

APPLICATIONS OF A DOUBLE-STRANDED ELASTIC ROD MODEL TO DNA

Jaspreet Singh

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

2020

Supervisor of Dissertation

Prashant K. Purohit, Professor Mechanical Engineering and Applied Mechanics

Graduate Group Chairperson

Jennifer R. Lukes, Professor Mechanical Engineering and Applied Mechanics

Dissertation Committee

Pedro Ponte Castañeda, Raymond S. Markowitz Faculty Fellow and Professor
Mechanical Engineering and Applied Mechanics

John L. Bassani, Richard H. and S. L. Gabel Professor Mechanical Engineering
and Applied Mechanics

Yale E. Goldman, Professor of Physiology

Acknowledgments

I am indebted to my advisor Prashant K. Purohit for the time and effort he invested in mentoring me. His scholastic aptitude, work ethic, and indefatigable drive made this journey a truly gratifying experience. With his support and encouragement, I felt confident to undertake several challenging projects.

Many thanks to my dissertation committee—Pedro Ponte Castañeda, John L. Bassani, and Yale E. Goldman—who took time out from their busy schedules to read the thesis and provided valuable feedback. They are the role models I look up to as I graduate.

During my Ph.D., I had opportunities to collaborate with John L. Bassani, Paulo E. Arratia, Igor Bargatin, and Jordan R. Raney. I benefited immensely from several stimulating discussions with them over the years. Special thanks to John L. Bassani whose comments and instruction have been instrumental in shaping my own research perspectives.

I thank my parents who actively supported me throughout the journey. Many thanks to my friends and lab-mates who were by my side through this lengthy and sometimes arduous marathon.

ABSTRACT

APPLICATIONS OF A DOUBLE-STRANDED ELASTIC ROD MODEL TO DNA

Jaspreet Singh,
Prashant K. Purohit

This thesis discusses the development and applications of a double-stranded elastic rod model for DNA, henceforth called *birod* model. The model highlights the role of DNA elasticity in allosteric interactions between two ligands bound to DNA (chapters 3 and 4) and—combined with statistical mechanics—provides insights into the micromechanics of DNA melting (chapter 5).

In chapters 3 and 4, the birod model for DNA is used to compute the allosteric interaction energy between two ligands on DNA. This interaction is quantified by measuring the change in free energy as a function of the distance between the binding sites for two ligands. The trends in this interaction energy can be explained using the birod model which accounts for the helical shape of DNA, elastic deformation of strands and base-pairs, and the stacking energy due to perturbations in position and orientation of the bases caused by the binding of ligands. The model predicts that the interaction energy between two ligands decays exponentially with the distance between them and oscillates with the periodicity of the double helix, which by appropriate parameter fitting is shown to quantitatively match with the experimental measurements. Furthermore, the decaying oscillatory trend in the perturbation of groove width in a protein-DNA complex predicted by the model is consistent with the results from molecular simulations.

In chapter 2, structural transitions in DNA are studied using ideas from the Zimm-Bragg helix-coil transition theory and the theory of fluctuating elastic rods. Experimental studies on single molecules of DNA have reported several cooperative structural transitions, including coexistence of three phases, when tensile force or twisting moment is applied to the molecule. The interface energy between two phases of DNA imparts the cooperative character to the force-extension curve or torque-rotation curve observed experimentally. In chapter 5, we choose one such structural transition from dsDNA to single-stranded DNA—called DNA melting—and study it using the statistical mechanics and continuum mechanics of an elastic birod. The detailed microscopic description of the outer-strands and base-pairs admissible in the birod model enables us to decipher

why the DNA oligomers with higher GC content are stiffer. Furthermore, a nonlinear asymmetric interaction between the outer strands leads to a sudden and highly cooperative melting transition. Furthermore, the model enables us to examine the effect of tensile force on the melting temperature.

Contents

Acknowledgements	ii
Abstract	iii
Contents	v
List of Tables	vii
List of Figures	viii
1 Introduction	1
2 Structural transitions in torsionally constrained DNA and their dependence on solution electrostatics	9
2.1 Model	9
2.2 Partition function for three phase coexistence	13
2.3 Effect of electrostatic interactions	17
2.4 Discussion	24
2.5 Conclusion	26
3 Elasticity as the basis of allostery in DNA	27
3.1 Introduction	27
3.2 Theory	29
3.2.1 Contact forces in the outer strands (\mathbf{n}^\pm)	32
3.2.2 Contact moments in the outer strands (\mathbf{m}^\pm)	32
3.2.3 Force and moment transferred by the web (\mathbf{f}, \mathbf{c})	33
3.2.4 Contributions from the stacking energy (l, h)	35
3.3 Results	35
4 Allosteric interactions in a birod model of DNA	43
4.1 Strategy to compute interaction energy	43
4.2 Interaction energy for two DNA binding proteins	45
4.2.1 Step 1: Deformation of the outer strands	45
4.2.2 Step 2: Rotation of strands	47
4.2.3 Step 3: Mechanics of base-pairing	49
4.2.3.1 Bending and twisting of the web	50

4.2.3.2	Extension of the web	51
4.2.3.3	Stacking energy	52
4.2.4	Step 4: Governing equations	52
4.2.5	Step 5: Boundary conditions	54
4.2.6	Step 6: Energy of the birod	55
4.3	Elastic constants	55
4.4	Results	57
4.5	Conclusion	63
5	Statistical mechanics of an elastic birod	65
5.1	Introduction	65
5.1.1	Kinematics	65
5.2	Force-displacement curve for birod	68
5.3	Application to DNA melting	71
5.4	Effect of tensile force on the melting temperature	74
5.5	Conclusion	75
6	Conclusion	78
A	Supplement to Chapter 2	81
A.1	Dependence of Gibbs free energy on GC content	81
A.2	Poisson-Boltzmann Equation	82
A.3	List of symbols used in the text	83
B	Supplement to Chapter 4	85
B.1	Exponential decay of interaction energy in a ‘ladder’	85
B.1.1	Step 1: Kinematic description of the two strands	86
B.1.2	Step 2: Rotation of the two strands	86
B.1.3	Step 3: Extension and rotation of the web	87
B.1.4	Step 4: Governing differential equations	88
B.1.5	Step 5,6 and 7: Interaction Energy	90
B.2	Kinematics of the – strand	90
B.3	Evaluation of material properties of the web	92
B.4	Choice of eigenvalues obtained in section 5	94
B.5	Results for $\alpha = \pi$ radians	95
C	Dependence of interaction energy on boundary conditions and elastic constants.	96
D	Supplement to Chapter 5	98
References		101

List of Tables

2.1	The key micro-structural feature distinguishing the three phases is the number of base-pairs per helical turn. Elastic constants for various phases of DNA.	10
A.1	Symbols used in the text.	84

List of Figures

1.1	Manifestations of DNA elasticity.	2
1.2	Coexistence lines for force (f) and torque (M) driven structural transitions in DNA. The three phases of DNA have different double helical structures characterized by base-pairs per helical turn: B-DNA–ground state of DNA, 10.3 bp/turn; S-DNA–over-stretched state, 37 bp/turn; P-DNA–over-twisted state, 3 bp/turn. For more information see Marko <i>et al</i> (Sarkar et al. (2001))	2
1.3	The birod model of DNA. The proteins P1 and P2 interact via allosteric interactions.	3
2.1	Coexistence curves for B-, S- and P-DNA obtained by equating the free energy for two phases at a time. We have used $\sigma_{BS} = \sigma_{BP} = \sigma_{PS} = 0.044$. Note that the exact values of σ_{BP} and σ_{SP} are not known and hence have been assumed to be equal to σ_{BS} . Later, we will examine the effects of varying σ_{PS} , σ_{BP} , etc. Experimental data from Bryant <i>et al.</i> (Bryant et al. (2003)) appear as red dots in the plot above and are in good agreement with our coexistence lines.	12
2.2	Torque-rotation curves for a B-DNA to P-DNA transition at constant force. The optical bead attached to one end of the molecule is twisted while maintaining a constant tension of 45 pN until B-DNA converts to P-DNA. Most of the twist takes place at constant torque. The lines in different colors correspond to different σ_{BP} values. The inset shows the trajectory on an $f - M$ plane.	15
2.3	Structural transitions in a torsionally constrained DNA molecule. (a) Force-extension curve, (b) $f - M$ plane, (c) Fractions of B-,S- and P-DNA. The points labeled A,B,C,D in each panel have the same λ_{ext} . At B the molecule enters the triple point and at C it exits the triple point as extension λ_{ext} increases. Notice that a mixture of B- and S-DNA transforms into a mixture of S- and P-DNA. In (a), as σ_{BS} increases, the transition width increases, as expected. This is shown using different color trajectories in (a) and (b).	17
2.4	Plot of overstretching force for a torsionally unconstrained DNA as a function of ion concentration. We use $q_S = 2.8$ and $a_S = 0.8$ nm. In the figure blue diamonds are the experimental data from Zhang et al (X. Zhang et al. (2012)). In the inset, the electrostatic free energy for S-DNA decreases as the concentration increases.	21

2.5 Effect of ion concentration on the structural transitions in torsionally constrained DNA. (a) force-extension curves, (b) $f - M$ plane, (c) f_3 versus ion concentration, and (d) f_4 versus ion concentration. These results assume $\sigma_{BS} = 0.044$ (Argudo & Purohit (2014a)), $\sigma_{BP} = 0.18$ (Bryant et al. (2003)) and $\sigma_{PS} = Ac_0^n$. Experimental data from King *et al* (King et al. (2016)) is also shown for comparison. The force corresponding two phase transitions T_3 and T_4 is marked by f_3 and f_4 on $f - \lambda_{ext}$ 22

