurements of the elasticity of single DNA-molecules by using magnetic beads. *Science* 258(5085):1122–1126.
51. Su, T., and P. K. Purohit, 2009. Mechanics of force unfolding of proteins. *Acta Biomaterialia* 5:1855–1863.

52. Marko, J., and E. Siggia, 1995. Statistical-mechanics of supercoiled DNA. *Phys. Rev. E* 52(3):2912–2938.
53. Neukirch, S., and E. L. Starostin, 2008. Writhe formulas and antipodal points in plectonemic DNA configurations. *Phys. Rev. E* 78(4):041912.
54. Andelman, D., 2006. Introduction to electrostatics in soft and biological matter, Taylor and Francis, Boca Raton, FL, 97–122. Scottish Graduate Series.
55. Marko, J., 2007. Torque and dynamics of linking number relaxation in stretched supercoiled DNA. *Phys. Rev. E* 76(2):021926.
56. Moroz, J., and P. Nelson, 1998. Entropic elasticity of twist-storing polymers. *Macromolecules* 31(18):6333–6347.
57. Bryant, Z., M. Stone, J. Gore, S. B. Smith, N. Cozzarelli, and C. Bustamante, 2003. Structural transitions and elasticity from torque measurements on DNA. *Nature* 424:338–341.
58. Argudo, D., and P. K. Purohit, 2013. DNA superhelical structures with non-constant helical pitch. *Math. Mech. Solids*. 18(6):649–673.
59. Argudo, D., and P. K. Purohit, 2013. Torsion of DNA modeled as a heterogeneous fluctuating rod. *J. Mech. Phys. Solids*. 228–256.
60. Argudo, D., and P. K. Purohit, 2014 (submitted). Equilibrium and kinetics of DNA overstretching modeled using a quartic energy landscape. *Biophys. J.* .
61. Koster, D., V. Croquette, C. Dekker, S. Shuman, and N. Dekker, 2005. Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. *Nature* 434:671–674.
62. Koster, D., A. Crut, S. Shuman, M. Bjornsti, and N. Dekker, 2010. Cellular Strategies for Regulating DNA Supercoiling: A Single-Molecule Perspective. *Cell* 142(4):519–530.
63. Strick, T., J. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, 1996. The elasticity of a single supercoiled DNA molecule. *Science* 271(5257):1835–1837.
64. Strick, T., M. N. Dessinges, G. Charvin, N. Dekker, J.-F. Allemand, D. Bensimon, and V. Croquette, 2003. Stretching of macromolecules and proteins. *Rep. Prog. Phys* 66(1):1–45.
65. Fraser, W., and D. Stump, 1998. The equilibrium of the convergence point in two-strand yarn plying. *Int. J. Solid. Struct.* 35(3-4):295–298.
66. Coleman, B., and D. Swigon, 2000. Theory of Supercoiled Elastic Rings with Self-Contact and Its Application to DNA Plasmids. *J. Elasticity* 60:173 – 221.
67. Thompson, J., G. van der Heijden, and S. Neukirch, 2002. Supercoiling of DNA plasmids: mechanics of the generalized ply. *Proc. R. Soc. Lond. A* 458(2020):959–985.
68. Goyal, S., N. Perkins, and C. Lee, 2005. Nonlinear dynamics and loop formation in Kirchhoff rods with implications to the mechanics of DNA and cables. *J. Comput. Phys.* 209(1):371–389.
69. Goyal, S., and N. Perkins, 2008. Looping mechanics of rods and DNA with non-homogeneous and discontinuous stiffness. *International Journal of Non-Linear Mechanics* 43:1121–1129.
70. van der Heijden, G., J. Thompson, and S. Neukirch, 2003. A variational approach to loaded ply structures. *J. Vib. Control* 9(1-2):175–185.

71. Purohit, P. K., 2008. Plectoneme formation in twisted fluctuating rods. *J. Mech. Phys. Solids* 56(5):1715–1729.
72. Purohit, P. K., 2010. Shape and Energetics of DNA Plectonemes. In K. Garikipati, and E. Arruda, editors, IUTAM Symposium on Cellular, Molecular and Tissue Mechanics. volume 16(3), 123–138.