2.6 Non-monotonic trend in external work done per unit reference contour length. Note that the length of the DNA molecule is of the order of thousands of nm which is why a small difference in the work done will get amplified. We take $\sigma_{BS} = 0.04$, $\sigma_{BP} = 0.18$ and $\sigma_{PS} = Ac_0^n$ 24

3.1 Birod model of DNA. The angle between the tangent \mathbf{t}^+ and \mathbf{e}_3 is k . A base pair in reference and deformed state is shown. The director frames attached to \pm ends of the base pair change from \mathbf{Q}_0 to \mathbf{Q}^\pm , respectively. The rigid rotation of the strand $\mathbf{Q} = (\mathbf{Q}^+\mathbf{Q}^{-T})^{\frac{1}{2}}\mathbf{Q}^-$ and micro-rotation $\mathbf{P} = (\mathbf{Q}^+\mathbf{Q}^{-T})^{\frac{1}{2}}$ 30

3.2 Distortions to the helical geometry of the + strand. We assume that the DNA remains straight after the binding of proteins. Therefore, the axis \mathbf{e}_3 remains undisturbed. The radius of the helix a changes to $a + r$, the phase angle ωx changes to $\omega x + \beta$, and the length of the infinitesimal element dx , shown in black, changes to $(1 + \xi)dx$ 31

3.3 Free-body diagrams which establish the connection between an elastic rod and an elastic birod. We deliberately show a straight ladder instead of helical birod to illustrate the mechanics. An elastic birod comprises two elastic rods + and -. The + strand exerts a moment \mathbf{c} and force \mathbf{f} on the - strand through an elastic web. This transfer of moment and force leads to deformation of the web. In the figure \mathbf{r}^\pm denotes the position vector for \pm strands, and \mathbf{n}^\pm and \mathbf{m}^\pm denote the contact forces and contact moments in \pm strands, respectively. The force and moment balance for + and - strand constitute the governing equations (eqn. (3.3)) for the elastic birod. For further discussion see Moakher and Maddocks (Moakher & Maddocks (2005)). 32

3.4 Variation of r, k_3, ξ and $\beta^+ = \beta^- = \beta$ for a single protein. The red curve corresponds to the boundary conditions $\beta_0 = 0, r_0 = 0.05$ nm and the green curve to $r_0 = 0, \beta_0 = 0.05$. The decay length is $l_d = \zeta^{-1} \approx 10$ bp which is close to that documented in literature (Kim et al. (2013); Gu et al. (2015)). 36

3.5 We show the deformed configuration of the double helix, red and green colors correspond to + and - strand, respectively. In the first figure, one protein binds at $x = 0$ with $r_0 = 0.2$ nm and $\beta_0 = 0$. In the second figure, two proteins bind at $x = \pm 1.5$ nm. In the third figure, two proteins bind at $x = \pm 3.5$ nm. Notice the overlap of deformations in the second figure which is absent in the third one. This overlap is manifests itself as interaction energy between the two proteins. The dotted lines denote the corresponding undeformed configuration. 37

3.6 The first figure shows the variation of interaction energy ΔG with distance p between the two proteins P_1 and P_2 . The boundary conditions $r_1 = 0.001$ nm, $\beta_1 = 0.0045$ for P_1 and $r_2 = 0.001$ nm, $\beta_2 = -0.0045$ for P_2 give the best fit to the experimental data for ΔG (Kim et al. (2013)). In the second figure, we show the variation of change in groove width $\rho(x) = g(x) - \frac{p}{2}$ when a protein with boundary conditions r_0, β_0 binds at $x = 0$. The decaying sinusoidal character is documented in previous work (Kim et al. (2013))(Gu et al. (2015)). The magnitude of the change in groove width (~ 3 Å) is consistent with estimates in (Kopka et al. (1985)). 38

3.7 Eqn. 3.15 shows that the strain parameters r, β , and ξ decay exponentially while oscillating with the periodicity of the double helix. Let us assume that the protein binding at $x = 0$ increases the radius of the double helix from a to $a + r_0$. This change in radius at $x = 0$ decays exponentially while oscillating with the periodicity of the double helix, away from the binding site. Similar behavior is observed for other strain parameters, β and ξ . Due to this sinusoidal modulation of the geometry, the binding of the second protein is facilitated at some locations, while inhibited at others; this manifests as an exponentially decaying oscillatory behavior observed in the allosteric interaction energy (ΔG). 39

3.8 Consider two proteins P1 and P2 binding at $x = 0$ and $x = p$, respectively. The first figure shows the variation of normalized free energy $\Delta G(p)_n = \frac{\Delta G}{|\Delta G|_{\max}}$ and correlation function $q(p)_n = \frac{q(p)}{|q|_{\max}}$ with the distance between the two proteins p . The correlation function $q(p) = (r_0^1 r_0^2 + r_p^1 r_p^2) + (\beta_0^1 \beta_0^2 + \beta_p^1 \beta_p^2)$ where r_p^1 is the change in radius caused by protein P1 at $x = p$. The boundary conditions for the two proteins are given in the figure. We find that the peaks and valleys of ΔG_n and q_n coincide, however the magnitudes are not identical. We find that the magnitudes are related as $\Delta G_n \approx \tanh(3q_n)$, as shown in the inset. We test this empirical relation for two different sets of boundary conditions and find a remarkable match. The diamonds denote the free energies computed using eqn. 4.40 and the solid line denotes the free energy computed using the normalized correlation function $q(p)_n$. This exercise shows that the correlation functions can be used as a surrogate for free energies. 41

4.1 A DNA molecule as a double helical elastic birod is shown on the left. The phosphate backbones are represented by outer strands while the complementary base-pairing is represented by the elastic web. The phase angle between the two helices is $\alpha = 2.1$ radians. Here $\mathbf{R}^+ = [\mathbf{n}_0^+ \quad \mathbf{b}_0^+ \quad \mathbf{t}_0^+]$ and $\mathbf{R}^- = [\mathbf{n}_0^- \quad \mathbf{b}_0^- \quad \mathbf{t}_0^-]$ are the Frenet-Serret frames attached to the + and - strands, respectively. Base-pairs in reference and current configuration are shown to the right. $\mathbf{Q}_0^+ = \mathbf{Q}_0^- = \mathbf{Q}_0$ in the reference configuration. In the current configuration, the rigid rotation of the base-pair is quantified by $\mathbf{Q} = \mathbf{Z}(\mathbf{1} + \mathbf{\Phi})\mathbf{Q}_0$ (eqn. 4.23) and the elastic moment \mathbf{c} is related to the Gibbs rotation vector of $\mathbf{P} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{\frac{1}{2}}$ (eqn. 4.19). 46

4.2 Variation of strain variables for a single protein. We plot the change in radius r , twist k_3 , stretch of centerline ξ and change in phase angle $\Delta\beta = \beta^+ - \beta^-$ for the double-helix. The red curve corresponds to the boundary conditions $k_{10} = r_0 = 0$ and $k_{20} = 0.1 \text{ nm}^{-1}$ at $x = 0$ and the green curve corresponds to $k_{10} = k_{20} = 0$ and $r_0 = 0.05 \text{ nm}$ at $x = 0$. The asymmetry of the double-helix (there is a major and minor groove in DNA) arising from the phase angle $\alpha = 2.1$ radian gives the curves a slight asymmetry about the site of protein binding. The curves are exactly symmetric about the site of protein binding if we choose phase angle $\alpha = \pi$ radians (which results in no major and minor groove) as shown in the appendix. 58

4.3 Variation of curvatures k_1 and k_2 for a single protein. The red curve corresponds to the boundary conditions $k_{10} = r_0 = 0$ and $k_{20} = 0.1 \text{ nm}^{-1}$ at $x = 0$ and the green curve corresponds to $k_{10} = k_{20} = 0$ and $r_0 = 0.05 \text{ nm}$ at $x = 0$. We find that the curvature decays exponentially and oscillates with a period $\approx 11 \text{ bp}$ 58

4.4 Variation of r, k_3 and $\Delta\beta$ for two proteins. Here a is the distance between the sites of protein binding. The strain variables decay exponentially away from the site of protein binding. When the distance between the proteins is large $10 \times 3.4 \text{ nm}$, the profile looks like a concatenation of two solutions for a single protein. 59

4.5 We plot the interaction energy between two proteins eqn. (4.40). In (a) we plot the behavior of ΔG for various boundary conditions. If the boundary conditions are specified on the curvatures we get an exponentially decaying profile oscillating with $5 - 6 \text{ bp}$ ($\approx 11/2 \text{ bp}$). The oscillatory behavior arises from the periodic geometry of DNA. In (b) the experimental data reproduced for comparison are from Kim *et al* (Kim et al. (2013)). We use $k_{11} = k_{21} \approx 0.02 \text{ nm}^{-1}$, $k_{12} = k_{22} = 0.05 \text{ nm}^{-1}$, $r_1 = -r_2 = 0.02 \text{ nm}$. The inset in (a) shows a protein DNA complex in which the proteins locally bend DNA. The inset in (b) shows that bending a DNA oligomer leads to widening of the groove on one side while narrowing it on the other. We find that the change in the groove width is approximately $\approx 2 \text{ \AA}$ which is close to the values reported by Kopka *et al* (Kopka et al. (1985)) ($0.5 - 2 \text{ \AA}$) 60

4.6 The inset in (b) shows a two protein complex. The boundary conditions are identical for both the proteins $k_{11} = k_{21} = k_{10}$, $k_{12} = k_{22} = k_{20}$, $r_1 = r_2 = r_0$; the legend in (a) contains the exact numerical values. For (b) the legend is the same as in (a). We examine behavior of $\hat{w}_3(x = 2 \text{ nm}, a)$ (eqn. 4.39) as a function of distance between the two proteins a for these boundary conditions. The strain variables oscillate with a period of 11 bp. We observe that in case of $r_0 = 0$, the strain parameter $\hat{w}_3(x = 2 \text{ nm}, a)$ decays as $e^{-\Gamma a} \psi(\omega a)$ where $\psi(\omega a)$ is a sinusoidal function, hence the combined energy of a two protein complex which is proportional to $(e^{-\Gamma a} \psi(\omega a))^2$ oscillates with a period of 5.5 bp (period of $\sin^2 x$ is half that of $\sin x$). If $k_{10} = k_{20} = 0$ the decay is exponential. If $r_0 \neq 0$ and k_{10} or $k_{20} \neq 0$, $\hat{w}_3(x = 2 \text{ nm}, a) \sim (e^{-\Gamma_1 a} \psi(\omega a) + e^{-\Gamma_2 a})$ and the energy of the two protein complex, which is proportional $(e^{-\Gamma_1 a} \psi(\omega a) + e^{-\Gamma_2 a})^2$, oscillates with a period of 11 bp. The behavior of the other strain variables in eqn (4.39) is similar. We plot the interaction energy $\Delta G(a)$ in (b) for the boundary conditions indicated in the legend of (a) and use it to verify the period we predict using this argument. 60

4.7 Decay length l_d is defined as the inverse of the eigenvalue with the least non-zero magnitude, for $\chi = 1$, $l_d = 1/|\lambda|_{min} = 1/0.34 \approx 9$ bp. χ is meant to account for the reduction in the elastic constants for AT base-pairs compared to GC base-pairs. The elastic constants for the web are $(\chi K_c, \chi K_e, \chi H_1, \chi H_2, \chi H_3, \chi L_1, \chi L_2, \chi L_3)$, eqn. (4.41) gives the numerical values for $\chi = 1$. We find that the decay length increases with a decrease in elastic constants for the web, thus AT-rich DNA sequences are expected to have higher decay lengths. Qualitative experimental and numerical evidence in support of the above conclusion is documented in (Kim et al. (2013)) and (Gu et al. (2015)), respectively. The inset shows how we extracted the the power law $l_d \sim \chi^{-2/3}$ 63

5.1 The cartoon of a straight birod acted upon by identical forces on both the strands. The two strands are referred to as \pm strands. 66

5.2 (a) Force-extension curve for the birod. The parameter values are $EI = 0.8 \text{ pNnm}^2$, $L_1 = 80 \text{ pN/nm}^2$, $L_2 = 1280 \text{ pN/nm}^2$, $H = 0.33 \text{ pN}$, $a = 1 \text{ nm}$, and $\lambda = 0.5 \text{ \AA}^{-1}$. The unit for the force values shown in the graph are $eV/\text{\AA}$ ($1 \text{ eV/\AA} = 1600 \text{ pN}$). Using the WLC chain formula (Grosberg et al. (1995)), the best fit to the data yields a persistence length of $l_p = 36 \text{ nm}$, which is close to the persistence length of of ds-DNA–50 nm (. (b) Effect of increase in L_1 . The force-displacement curve is independent of stretch modulus L_2 and the bending rigidity H of the web. (c) Effect of increase in temperature (d) Effect of increase in the stiffness of the outer strands. (d) $\langle w^2 \rangle$ vs x for various values of tensile force F with hinged-hinged boundary conditions. 72

5.3 The average distance between the two strands $\langle v \rangle$ as a function of the temperature T 73

5.4 The birod cartoon for studying DNA melting. Note that the force is applied only on one strand. 74

5.5 (a) The effect of force on the the $\langle v \rangle$ vs T curves (b) The effect of the tensile force on the melting temperature. The experimental data is from X. Zhang et al. (2012). 76

B.1	A straight birod, referred to as a ladder, being pulled by two force pairs separated by a distance a . We show that the interaction energy between the two force pairs given by $\Delta G = E_a^2 - E_0^1 - E_a^1$ decreases exponentially with a	86
B.2	Variation of strain variables for $\alpha = \pi$ radians. Notice that the curves are symmetric about the site of protein binding. As mentioned in section 6, the curves are not symmetric if we choose $\alpha = 2.1$ radians.	95
C.1	Effect of elastic constants on the decay length l_d (nm) and period ν . ω and P are the period and pitch, respectively for dsDNA. The elasticity of the web is parametrized by a single parameter A such that $L_1 = L_2 = L_3 = A pN/nm^2$ and $H_1 = H_2 = H_3 = A pN$. We choose $A = 10$ since it gives decay length in the right ballpark Kim et al. (2013).	97

Chapter 1

Introduction

This thesis uses the theory of a double stranded elastic rod ([Moakher & Maddocks \(2005\)](#)) to develop a mechanics-based model of DNA, henceforth called *birod* model. The model describes the geometry and elasticity of DNA across several length scales, which when combined with statistical mechanics reveals the micromechanics behind several experimentally observed phenomena such as structural transitions in DNA, allosteric interactions, and DNA melting. DNA is a double-helical long thread-like molecule with a radius of 1 nm and pitch of 3.4 nm capable of bending, twisting, and stretching elastically. At the level of a few hundred nanometers, DNA elasticity plays an important role in determining geometry of the protein-DNA complexes as shown in [fig.1.1\(a\)](#). This geometry is crucial to biological processes such as DNA replication, ultimately influencing gene expression ([Kim et al. \(2013\)](#)). The lactose metabolism system in *E. coli* is controlled by the Lac Repressor protein which binds to two different sites on DNA molecule forming a loop as shown in [fig. 1.1\(b\)](#). The gene expression is hence directly linked to the looping ability (resistance) of DNA ([Purohit & Nelson \(2006\)](#)). The bending and twisting resistance of DNA determines the forces necessary to package the viral genome into a capsid ([Purohit et al. \(2003\)](#)). On a macroscopic length scale spanning a couple of hundred micrometers, DNA elasticity manifests itself as cooperative structural transitions, which have been used to construct several synthetic molecular motors capable of performing mechanical work *in vivo* ([Bryant et al. \(2003\)](#); [Rich & Zhang \(2003\)](#)) shown in [fig.1.1\(c,d\)](#).

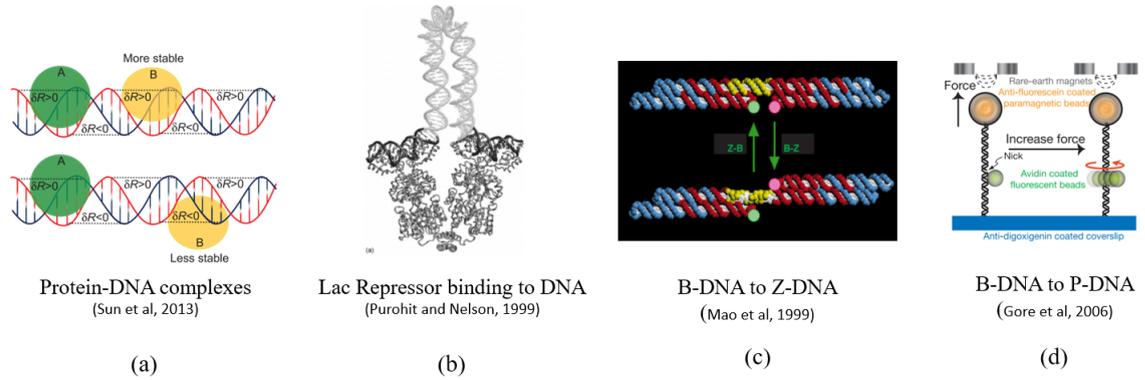


FIGURE 1.1: Manifestations of DNA elasticity.

In this document, we discuss three problems related to DNA elasticity:

1. **Structural transitions in DNA** (See fig.1.2) : When a force or torque is applied to a DNA molecule, it extends or twists elastically up to a certain limit beyond which it suffers a sudden structural transition whose products depend on the boundary conditions on the two ends of the molecule. These structural transitions, akin to first order phase transitions, are cooperative and are characterized by a sudden jump in the displacement variables followed by a change in the internal structure of the DNA. The primary concern in chapter 2 is to examine these structural transitions using a homogeneous elastic rod model for DNA.

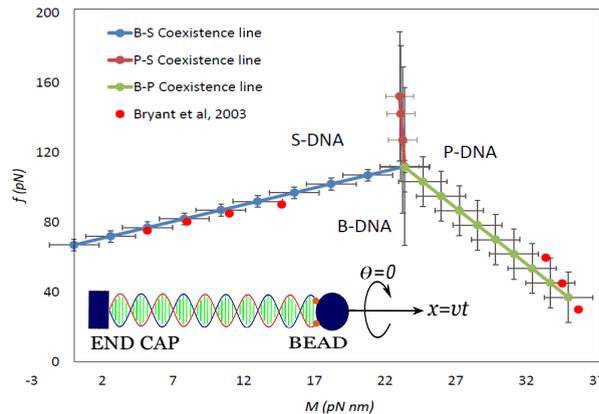


FIGURE 1.2: Coexistence lines for force (f) and torque (M) driven structural transitions in DNA. The three phases of DNA have different double helical structures characterized by base-pairs per helical turn: B-DNA—ground state of DNA, 10.3 bp/turn; S-DNA—over-stretched state, 37 bp/turn; P-DNA—over-twisted state, 3 bp/turn. For more information see Marko *et al* (Sarkar *et al.* (2001))

2. **Allosteric interactions in DNA** (See fig.1.3): When a ligand binds to DNA it induces conformational changes at the binding site which propagate to regions tens of base-pairs away, thereby encouraging or inhibiting the binding of a second ligand

in those places. Such interactions between two binding agents are called *allosteric* interactions. In chapters 3 and 4, we use an elastic birod to study the allosteric interactions in DNA. Our model provides a framework for viewing allosteric interactions in DNA within the ambit of configurational forces of continuum elasticity.