73. Clauvelin, N., B. Audoly, and S. Neukirch, 2009. Elasticity and Electrostatics of Plectonemic DNA. *Biophys. J.* 96(9):3716–3723.
74. Daniels, B., S. Forth, M. Y. Sheinin, D. Wang, and J. P. Sethna, 2009. Discontinuities at the DNA supercoiling transition. *Phys. Rev. E.* 80:040901(R).
75. Chouaieb, N., A. Goriely, and J. Maddocks, 2006. Helices. *Proc. Natl. Acad. Sci.* 103(25):9398–9403.
76. van der Heijden, G., 2001. The static deformation of a twisted elastic rod constrained to lie on a cylinder. *Proc. R. Soc. Lond. A* 457(2007):695–715.
77. Pirański, P., 1998. In search of ideal knots, World Scientific, Singapore, volume 19 of *Series on Knots and everything*, 20–41.
78. Coyne, J., 1990. Analysis of the Formation and Elimination of loops in Twisted Cable. *IEEE J. Oceanic Eng.* 15:72–83.
79. Kúlic, I. M., H. Mohrbach, R. Thaokar, and H. Schiessel, 2007. Equation of state of looped DNA. *Phys. Rev. E.* 75(1):011913.
80. Agrawal, N. J., R. Radhakrishnan, and P. Purohit, 2008. Geometry of mediating protein affects the probability of loop formation in DNA. *Biophys. J.* 94:3150–3158.
81. Kornyshev, A. A., D. J. Lee, S. Leikin, and A. Wynveen, 2007. Structure and interactions of biological helices. *Rev. Mod. Phys.* 79(3):943–996.
82. Brenner, S., and V. Parsegian, 1974. A physical method for deriving the electrostatic interaction between rod-like polyelectrolytes at all mutual angles. *Biophys. J.* 14(4):327–334.
83. Podgornik, R., and V. Parsegian, 1990. Molecular fluctuations in the packing of polymeric liquid crystals. *Macromolecules* 23(8):2265–2269.
84. van der Maarel, J. R. C., 2008. Introduction to Biopolymer Physics. World Scientific, Singapore.
85. Zhang, H., and J. Marko, 2008. Maxwell relations for single-DNA experiments: Monitoring protein binding and double-helix torque with force-extension measurements. *Phys. Rev. E* 77(3 Pt 1):031916.
86. Vologodskii, A., and N. Cozzarelli, 1995. Modeling of long-range electrostatic interactions in DNA. *Biopolymers* 35(3):289–296.
87. Rudoi, Y., and A. Sukhanov, 2000. Thermodynamic fluctuation within the Gibbs and Einstein approaches. *Usp. Fiz. Nauk.* 43(12):1169–1199.
88. Landau, L. D., and E. Lifshitz, 1980. Statistical Physics. Pergamon Press, New York, 3rd edition. Pt. 1.
89. Burak, Y., G. Ariel, and D. Andelman, 2004. Competition between condensation of monovalent and multivalent ions in DNA aggregation. *Curr. Opin. Colloid Interface Sci.* 9:53–58.

90. Besteman, K., K. van Eijk, and S. Lemay, 2007. Charge inversion accompanies DNA condensation by multivalent ions. *Nature Physics* 3:641–644.
91. Savelyev, A., and G. Papoian, 2007. Inter-DNA Electrostatics from Explicit Solvent Molecular Dynamics Simulations. *J. Am. Chem. Soc.* 129:6060–6061.
92. Sottas, P., E. Larquet, A. Stasiak, and J. Dubochet, 1999. Brownian Dynamics Simulation of DNA Condensation. *Biophys. J.* 77:1858–1870.
93. Luan, B., and A. Aksimentiev, 2010. Electric and electrophoretic inversion of the DNA charge in multivalent electrolytes. *Soft Matter* 6:243–246.
94. Laughton, C. A., and S. A. Harris, 2011. Atomistic simulation of DNA. *Atomistic simulation of DNA* 1:6590–600.
95. Noy, A., and R. Golestanian, 2010. The Chirality of DNA: Elasticity Cross-Terms at Base-Pair Level Including A-tracts and the Influence of Ionic Strength. *J. Phys. Chem. B.* 114:8022–8031.