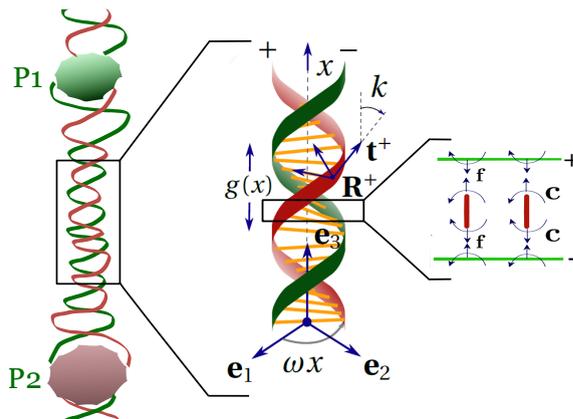


FIGURE 1.3: The birod model of DNA. The proteins P1 and P2 interact via allosteric interactions.

3. **Melting transition in dsDNA:** Beyond a certain temperature, called the melting temperature, the complimentary base-pairing holding together two phosphate backbones gets disrupted and dsDNA disintegrates into two single strands. This transition, called the melting transition, is highly cooperative and resembles a first-order phase transition. A combination of statistical mechanics and continuum mechanics of an elastic birod is used to study various characteristics of the melting transition. The model predicts that the melting temperature decreases with increasing tensile loads, in agreement with reported experimental observations.

To study structural transitions in chapter 2, we use a homogeneous elastic rod model while we use an elastic birod model to study the allosteric interactions. The reason for using two different models is the difference in the length scales of the two problems. In the first problem involving structural transitions, we are dealing with molecules with lengths ranging from a few tens of μm to a few mm, while allosteric interactions are restricted to only 1-10 nm (1-3 helical turns of DNA). Therefore, in the former case, we neglect the double helical geometry and mechanics of base-pairs, and use a homogeneous rod model for DNA. In the problem involving allosteric interactions (~ 10 nm), we need an additional layer of magnification to deal with the mechanics of strands and base-pairs separately, which we note, in retrospect, play a crucial role. We attempt to bridge the gap between the two models by connecting the stretch modulus, twist modulus and twist-stretch coupling modulus used in the homogeneous elastic rod model to

the elastic constants in the birod model. The melting transition, observed at various length scales, is also studied using an elastic birod model because we need to compute the average distance between the two strands as a function of temperature, which can be done only in a birod. Furthermore, we show using a statistical mechanical approach that the fluctuating elastic rod model (or WLC model) can be recovered from the birod in the limit of long contour length.

The DNA molecule exhibits a double-stranded (ds) right handed helical structure with a helical repeat of 3.5 nm and 10.5 base pairs (bp) per helical turn in its well known B-form, the two phosphate backbones being held together by hydrogen bonding arising from complimentary base pairing and base stacking interactions. The geometrical properties of the double-stranded helix such as pitch, base pairs per helical turn, and contour length per base pair characterize different phases of DNA such as B-DNA, S-DNA and P-DNA *etc.* (J. F. Marko & Neukirch (2013)). The most commonly observed form of DNA in living cells is B-DNA. S-DNA can be formed as a product of a force driven cooperative overstretching transition of torsionally unconstrained B-DNA at lower temperatures which preserves the base pairing (King et al. (2013)). Its contour length is 1.7 times that of B-DNA with 37 base pairs per helical turn (J. F. Marko (2007)). P-DNA can be formed as a product of a torque driven overstretching transition of B-DNA (Bryant et al. (2003)). Its contour length is also 1.7 times that of B-DNA, but with 3 base pairs per helical turn. M-DNA (melted-DNA) can be formed as a product of a force driven overstretching transition of B-DNA at higher temperatures and low ionic concentrations (King et al. (2013)). It is characterized by the absence of base-pairing, but both the strands are under tension.

The single molecule experiment in which these phases are produced involves fixing one end of the DNA molecule and attaching a bead on the other, which is pulled using optical or magnetic tweezers. The resulting structures can be analysed using various proteins or dyes which selectively bind to a specific phase of DNA (King et al. (2013, 2016)). In such experiments the rotation of the glass bead can be held fixed, which corresponds to a torsionally constrained bead, or it can be allowed to rotate freely, which corresponds to a torsionally unconstrained bead. The resulting force-extension curves show a strong sigmoidal nature wherein the displacement suffers a sudden increase over a small interval of force, which is suggestive of a first-order phase-transition. Such a transition with a distinct plateau has been referred to as an ‘overstretching transition’ in the literature (J. F. Marko (2007)).

Several theoretical models have been proposed to study these transitions from the standpoint of statistical mechanics (Sarkar et al. (2001); Leger et al. (1999)). However, these models fail to account for the chirality of a DNA molecule which manifests itself as a twist-stretch coupling (Gore et al. (2006); Kamien et al. (1996)). Thermal undulations further modify the energetic landscape, the explicit calculations for which using the

elastic rod model are presented in (Argudo & Purohit (2014b)). Other models such as the worm-like-chain model (WLC) and freely-jointed-chain model (FJC) can also be used to understand the various aspects of the overstretching transition in the torsionally unconstrained case (Rouzina & Bloomfield (2001a,b); Storm & Nelson (2003)). A phase diagram for DNA on a force-torque plane is calculated in (J. F. Marko & Neukirch (2013)) by equating the free energies comprising of the sum of the elastic energies of torsion and stretching (using the WLC model).

The ion concentration plays a critical role in determining the structure of products and underlying mechanism of the overstretching transition. Experimental observations reported in (King et al. (2013); X. Zhang et al. (2013, 2012)) suggest that in a torsionally unconstrained molecule, the effect of increasing the ion concentration is two-fold: a) it leads to an increase in the overstretching force, and b) it dictates whether the over-stretched state will be ss-DNA, M-DNA or S-DNA. The change in mechanism from strand unpeeling to cooperative overstretching with increase in ion concentration from 10 mM to 150 mM results in different force-displacement curves (King et al. (2013)).

The objectives of the study presented in chapter 2 are: a) we investigate the overstretching transition in a torsionally constrained molecule using the elastic rod model postulated in (Kamien et al. (1996)) and (Argudo & Purohit (2014b)), and b) we quantify the effect of changes in ion concentration of the surrounding media on the overstretching transition. The primary contribution of our model is to resolve the question: Why does overstretching force and associated work done vary non-monotonically with ion concentration in the case of a torsionally constrained DNA (King et al. (2016)), while in the case of a torsionally unconstrained DNA the aforementioned quantities increase monotonically with ion concentration (Rouzina & Bloomfield (2001b)). Our methodology uncovers how the fractions of B-,S- and P-DNA change while the molecule is pulled through the triple point in the phase diagram. We conclude the chapter with a form for the dependence of interfacial energy on ion concentration which leads to falsifiable predictions.

While the homogeneous elastic rod model for DNA suffices for certain applications such as structural transitions (King et al. (2016)) and supercoiling of DNA oligomers (Purohit (2008)), its application is inherently limited for studying phenomenon that are fundamentally linked to the double stranded structure of DNA such as allosteric interactions between two proteins on DNA (Kim et al. (2013)), temperature driven strand separation (Rouzina & Bloomfield (2001a)), and origin of negative twist-stretch coupling (Bryant et al. (2003)). We apply the birod theory proposed Moakher and Maddocks (Moakher & Maddocks (2005)) to these problems. In a double stranded elastic birod model, in addition to the standard variables of the Cosserat rod theory (i.e., center line of the rod cross-section $\mathbf{r}(s, t)$ and a material frame $[\mathbf{d}_1(s, t) \quad \mathbf{d}_2(s, t) \quad \mathbf{d}_3(s, t)]$), there are two micro-structural variables— $\mathbf{w}(s, t)$, a micro-displacement measuring the change

in distance between the two strands, and $\mathbf{P}(s, t)$, a micro-rotation measuring the change in orientation of one strand relative to the other. Fortunately, the forces conjugate to these micro-structural variables obey balance laws that look similar to the balance of forces and moments equations of a standard Cosserat rod. Moakher and Maddocks (Moakher & Maddocks (2005)) have provided hyper-elastic constitutive laws for these micro-structural variables that are based on *quadratic* energies.

In chapters 3 and 4, our motive is to use the elastic birod model to study allosteric interactions in DNA within the scope of configurational forces of continuum elasticity. Just as defects produce local elastic fields in a solid, proteins binding to DNA also deform it locally. Since DNA behaves like an elastic rod at scales of a few tens of nanometers (Purohit et al. (2008)), we expect that if two proteins bind to DNA separated by a distance p then the deformation fields created by them will overlap and lead to an interaction energy which depends on p in a clearly quantifiable way. This problem has not been theoretically addressed so far, but there is experimental evidence of the interaction. Some of this experimental evidence has been extracted by connecting the interaction energy with the kinetics of protein binding/unbinding. In spirit, this is similar to continuum elasticity in which configurational forces often determine defect dynamics through a kinetic law (Phillips & Rob (2001); Gurtin (1999)). Kim *et al.* (Kim et al. (2013)) have exploited this connection of interaction energies to kinetics to show that gene expression, which depends on RNA polymerase binding affinity to DNA in live bacteria, is a function of the proximity of LacR and T7 RNA polymerase bound to DNA.

In the experiments of Kim *et al.* (Kim et al. (2013)) one end of a DNA molecule is attached to the passivated surface of a flow cell and binding sites are provided for two specific proteins to bind. The length of the DNA between these binding sites, a , is increased in 1bp increments between 7 base-pairs (bp) and 45bp. First, one type of fluorescently labeled protein (call it A) is flowed into the cell so that it binds to the DNA. Then, the second protein (call it B) is flowed in at a specific concentration. The dissociation times of the fluorescent protein are then monitored as a function of a . This dissociation time depends on the free-energy change ΔG of the DNA + two protein complex from the state when the two proteins are bound to that when protein A is unbound. Now, in general, the free energy ΔG of the ternary complex formed by the DNA and proteins A and B consists of three parts (Kim et al. (2013)):

$$\Delta G = \Delta G_A + \Delta G_B + \Delta\Delta G_{AB}(a), \quad (1.1)$$

where ΔG_A and ΔG_B are the free energy changes caused by binding of A and B alone respectively to the DNA. These are constants. The last term $\Delta\Delta G_{AB}(a)$ is the portion of the free energy change that accounts for the interaction of the two proteins bound

to the DNA while being separated by a distance a . In chapters 3 and 4, we give a mechanical origin for the variation of $\Delta\Delta G(a)$.

Allosteric effects and their relation to protein DNA interactions have been studied using molecular dynamic (MD) simulations (Gu *et al.* (2015); Hancock *et al.* (2013)). Gu *et al.* (Gu *et al.* (2015)) have studied various kinds of deformations which include shift, roll, rise, twist, slide, and tilt of the DNA bases. They observed a sinusoidal correlation in the major groove widths similar to the one observed by Kim *et al.* (Kim *et al.* (2013)). Furthermore, Gu *et al.* point out that the presence of GC rich sequences dampens the allosteric effects which is what Kim *et al.* observe experimentally. Kim *et al.* (Kim *et al.* (2013)) classify the DNA binding proteins into two categories: ones that bind to straight DNA and others that bend DNA. In chapter 3, we deal with the simpler case involving proteins that bind to straight DNA. In chapter 4, we extend that formulation by incorporating the 3-D bending and torsion of DNA molecules.

In chapter 5, we combine statistical mechanics (Landau & Lifshitz (1969)) with the continuum mechanics of an elastic birod (Moakher & Maddocks (2005)) to study temperature driven strand separation in dsDNA. This phenomenon is called DNA melting and there is plenty of experimental evidence for it (Williams *et al.* (2001); Rouzina & Bloomfield (2001b)). Additionally, several models have also been proposed which capture the experimental data. Experimental evidence suggests that this entropic transition from dsDNA to ssDNA is strongly cooperative (Gibbs-Davis *et al.* (2007); Nishigaki *et al.* (1984)), and the increase in the inter-strand distance with temperature is sudden as opposed to a mere incremental linear thermal expansion.

Although thermal melting of DNA has been studied extensively for some decades now (Lehman & McTague (1968); Crothers (1968)), it has not been analyzed within the birod framework of continuum mechanics. The sharp transition characterizing the cooperative DNA melting depends on various factors such as (X. Zhang *et al.* (2012)), i) the internal base-pair sequence: DNA fragments with higher GC content have higher melting temperatures, ii) the tensile force: experimental evidence suggests that the melting temperature decreases as the tensile load increases, and iii) the ion concentration: the ion concentration regulates the balance between the transition to S-DNA and ssDNA—increasing the ion concentrations favors the former. Most attempts to model the melting transition rely on using the Clausius-Clayperon equation to get empirical relations between the various quantities of interest such melting temperature versus tensile force and over-stretching force vs the ion concentration *et cetera* (Rouzina & Bloomfield (2001a,b)). Since these empirical models interface directly with the experimental data, they have superior predictive capabilities and produce experimentally verifiable predictions. However, this approach obscures the underlying mechanics or biochemistry driving the process. To circumvent this drawback to some extent, several statistical mechanics based models such as (Dauxois *et al.* (1993); Peyrard & Bishop (1989)) begin by positing

an energy functional that reproduces the characteristic cooperative melting transition behavior observed experimentally. The detailed description sheds light on the underlying mechanism—cooperative H-bond disruption—and enables one to clearly discern the effect of various parameters such as the cooperativity parameter and asymmetric interactions on the melting transition. However, the inherent analytical intractability involved in evaluating the path integrals central to these approaches limits their application to relatively simple energy functionals. Molecular dynamics simulations have also been used to study several problems related to melting transition, unzipping of DNA, and other structural transformations in DNA (Pérez et al. (2012); F. Zhang & Collins (1995)). The interaction potentials are available for various chemical bonds in DNA thereby permitting a very detailed description of the structure, shape, and chemistry crucial to the problem. The results from the simulations agree well with the experimental data, but, they entail a huge computational burden.

Here, the DNA is modeled using an elastic birod; for analytical tractability it is assumed to be a straight ladder-like birod. However despite this simplifying assumption, the model is able to capture the key mechanical characteristics of DNA such as the worm-like-chain force-extension curve for long DNA oligomers, the cooperativity in the melting transition, and the effect of tensile force on the melting temperature. The force-extension curve from the model can be approximated by a worm-like-chain chain formula with a persistence length of a few tens of nm which is close the actual persistence length of dsDNA (50 nm). The model improves upon the existing statistical mechanics approaches by computing (as opposed to positing) the energy functional from the kinematic description. Additionally, the novelty achieved in our model is two-fold: i) we find that the elasticity of the outer strands is responsible for the cooperativity observed in the melting transitions, and ii) our model accounts for the effect of tensile force on the melting temperature. However, since the model uses a straight elastic birod restricted to planar deformations, instead of a helical birod with 3-D deformations and rotations, it can not be used to study phenomena such as phase transitions among B-, S-, and P-DNA (Sarkar et al. (2001)). Furthermore, it can not account for a negative twist-stretch coupling of DNA helix (Bryant et al. (2003)).

Chapter 2

Structural transitions in torsionally constrained DNA and their dependence on solution electrostatics

2.1 Model

The single molecule experiment under consideration involves attaching one end of the DNA molecule to an end cap while the two strands on the other end are bound to an optically or magnetically trapped bead. We conceive of a displacement controlled experiment wherein the bead is pulled at a constant velocity in a quasistatic manner resulting in a uniaxial extension of the molecule. As the displacement on the bead increases, various structural transitions occur in the molecule whose study is the object of central concern in this section. The temperature is assumed to be constant at 300 K throughout the experiment.

The behavior of a DNA molecule in response to force and torque is markedly different from both an elastic rod and an entropic polymer. The ability of DNA to resist torsional loading distinguishes it from classical polymers wherein free rotation is permitted at the joints of the monomers. As a result, the persistence length of DNA is much larger than the persistence length of classical polymers. On the other hand, a pronounced effect of entropic disturbances on a molecular length scale distinguishes its response from that of an ideal elastic rod.

In order to incorporate the effects of entropic disturbances on the slender topology of the molecule, we borrow ideas from theory of fluctuating elastic rods ([Purohit et al. \(2008\)](#));

[Argudo & Purohit \(2014b\)](#)) to posit a free energy function which is a sum of the elastic component G_j^e and the entropic component G_j^f , $G_j = G_j^e + G_j^f$, where j denotes the phase of DNA. G_j is the Gibbs free energy of the molecule per unit reference contour length which is approximately 3 times ($1/0.34 = 2.94$ bp/nm) the energy per base pair. Such a treatment inherently presumes the DNA molecule to be a long homogeneous elastic rod capable of undergoing phase transitions and responding to thermal fluctuations.

The linear elastic component of Gibbs free energy per unit reference length is quadratic in force f and torque M .

$$G_j^e = -\frac{C_j}{2(S_j C_j - g_j^2)} f^2 + \frac{g_j}{S_j C_j - g_j^2} f M - \frac{S_j}{2(S_j C_j - g_j^2)} M^2 - \lambda_j^0 f - \kappa_j^0 M, \quad (2.1)$$

where C_j , S_j and g_j denote the elastic moduli for twisting, stretching and twist-stretch coupling for the j phase. The units for these constants are pNm², pN and pNm unless otherwise stated. λ_j^0 and κ_j^0 denote inherent axial stretch and torsional strain of the j phase of DNA with respect to B-DNA. These constants depend on the geometrical structure for each phase and are archived in table 2.1. b_j denotes the contour length per base pair for the j phase, $b_B = 0.34$ nm/bp. The stretch and torsion are calculated with respect to the contour length in the reference configuration, i.e. B-DNA.

We have assumed that bending energy does not make significant contributions to the

	bp/turn	A_j	S_j	C_j	g_j	λ_j^0	κ_j^0	Sources
B-DNA	10.3	205	1400	369	-84	1	0	Gore et al. (2006) , Argudo & Purohit (2014b) , Leger et al. (1999) , J. F. Marko & Neukirch (2013) .
S-DNA	37	28.7	3030	140	-122	1.7	-1.4	Fu et al. (2011) , Zhao & Purohit (2016) , J. F. Marko & Neukirch (2013) .
P-DNA	3	61.5	1600	900	-120	1.6	4.4	Zhao & Purohit (2016) , Leger et al. (1999) , Allemand et al. (1998) .