96. Baumann, C. G., S. B. Smith, V. A. Bloomfield, and C. Bustamante, 1997. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. USA* 94:6185–6190.
97. Brewer, L., 2010. Deciphering the structure of DNA toroids. *Intgr. Biol.* published online doi:10.1039/c0ib00128g.
98. Battle, C., B. van der Broek, M. Noom, J. van Mameren, G. Wuite, and F. MacKintosh, 2009. Unraveling DNA tori under tension. *Phys. Rev. E.* 80:031917.
99. Bloomfield, V., 1997. DNA condensation by multivalent cations. *Biopolymers* 44:269–282.
100. Rau, D., and V. Parsegian, 1992. Direct measurements of the intermolecular forces between counterion-condensed DNA double helices. Evidence for long-range attractive hydration forces. *Biophys. J.* 61(1):246–259.
101. Murayama, Y., Y. Sakamaki, and S. M., 2003. Elastic response of single DNA molecules exhibits a reentrant collapsing transition. *Phys. Rev. Lett.* 90:018102.
102. Thomas, T., and V. Bloomfield, 1983. Chain flexibility and hydrodynamics of the B and Z forms of poly(dG-dC).poly(dG-dC). *Nucleic Acid Res.* 11:1919–1930.
103. DeRouchey, J., and D. Rau, 2011. Role of amino acid insertions on intermolecular forces between arginine peptide condensed DNA helices: implications for protamine-DNA packaging in sperm. *J. Biol. Chem.* 286(49):41985–41992.
104. Shen, M., K. Downing, R. Balhorn, and N. Hud, 2000. Controlling the size of nanoscale toroidal DNA condensates with static curvature and ionic strength. *J. Am. Chem. Soc.* 122:4833–4834.
105. Hud, N., K. Downing, and R. Balhorn, 1995. A constant radius of curvature model for the organization of DNA in toroidal condensates. *Proc. Natl. Acad. Sci.* 92:3581–3585.
106. Conwell, C., I. Vilfan, and N. Hud, 2003. Controlling the size of nanoscale toroidal DNA condensates with static curvature and ionic strength. *Proc. Natl. Acad. Sci.* 100:9296–9301.
107. Hud, N., and V. I., 2005. Toroidal DNA condensates: unravelling the fine structure and the role of nucleation in determining size. *Annu. Rev. Biophys. Biomol. Struct.* 34:295–318.
108. Baumann, C., V. Bloomfield, S. Smith, B. C., M. Wang, and M. Block, S., 2000. Stretching of Single Collapsed DNA Molecules. *Biophys. J.* 78:1965–1978.

109. Fu, W., X. Wang, X. Zhang, S. Ran, J. Yan, and M. Li, 2006. Compaction Dynamics of Single DNA Molecules under Tension. *J. Am. Chem. Soc.* 128:15040–15041.
110. Fu, W., H. Chen, M. Li, and J. Yan, 2010. Kinetics of Single DNA Compaction by Hexamine Cobalt Chloride. *J. Comput. Theor. Nanosci.* 7:213–217.
111. Grosberg, A., and A. Zhestkov, 1986. On the compact form of linear duplex DNA - Globular states of the uniform elastic (persistent) macromolecule. *J. Biomol. Struct. Dyn.* 3:859–872.
112. Mamasakhlisov, Y., B. Todd, A. Badasyan, A. Mkrtchyan, V. Morozov, and V. Parsegian, 2009. DNA stretching and multivalent-cation induced condensation. *Phys. Rev. E* 80:031915.
113. Fuller, F., 1971. Writhing number of a space curve. *Proc. Natl. Acad. Sci.* 68(4):815–819.
114. Purohit, P., M. Inamdar, P. Grayson, T. Squires, J. Kondev, and R. Phillips, 2005. Forces during Bacteriophage DNA Packaging and Ejection. *Biophys. J.* 88(2):851–866.
115. Purohit, P., J. Kondev, and R. Phillips, 2003. Mechanics of DNA Packaging in Viruses. *Proc. Natl. Acad. Sci.* 100(6):3173–3178.
116. Cherstvy, A., 2005. Structure of DNA toroids and electrostatic attraction of DNA duplexes. *J. Phys. Condens. Matter.* 17:1363.
117. Park, S., D. Harries, and M. Gelbart, 1998. Topological Defects and the Optimum Size of DNA Condensates. *Biophys. J.* 75:714–720.
118. Cárdenas-Lizana, P., and P. Hsiao, 2009. Stick-Release Pattern in Stretching Single Condensed Polyelectrolyte Toroids. *Macromolecules* 42:3211–3214.
119. Rouzina, I., and V. Bloomfield, 1998. DNA Bending by Small, Mobile Multivalent Cations. *Biophys. J.* 74:3152–3164.
120. Ubbink, J., and T. Odijk, 1996. Deformation of toroidal DNA condensates under surface stress. *Europhys. Lett.* 33(5):353–358.
121. Ubbink, J., and T. Odijk, 1995. Polymer- and Salt-Induced toroids of Hexagonal DNA. *Biophys. J.* 68:54–61.
122. Schnurr, B., F. Gittes, and F. MacKintosh, 2002. Metastable intermediates in the condensation of semiflexible polymers. *Phys. Rev. E.* 65:061904.
123. Montesi, A., M. Pasquali, and F. MacKintosh, 2004. Collapse of a semiflexible polymer in poor solvent. *Phys. Rev. E.* 69:021916.
124. Pereira, G., and D. R. M. Williams, 2001. Toroidal Condensates of Semiflexible Polymers in Poor Solvents: Adsorption, Stretching, and Compression. *Biophys. J.* 80:161–168.
125. Marko, J., 1997. Supercoiled and braided DNA under tension. *Physical Review Letters E* 55:1758–1772.
126. Coleman, B., and D. Swigon, 2002. Theory of Self-contact in DNA molecules Modeled as Elastic Rods. *In Nuovi progressi nella fisica matematica dall’eredit di Dario Graffi.* volume 177, 281–295.
127. Gelfand, I., and S. Fomin, 2000. *Calculus of Variations.* Dover Publications, Mineola, New York.
128. Bolza, O., 1961. *Lectures on the Calculus of Variations.* Dover Publications, New York.

129. Hunt, N., and J. Hearst, 1991. Elastic model of DNA supercoiling in the infinite-length limit. *J. Chem. Phys.* 95(12):9329–9336.
130. Purohit, P. K., M. Arsenault, Y. Goldman, and H. Bau, 2008. The mechanics of short rod-like molecules in tension. *Int. J. Nonlin. Mech.* 43:1056–1063.
131. Nizette, M., and A. Goriely, 1999. Towards a classification of Euler-Kirchhoff filaments. *J. Math. Phys.* 40(6):2830–2866.
132. Enneper, A., and F. Muller, 1890. Elliptische Functionen. Theorie und Geschichte. Halle a.S.
133. Greenhill, A., 1892. The Application of Elliptic Functions. Macmillan and Co, New York.
134. Marko, J., 1998. DNA under high tension: Overstretching, underwinding, and relaxation dynamics. *Phys. Rev. E.* 57:2134–2149.
135. Manning, R., J. Maddocks, and J. Kahn, 1996. A continuum rod model of sequence-dependent DNA structure. *J. Chem. Phys.* 105:5626–.
136. Frank-Kamenetskiia, M., A. Lukashina, V. Anshelevicha, and A. Vologoskii, 1985. Torsional and Bending Rigidity of the Double Helix from Data on Small DNA Rings. *J. Biomed. Struct. Dyn.* 2:1005–.
137. Coleman, B., W. Olson, and D. Swigon, 2003. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* 118:7127–.
138. Olson, W., and V. Zhurkin, 2000. Modeling DNA deformations. *Curr. Opin. Struct. Biol.* 10:286–297.
139. Gonzalez, O., D. Petkeviciute, and J. Maddocks, 2013. A sequence-dependent rigid-base model of DNA. *J. Chem. Phys.* 138:055102–.
140. Moakher, M., and J. Maddocks, 2005. A Double-Strand Elastic Rod Theory. *Arch. Rational Mech. Anal.* 177:53–91.
141. Wolfe, K., W. Hastings, S. Dutta, A. Long, B. Shapiro, T. Woolf, M. Guthold, and G. Chirikjian, 2012. Multiscale Modeling of Double-Helical DNA and RNA: A Unification through Lie Groups. *J. Phys. Chem. B* 116:8556–.