TABLE 2.1: The key micro-structural feature distinguishing the three phases is the number of base-pairs per helical turn. Elastic constants for various phases of DNA.

elastic energy. Since the magnitude of force in these single molecule experiments is large (> 10 pN), the formation of plectonemes has also been neglected (for more information see [Purohit \(2008\)](#)). Our formulation does not account for observed sequence dependent behavior of DNA, such as the variation of mechanical properties with base-pair sequence ([Weber et al. \(2009\)](#); [Lankas et al. \(2000\)](#); [Eslami-Mossallam et al. \(2016\)](#)), or the dependence of the end products on the sequence and location of basepairs ([X. Zhang et al. \(2013\)](#)). For example, it has been pointed out that certain sequences of base pairs, such as ATAT..., have lower elastic constants than others ([Weber et al. \(2009\)](#));

[Eslami-Mossallam et al. \(2016\)](#)). As another example, the stretching, bending, twisting stiffnesses, etc., of AA, AT, GC and GG sequences (with just two base-pairs) have been tabulated in Lankas *et al.* ([Lankas et al. \(2000\)](#)) in their table 1. In order to capture such effects in the parameters in our Table 2.1, which represents much longer DNA fragments (thousands of base-pairs), we could follow the approach of Manning *et al.* ([Manning et al. \(1996\)](#)) who came up with homogenized effective elastic constants starting from the geometric and kinematic details of base stacking. However, the stiffness values in table 1 of ([Lankas et al. \(2000\)](#)) are only for B-DNA. Similar stiffness values for the S-, P- or other phases are not known; similarly, parameters relating to the stacking geometry, such as, the relative rotation of one base-pair relative to the other are well-documented only for B-DNA, not for S-, P-, or other phases of DNA. Hence, we present a different calculation that accounts for sequence dependence in the expression for the Gibbs free energy of the DNA molecule in Appendix A. This calculation relies on experimental data for how the melting temperature of DNA depends on the fraction of GC base-pairs r_{GC} in the molecule ([X. Zhang et al. \(2012\)](#)). Although, the above approach cannot give us the details of how the twist, bend, stretching moduli depend on r_{GC} , but can still provide useful information to study the phase behavior of DNA, at least when it is torsionally unconstrained. On the other hand, many features of the phase behavior of DNA under torsional constraints can be captured using our simpler formulation (without sequence dependence), as was shown successfully for other experimental conditions in ([Argudo & Purohit \(2014b\)](#); [Zhao & Purohit \(2016\)](#)).

For a force f and torque M , the entropic contribution to the free energy of a homogeneous rod with reference contour length L subjected to tension and twist is $\beta G_j^L = C_1 - \ln Q_j + \ln \left[\left(\exp \frac{LQ_j}{A_j} \right) (1 - \exp(-\frac{2LQ_j}{A_j})) \right]$ ([Argudo & Purohit \(2014b\)](#)). Here $Q_j = \sqrt{A_j f - M^2/4}$, L is the reference contour length, C_1 is a constant independent of force and torque and $\beta = \frac{1}{k_B T}$ where k_B is the Boltzmann constant and T is the absolute temperature. The DNA segment in the experiments under consideration is held under high tension to prevent the torsional buckling which occurs at $M_c = 2\sqrt{A_j f}$. The applied torque $M \ll M_c$ which implies $Q \gg 1$ and the molecule is stable under small perturbations caused by thermal bombardment, then

$$G_j^f \approx \frac{G_j^L}{L} = \frac{Q_j}{\beta A_j} - \frac{\ln Q}{L}, \quad (Q \gg 1). \quad (2.2)$$

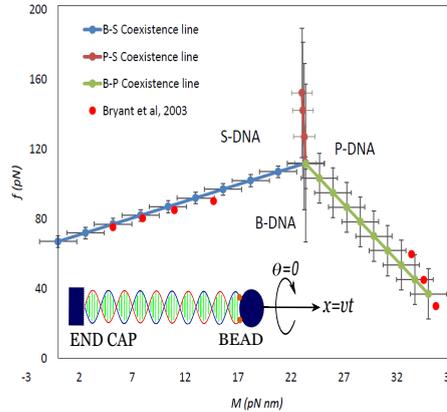


FIGURE 2.1: Coexistence curves for B-, S- and P-DNA obtained by equating the free energy for two phases at a time. We have used $\sigma_{BS} = \sigma_{BP} = \sigma_{PS} = 0.044$. Note that the exact values of σ_{BP} and σ_{SP} are not known and hence have been assumed to be equal to σ_{BS} . Later, we will examine the effects of varying σ_{PS} , σ_{BP} , etc. Experimental data from Bryant *et al.* (Bryant *et al.* (2003)) appear as red dots in the plot above and are in good agreement with our coexistence lines.

The total Gibbs free energy per unit contour length can be obtained by adding the contributions from the elastic and entropic components:

$$\begin{aligned}
 G_j &= G_j^e + G_j^f \\
 &= -\frac{C_j}{2(S_j C_j - g_j^2)} f^2 + \frac{g_j}{S_j C_j - g_j^2} f M - \frac{S_j}{2(S_j C_j - g_j^2)} M^2 - \lambda_0^j f - \kappa_0^j M + \frac{Q_j}{\beta A_j} - \frac{\ln Q}{L}.
 \end{aligned}
 \tag{2.3}$$

We obtain the coexistence curves for B- and S-DNA, B- and P-DNA and for S- and P-DNA by equating the respective free energies per base pair. The results are shown in figure 2.1 as solid lines with a triple point for B-, S- and P-DNA occurring at $f = 120$ pN, $M = 24$ pNnm. Now, we calculate the range of force and torque over which the structural transitions occur. We call these ranges ‘widths’, henceforth. We employ the Zimm-Bragg theory for helix-coil transitions (Grosberg *et al.* (1995)) to compute the widths. Consider a point on the B-S coexistence line wherein B- and S-DNA are in equilibrium with each other as we go from $B \rightarrow S$. We define two Zimm-Bragg parameters $s = e^{-\beta g_{BS}(f, M)}$ and $\sigma = e^{-\beta f_{BS}}$ where $g_{BS}(f, M) = b_B(G_S(f, M) - G_B(f, M))$ is the change in free energy per base pair and f_{BS} is the energy of one junction between the B- and the S- phase of DNA. Note that a conversion factor of $b_B = 0.34$ nm/bp is required to convert free energy per unit reference contour length to free energy per base pair. The fraction of S-DNA is given by $r_S(f, M) = 1/2 + (s - 1)/[2\sqrt{(s - 1)^2 + 4s\sigma^2}]$ (Grosberg *et al.* (1995)) and is obtained from an Ising-like model for the structural transition. The values of σ_{BS} are archived in (Argudo & Purohit (2014a)), however, the values of σ_{SP} and σ_{BP} are not known, which is why each has been assumed to be equal to σ_{BS} for the purpose of obtaining the phase diagram in figure 2.1. The force and torque widths

for the B-S coexistence curve are defined to be $\Delta f = \left. \frac{\partial f}{\partial r_S} \right|_{s=1}$ and $\Delta M = \left. \frac{\partial M}{\partial r_S} \right|_{s=1}$. The slope of the force f versus the fraction r_S curve for a force driven structural transition at $s_S = 1$, (which is equivalent to saying $\Delta g_{BS} = 0$ or $r_S = 1/2$), is directly proportional to the fluctuation of the force Δf during the phase transition. Analogous reasoning holds for a torque driven transition in which we are interested in ΔM . A similar procedure is followed for the P-S and B-P coexistence lines as well. These widths are shown in figure 2.1 for each transition. Our phase diagram is in good agreement with the experimental data reported by Bryant *et al* (Bryant *et al.* (2003)).

The kinematic quantities, stretch λ_j and twist κ_j , conjugate to the force f and torque M respectively, can be obtained from the free energy function for the j phase. It is worth pointing out that the above expressions are valid in the limit when the force is high enough to suppress large bending fluctuations i.e., the terms of $O(1/Q_j^2)$ are neglected.

$$\begin{aligned}\lambda_j &= -\frac{\partial G_j}{\partial f} = \lambda_0^j + \left[\frac{C_j f - g_j M}{S_j C_j - g_j^2} \right] - \frac{1}{2\beta Q_j}, \\ \kappa_j &= -\frac{\partial G_j}{\partial M} = \kappa_0^j + \left[\frac{S_j M - g_j f}{S_j C_j - g_j^2} \right] + \frac{M}{4\beta A_j Q_j}.\end{aligned}\tag{2.4}$$

2.2 Partition function for three phase coexistence

In this section, we build upon the foundational ideas behind the Zimm-Bragg model (Grosberg *et al.* (1995)) and apply it to our 1-D continuum model of DNA. Conventionally, the Zimm-Bragg model has been used to study helix-coil transitions in macromolecules, especially polymer chains. A potent feature of this model is a closed-form transfer matrix that enables us to conveniently sum the partition function of a macromolecule over the admissible micro-states and calculate the fraction in helix (coil) phase. We have already utilized these expressions to compute the widths of the coexistence lines along the force and torque axes in the previous section. Now, we want to admit the possibility that three phases of DNA can coexist at some force and torque as they do at a triple point in the phase diagram. We conjecture an analog of the Zimm-Bragg model with three distinct phases viz. B-, S- and P-DNA with respective interfaces between them. Our ensemble is a molecule with fixed number of base-pairs and each of them can be in either B-,S- or P- state; the preference for each is dictated by free energy and interfacial energy considerations. We propose the following rules for calculating the partition function for one molecule.

- We assume the ground state be B-DNA and attach our energetic datum to it. Consequently, the statistical weight of B-DNA is 1.