142. Travers, A., 2004. The structural basis of DNA flexibility. *Phil. Trans. R. Soc. Lond. A* 362:1390–.
143. Flory, P., 1976. Statistical thermodynamics of random networks. *Proc. R. Soc. Lond. A.* 351:351–380.
144. Zhang, Y., and D. Crothers, 2003. Statistical Mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* 84:136–153.
145. Lee, M., S. Kim, and S. Hong, 2010. Minute negative superhelicity is sufficient to induce the B-Z transition in the presence of low tension. *Proc. Natl. Acad. Sci.* 107:4985.
146. Nelson, P., 2008. Biological Physics: Energy, Information, Life. W.H. Freeman and Company, New York.
147. Kim, J., and G. Chirikjian, 2006. Conformational analysis of stiff chiral polymers with end constraints. *Mol. Simul.* 32:1139–1154.

148. van der Heijden, G., and J. Thompson, 2000. Helical and localised buckling in twisted rods: A unified analysis of the symmetric case. *Nonlinear Dynam.* 79(1):71–99.
149. Emanuel, M., 2012. The effect of thermal fluctuations on elastic instabilities of biopolymers. Ph.D. thesis, Universiteit Leiden.
150. Lipfert, J., G. Skinner, J. Keegstra, T. Hensgens, T. Jagger, D. Dullin, and N. Dekker, 2013. Probing the Response of Double-Stranded RNA to Force and Torque at the Single-Molecule Level. *Biophys. J.* 104:1a–2a.
151. Gore, J., Z. Bryant, M. Nollmann, M. Le, N. Cozzarelli, and C. Bustamante, 2006. DNA overwinds when stretched. *Nature* 442:836–839.
152. Gross, P., N. Laurens, L. Oddershede, U. Bockelmann, E. Peterman, and G. Wuite, 2011. Quantifying how DNA stretches, melts and changes twist under tension. *Nat. Phys.* 7:731–736.
153. Atwell, S., L. Disseau, A. Stasiak, A. Stasiak, A. Renodon-Corniere, M. Takahashi, J. Viovy, and G. Capello, 2012. Probing Rad51-DNA interactions by changing DNA twist. *Nucleic Acids Res.* 10:1–8.
154. Triebel, H., H. Bar, R. Geuther, and G. Burckhardt, 1995. Netropsin-induced changes of DNA supercoiling; sedimentation studies. *Progr. Colloid. Polym. Sci.* 99:45–54.
155. Yan, J., and J. Marko, 2004. Localized Single-Stranded Bubble Mechanism for Cyclization of Short double helix DNA. *Phys. Rev. Lett.* 93:108108.
156. Storm, C., and P. Nelson, 2003. Theory of high-force DNA stretching and overstretching. *Phys. Rev. E.* 67:051906–.
157. Calderon, C., W. Chen, K. Lin, N. Harris, and C. Kiang, 2009. Quantifying DNA melting transitions using single-molecule force spectroscopy. *J. Phys. Condens. Matter.* 21:24114.
158. Fu, H., H. Chen, X. Zhang, Z. Qu, J. Marko, and J. Yan, 2011. Transition dynamics and selection of the distinct S-DNA and strand unpeeling modes of double helix overstretching. *Nucleic Acids Res.* 39:3473–3481.
159. Sarkar, A., J. Leger, D. Chatenay, and J. Marko, 2001. Structural transition in DNA driven by external force and torque. *Phys. Rev. E.* 63:051903–.
160. Wiggins, P., R. Phillip, and P. Nelson, 2005. Exact theory of kinkable elastic polymers. *Phys. Rev. E* 71:021909–.
161. Nelson, P., 2012. Spare the (Elastic) Rod. *Science* 337:1045–.
162. Neukirch, S., and J. Marko, 2013. Global force-torque phase diagram for the DNA double helix: Structural transitions, triple points, and collapsed plectonemes. *Phys. Rev. E.* 88:062722–.
163. Smith, S., Y. Cui, and C. Bustamante, 1996. Overstretching B-DNA: the elastic response of the individual double-stranded and single-stranded DNA molecules. *Science* 271:795–799.