- The statistical weight of S- and P-DNA is $s_S = e^{-\beta\Delta g_{BS}}$ and $s_P = e^{-\beta\Delta g_{BP}}$, respectively where $\Delta g_{BS} = b_B(G_S(f, M) - G_B(f, M))$ and $\Delta g_{BP} = b_B(G_P(f, M) - G_B(f, M))$ are the free energy differences per base pair between S- and B-DNA and P- and B-DNA ($b_B=0.35$ nm/bp).
- There are three possible interfaces B-S, S-P and B-P, each of which has an energy associated with it that must be penalized. Every interface has a statistical weight of $\sigma_{ij} = e^{-\beta\Delta\gamma_{ij}}$ where $i \neq j$ and $i = B, S, P$ and $j = B, S, P$ and $\Delta\gamma_{ij}$ is the free energy of the interface between i and j phase. If the interfacial energy is small i.e. ($\sigma_{ij} \sim k_B T$), there could be multiple interfaces throughout the length of the molecule and the associated sigmoidal transition is broad with a large variance or width. On the contrary, if the interfacial energy is large ($\sigma_{ij} \gg k_B T$), then there exists only one interface and the corresponding sigmoidal transition proceeds via a steep jump, or it has a small width.

The partition function Z_{n+1} for a molecule with $(n + 1)$ base pairs can be obtained by raising the transfer matrix to power n :

$$Z_{n+1} = \sum_{k \in \Omega_{n+1}} e^{-\beta E_k} = \begin{bmatrix} 1 & s_S & s_P \end{bmatrix} \mathbf{T}^n \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad (2.5)$$

where \mathbf{T} is the transfer matrix given by

$$\mathbf{T} = \begin{bmatrix} P(B|B) & P(S|B) & P(P|B) \\ P(B|S) & P(S|S) & P(P|S) \\ P(B|P) & P(S|P) & P(P|P) \end{bmatrix} = \begin{bmatrix} 1 & \sigma_{BS} s_S & \sigma_{BP} s_P \\ \sigma_{BS} & s_S & \sigma_{SP} s_P \\ \sigma_{BP} & \sigma_{SP} s_S & s_P \end{bmatrix}. \quad (2.6)$$

In the above $P(S|B)$ represents the probability that the current base-pair is in the S-DNA state, given that the previous base-pair is in the B-DNA state and similarly for $P(P|B)$, etc. In the limit $n \rightarrow \infty$ the behavior of the system is governed by the largest eigenvalue of \mathbf{T} (say Λ) and the corresponding eigenvector is used to calculate the fractions of B-, S- and P-DNA states as in the classical Zimm-Bragg theory. Thus, if $[\psi_B \ \psi_S \ \psi_P]^T$ is the eigenvector, then

$$\begin{bmatrix} 1 & \sigma_{BS} s_S & \sigma_{BP} s_P \\ \sigma_{BS} & s_S & \sigma_{SP} s_P \\ \sigma_{BP} & \sigma_{SP} s_S & s_P \end{bmatrix} \begin{bmatrix} \psi_B \\ \psi_S \\ \psi_P \end{bmatrix} = \Lambda \begin{bmatrix} \psi_B \\ \psi_S \\ \psi_P \end{bmatrix}. \quad (2.7)$$

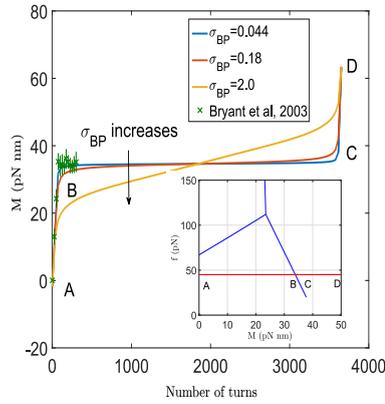


FIGURE 2.2: Torque-rotation curves for a B-DNA to P-DNA transition at constant force. The optical bead attached to one end of the molecule is twisted while maintaining a constant tension of 45 pN until B-DNA converts to P-DNA. Most of the twist takes place at constant torque. The lines in different colors correspond to different σ_{BP} values. The inset shows the trajectory on an $f - M$ plane.

Let the fractions of B-,S- and P-DNA be denoted by r_B, r_S and r_P , respectively. We follow the standard scheme given in (Grosberg et al. (1995)) to get

$$\begin{bmatrix} r_B(f, M) \\ r_S(f, M) \\ r_P(f, M) \end{bmatrix} = \begin{bmatrix} \psi_B^2 \\ \psi_S^2 s_S \\ \psi_P^2 s_P \end{bmatrix} \quad (2.8)$$

We already know the functional dependence of $g_{BS}(f, M)$, $g_{BP}(f, M)$, consequently of s_S , s_P , on the force f and torque M applied on the molecule. In the most general case, we expect that σ_{ij} is also a function of force and torque. While there are several estimates for σ_{BS} (Argudo & Purohit (2014a); Rouzina & Bloomfield (2001a)), the values of σ_{BP} and σ_{PS} are not known.

In order to estimate σ_{BP} , we use the experimental data from Bryant *et al* (Bryant et al. (2003)) in which a bead attached to a torsionally unconstrained DNA molecule is twisted to convert B-DNA to P-DNA while the tension is maintained constant throughout using a feedback mechanism. The transition from B-DNA to P-DNA is highly cooperative and takes place at a constant torque (32 pNnm) accompanied by a sudden jump in the twist which is characteristic of a first-order phase transition. The corresponding trajectory on an $f - M$ plane is a straight line parallel to torque axis (see inset of figure 2.2). Now to estimate σ_{BP} , we pick the experimental data ($\Delta M = 2$ pNnm at $f = 45$ pN and $M = 34$ pNnm) from Bryant *et al.*(Bryant et al. (2003)). We substitute these values in $\Delta M = \frac{\partial M}{\partial r_P}|_{s=1}$ to get $\sigma_{BP} = 0.18$. The resulting torque-rotation curve is plotted in figure 2.2 together with the data from Bryant *et al.*(Bryant et al. (2003)). Using our model we predict that almost 3.7×10^3 turns are required to transform 5100 nm of B-DNA completely into P-DNA, the corresponding experimental value being 4.0×10^3

(Bryant *et al.* (2003)). A notable observation is that most of the twist goes into changing the phase of the molecule from B-DNA to P-DNA. In figure 2.2 we also explore the effects of the variation of interfacial energy which leads to changes in σ_{BP} . A large $\sigma_{BP} \sim 2.0$, corresponds to a small interfacial energy, which leads to increase in the range of force (or width of the coexistence line along the force axis) over which the structural transition occurs, which means there is a possibility of multiple interfaces throughout the molecule. On the other hand, a small $\sigma_{BP} \sim 0.044$ corresponds to a large interfacial free energy and the associated structural transition is sharp.

We now apply this model to study the response of a torsionally constrained DNA molecule. For a prototypical experiment described in figure 2.1 (inset), the sum of the displacements of the three phases is equal to the externally imposed displacement and the sum of the twists of the individual phases is zero due to the torsional constraint, thus

$$\begin{aligned} r_B \lambda_B + r_S \lambda_S + r_P \lambda_P &= \lambda_{ext}, \\ r_B \kappa_B + r_S \kappa_S + r_P \kappa_P &= 0. \end{aligned} \tag{2.9}$$

The stretches ($\lambda_B, \lambda_S, \lambda_P$), twists ($\kappa_B, \kappa_S, \kappa_P$) and fractions (r_B, r_S, r_P) of the individual phases are functions of force f and torque M , the explicit expressions for which have been computed in eqn. (2.4) and eqn. (2.8). Then, for an imposed λ_{ext} the two equations given above are two non-linear equations in two unknowns, f and M . We solve these equations and plot the results in figure 2.3. Figure 2.3a shows the force-extension curve, figure 2.3b shows the corresponding trajectory on $f - M$ plane and figure 2.3c shows how the fraction of base-pairs in each phase evolves with the extension λ_{ext} . In the beginning, the whole molecule is in its ground state, i.e. B-DNA, which transforms into a mixture of B- and S-DNA which, subsequently, converts to a mixture of S- and P-DNA while passing through the triple point as shown in figure 2.3a and figure 2.3c. These two transitions have been experimentally observed using fluorescence microscopy (King *et al.* (2016)). A bulk of the extension, characterized by a 70% jump in the contour length of the molecule, takes place at the triple point (24 pNnm, 120 pN) of B-, S- and P-DNA (see figure 2.3b). A linear trend in the variation of the fractions of various phases at the triple point in figure 2.3c is a direct consequence of the fact that at fixed force and torque the linking number of the molecule is conserved. We observe that before entering the triple point, the molecule is a mixture of B- (90%) and S-DNA (10%) which is converted to a mixture of S-DNA (80%) and P-DNA (20%) which is exactly what has been reported by King *et al.* (King *et al.* (2016)). The sigmoidal plateau observed in the force-displacement curves corresponds to the extension of the molecule at the triple point. To demonstrate the utility of our formulation, we investigate the effect of the