164. Williams, M., J. Wenner, I. Rouzina, and V. Bloomfield, 2001. Entropy and Heat Capacity of DNA Melting from Temperature Dependence of Single Molecule Stretching. *Biophys. J.* 80:1932–1939.
165. Cluzel, P., A. Lebrun, R. Lavery, J. Viovy, F. Chatenay, and F. Caron, 1996. DNA: an extensible molecule. *Science* 271:792–794.

166. Ahsan, A., J. Rudnick, and R. Bruinsma, 1998. Elasticity theory of the B-DNA to S-DNA transition. *Biophys. J.* 74:132–137.
167. Storm, C., and P. Nelson, 2003. The bend stiffness of S-DNA. *Europhys. Lett.* 62(5):760–766.
168. Cizeau, P., , and J. Viovy, 1997. DNA: an extensible molecule. *Biopolymers* 42:383–385.
169. Fiascoaro, A., and F. Falo, 2012. Dynamical model for the full stretching curve of DNA. *Phys. Rev. E* 86:032902–.
170. Pupo, A., F. Falo, and A. Fiascoaro, 2013. DNA overstretching transition induced by melting in a dynamical mesoscopic model. *J. Chem. Phys.* 139:095101–.
171. Rouzina, I., M. Mccauley, and M. Williams, 2013. Kinetics of DNA Overstretching: Melting vs B-To-S Transition. *Biophys. J.* 104:165a–.
172. Chan, H., S. Bromberg, and K. Dill, 1995. Models of Cooperativity in Protein Folding. *Phil. Trans. R. Soc. Lond. B* 348:61–70.
173. Jackson, M., 2006. *Molecular and Cellular Biophysics*. Cambridge University, New York.
174. Hanggi, P., P. Talkner, and M. Borkovec, 1990. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* 62:251–341.
175. Tuszynski, J. A., M. Clouter, and H. Kiefte, 1985. Non-Gaussian models for critical fluctuations. *Phys. Rev. B* 33:3423–3435.
176. Odijk, T., 1995. Stiff Chains and Filaments under Tension. *Macromolecules* 28:7016–7018.
177. Roy, B. N., 2002. *Fundamentals of Classical and Statistical Thermodynamics*. Wiley, New York.
178. Ryzhik, I., and I. Gradshteyn, 1980. *Tables of Integrals, Series, and Products*. Academic Press, Inc, London.
179. Olver, F. W. J., D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors, 2010. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY.
180. van Mameren, J., P. Grossr, G. Farge, G. Hoojiman, M. Modesti, M. Falkenberg, G. Wuite, and E. Peterman, 2009. Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc. Nat. Acad. Sci.* 106:18231–18236.
181. Wenner, J., M. Williams, I. Rouzina, and V. Bloomfield, 2002. Entropy and Heat Capacity of DNA Melting from Temperature Dependence of Single Molecule Stretching. *Biophys. J.* 82:3160–3169.
182. Khokhlov, A. R., and A. Y. Grosberg, 1994. *Statistical Physics of Macromolecules*. American Institute of Physics, New York.
183. Kramer, H., 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7:284–.
184. Schurr, J., J. Delrow, B. Fujimoto, and A. Benight, 1997. The question of long-range allosteric transitions in DNA. *Biopolymers* 44(3):283–308.
185. Leger, J., J. Robert, L. Bourdieu, D. Chatenay, and J. Marko, 1998. RecA binding to a single double-stranded DNA molecule: A possible role of DNA conformational fluctuations. *Proc. Nat. Acad. Sci.* 95:12295–12299.

186. Lee, M., J. Lipfert, H. Sanchez, C. Wyman, and N. Dekker, 2013. Structural and torsional properties of the RAD51-dsDNA nucleoprotein filament. *Nucleic Acids Res.* 12:7023–7030.
187. van Loenhout, M., M. de Grunt, and C. Dekker, 2012. Dynamics of DNA Supercoils. *Science* 338:94–97.
188. Neukirch, S., G. van der Heijden, and J. Thompson, 2002. Writhing instabilities of twisted rods: from infinite to finite length. *J. Mech. Phys. Solids* 50(6):1175–1191.
189. van der Heijden, G., S. Neukirch, and J. Thompson, 2003. Instability and self-contact phenomena in the writhing of clamped rods. *Int. J. Mech. Sci.* 45(1):161–196.
190. Manning, G., 1969. Limiting laws and counterion condensation in polyelectrolyte solutions .I. Colligative properties. *J. Chem. Phys.* 51(3):954.
191. Salermo, D., A. Tempestini, I. Mai, D. Brogioli, R. Ziano, V. Cassina, and Mantegazza, 2012. Single-Molecule Study of the DNA Denaturation Phase Transition in the Force-Torsion Space. *Phys. Rev. Lett.* 109:118303.
192. Moradi, M., V. Babin, C. Roland, and S. Celeste, 2013. Reaction path ensemble of the B-Z-DNA transition: a comprehensive atomistic study. *Nucleic Acids Res.* 41:33–43.
193. Bergou, M., M. Wardetzky, S. Robinson, B. Audoly, and E. Grinspun, 2008. Discrete elastic rods. *ACM Transactions on Graphics* 27(3):63–.
194. Chaichian, M., and A. Demichev, 2001. Path Integrals in Physics Volume 1: Stochastic Processes and Quantum Mechanics. IOP, Bristol.
195. Usmani, R., 1994. Inversion of Jacobi's Tridiagonal Matrix. *COmputers Math. Applic.* 27:59–66.
196. Maier, R., and D. L. Stein, 2003. The effects of weak spatiotemporal noise on a bistable one-dimensional system. *In Noise in Complex Systems and Stochastic Dynamics.* volume 5114, 67–78.
197. Hanggi, H., 1986. Escape from a Metastable State. *Journal of Statistical Physics* 42:105–.
198. Krumhansl, J., and J. Schrieffer, 1975. Dynamics and statistical mechanics of a one-dimensional model Hamiltonian for structural phase transitions. *Phys. Rev. B* 11:3535–3545.
199. Tuszynski, J. A., M. Clouter, and H. Kiefte, 1985. Beyond the Gaussian approximation: Some new results in the Landau theory of phase transitions. *Phys. Rev.* 108A:272–276.
200. Stein, D., 2005. Large Fluctuations, Classical Activation, Quantum Tunneling, and Phase Transitions. *Braz. J. Phys.* 35:242–252.
201. Stein, D., 2004. Critical Behavior of the Kramers Escape Rate in Asymmetric Classical Field Theories. *J. Stat. Phys.* 114:1537–1556.
202. Maier, R., and D. L. Stein, 2001. Droplet Nucleation and Domain Wall Motion in a Bounded Interval. *Phys. Rev. Lett.* 87:1270601–.
203. Berglund, N., and B. Gentz, 2009. Anomalous behavior of the Kramers rate at bifurcations in classical field theories. *J. Phys. A: Math. Theor.* 42:052001–.
204. Kardar, M., 2007. Statistical physics of fields. Cambridge University Press , Cambridge.

205. Langer, J., 1991. An introduction to the kinetics of first-order phase transitions, Cambridge University Press, Cambridge, UK, 297–363. Aléa Saclay: Monographs and text in Statistical Physics 1.
206. Currie, J., J. Krumhansl, A. Bishop, and S. Trullinger, 1980. Statistical mechanics of one-dimensional solitary-wave-bearing scalar fields: Exact results and ideal-gas phenomenology. *Phys. Rev. B* 22:477–496.
207. Habib, S., and L. Grant, 2000. Dynamics of Kinks: Nucleation, Diffusion, and Annihilation. *Phys. Rev. Lett.* 84:1070–1073.
208. Kleinert, H., and H. Chervyakov, 1998. Simple Explicit Formulas for Gaussian Path Integrals with Time-Dependent Frequencies. *Phys. Lett. A.* 245:345–357.
209. Schulman, L., 1981. Techniques and Application of Path Integration. Wiley, New York.
210. Mckane, A., and M. Tarlie, 1995. Regularization of functional determinants using boundary perturbations. *J. Phys. A: Math. Gen.* 28:6931–6942.
211. Bosaeus, B., 2013. Overstretching single-molecule force spectroscopy studies. Ph.D. thesis, Chalmers University of Technology.