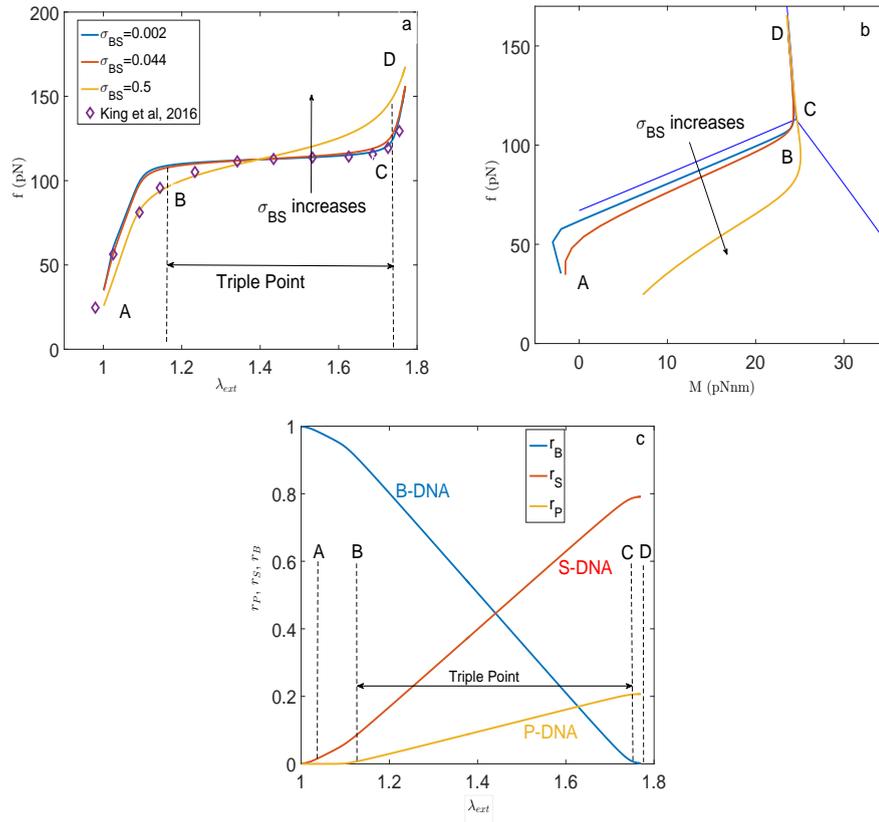


FIGURE 2.3: Structural transitions in a torsionally constrained DNA molecule. (a) Force-extension curve, (b) $f - M$ plane, (c) Fractions of B-,S- and P-DNA. The points labeled A,B,C,D in each panel have the same λ_{ext} . At B the molecule enters the triple point and at C it exits the triple point as extension λ_{ext} increases. Notice that a mixture of B- and S-DNA transforms into a mixture of S- and P-DNA. In (a), as σ_{BS} increases, the transition width increases, as expected. This is shown using different color trajectories in (a) and (b).

variation of σ_{BS} on the response of the molecule. As expected, the transition becomes more diffuse on the force axis as σ_{BS} increases.

2.3 Effect of electrostatic interactions

Our objective, in this section, is to present a systematic approach to account for the electrostatic interactions between a negatively charged DNA molecule and the ions present in the surrounding media. We assume the DNA molecule to be a polyelectrolyte which dissociates in an aqueous solution to form a polyanion wherein the phosphate groups attached to the sugar rings form the repeated electrolyte group. The behavior of such a charged polyion is markedly different from a neutral polymer, especially in polar solvents, and varies as the concentration of ions in the solution changes.

The implications of changing the ion concentration of the solution are immense and various experiments have been carried out to understand its effect on the overstretching transition in a torsionally unconstrained DNA molecule (X. Zhang et al. (2012); King et al. (2013)). As the ion concentration increases the mechanism for overstretching changes from a force induced melting leading to formation of ss-DNA (single stranded DNA) to a highly cooperative transition leading to S-DNA. The interactions between the cations in the solution and the negatively charged DNA backbone stabilizes S-DNA at higher ion concentration. The melting temperature of DNA also increases as the ion concentration increases and is related linearly to the logarithm of the concentration (X. Zhang et al. (2012)). Since base-pair disruption is highly sensitive to the ion concentration the overstretching transition could be non-hysteretic or hysteretic depending on ion concentration (King et al. (2016); X. Zhang et al. (2012)).

In torsionally constrained DNA, there are two transitions – first in which B-DNA converts to a mixture of B- and S-DNA (or melted M-DNA), and second in which the mixture of B- and S-DNA (or M-DNA) undergoes overstretching at the triple point (110 pN) to a mixture of S- and P-DNA – while conserving linking number throughout (King et al. (2016)). The ion concentration governs the competition between the formation of S-DNA and M-DNA, the former being preferred at higher ion concentration. A notable observation pertaining to the second overstretching transition at 110 pN is the non-monotonic trend in stability of P-DNA versus ion concentration (King et al. (2016)). This is in contrast to S-DNA, whose stability increases as ion concentration increases. We seek a quantitative explanation for these experimental observations by accounting for electrostatics in solution.

We use the Poisson-Boltzmann equation to describe the electrostatics in solution and employ the polyelectrolyte model of DNA described in (Frank-Kamenetskii (n.d.)) to get the contribution of ion concentration to the Gibbs free energy. This treatment presumes a DNA molecule to be an infinite rod of radius a carrying a negative (non-dimensional) charge per unit length, q , immersed in a dielectric continuum, which in our case is an aqueous solution of pH 7 with dielectric constant $D = 80$ (for further information see Frank-Kamenetskii (n.d.)). To describe the electrostatics of one phase of DNA, we need two constants:

1. The dimensionless charge $q = l_B/(b/2)$ where $l_B = e^2/Dk_B T$ is Bjerrum length of water at 300 K, e is the charge of an electron and $b/2$ is half the distance between successive base pairs. Thus, big changes in the dimensionless charge density due to structural transitions in DNA enter through a change in b . For B-DNA, $q_B = 4.2$ (Frank-Kamenetskii (n.d.)).

2. The radius of the molecule a . For B-DNA the radius is $a_B = 1\text{nm}$ (Frank-Kamenetskiĭ (n.d.)).

Since the magnitude of elastic stretching for individual phases is extremely small even at forces as high as 100 pN ($(C_B f / (S_B C_B - g_B^2)) \approx 0.07$ at $f = 100$ pN), we can safely ignore their effect on q . The treatment above presumes the negative charge to be independent of the ion concentration. More generally, the charge could depend on the interplay between the ionization constant (pK_a) of the DNA base and pH of the solution (Grosberg et al. (1995)). We introduce a dimensionless potential $u = eU/k_B T$ and employ the Poisson-Boltzmann equation in cylindrical geometry to model the electrostatics for such a rod (see details in A.2. A full list of symbols used here is given in A.3.):

$$u''(y) + r^{-1}u'(y) = -4\pi l_b(c^+(y) - c^-(y)). \quad (2.10)$$

Here y is the distance from the center of the cylinder and $c^+(y)$ and $c^-(y)$ are the concentrations of positively and negatively charged ions which are functions of y . We assume that the solution consists of only monovalent ions such as K^+ , Na^+ and Cl^- . Following standard assumptions of the Poisson-Boltzmann theory (Frank-Kamenetskiĭ (n.d.)), $c^p(y) = c_0 e^{-pu(y)}$ where $p = -, +$, and c_0 is the concentration of both species at infinity, and u is dimensionless. It is now possible to pose a boundary value problem to describe the above system.

$$\begin{aligned} u''(y) + r^{-1}u'(y) &= \chi^2 \sinh u(y), \\ u'(a) &= 2q/a, \quad u(R) = 0, \quad R \rightarrow \infty, \\ \frac{1}{\chi} &= r_D = \frac{1}{\sqrt{8\pi l_B c_0}}. \end{aligned} \quad (2.11)$$

We note here that since we model the DNA molecule as a homogeneous elastic rod with a constant charge per unit length we cannot capture sequence dependent responses. However, this way of accounting for the electrostatics predicts important aspects of the B-DNA to Z-DNA transition (Frank-Kamenetskiĭ (n.d.)) and the dependence of DNA melting temperature on ion concentration (X. Zhang et al. (2012)).

The electrostatic free energy of the cylindrical rod G^{el} per base pair can thus be calculated by charging the model cylinder incrementally from 0 to q as follows:

$$G^{el} = -2k_B T \int_0^1 u(a, tq, c_0) dt. \quad (2.12)$$

The accuracy of the theory is contingent on precisely knowing the values of a and q for a particular phase of DNA. The Gibbs free energy function can now be conveniently

decomposed into elastic, entropic and electrostatic components:

$$G_j(f, M, c_0) = G_j^e(f, M) + G_j^f(f, M) + G_j^{el}(c_0). \quad (2.13)$$

Notice that the electrostatic component is independent of force and torque which is why adding it does not change the constitutive relations in eqn.2.4. However, the coexistence lines in the phase diagram will move depending on the ion concentration because now the free energy per base pair has an electrostatic contribution that we did not consider earlier. This electrostatic part of the free energy will change depending on the phase of the DNA.

As an application of this idea consider an experiment involving torsionally unconstrained pulling of a DNA molecule. The overstretching transition from B-DNA to S-DNA at 0.1 M occurs at 64 pN. In many previous papers it has been shown that the force at which the overstretching transition occurs is related linearly to the logarithm of the ion concentration (Rouzina & Bloomfield (2001b); X. Zhang et al. (2012)) (see the appendix for how the above model leads to the logarithmic dependence of f_{ov} on ion concentration). For a given ion concentration we can find the overstretching force f_{ov} by equating the free energy per base pair of B- and S-DNA; conversely, since the dependence of f_{ov} on the ion concentration c_0 is known from experiments to be $f_{ov} = 71 + 7 \log c_0$ (X. Zhang et al. (2012)), we can fit the constants q_S and a_S for S-DNA by equating free energies of the two phases. A rough estimate for q_S can be obtained from the definition $q = \frac{l_B}{b/2}$ where $l_B = 0.71$ nm is the Bjerrum length and $b/2$ is half the distance between successive base pairs. For S-DNA, $b = 0.34 \times 1.7$ nm which gives $q_S \approx 2.5$. Also, it has been pointed out in (X. Zhang et al. (2012)) and (Rouzina & Bloomfield (2001b)) that inter-strand distance of overstretched DNA is less than 1 nm. Using $q_S = 2.8$ and $a_S = 0.8$ nm, we plot the overstretching force as a function of ion concentration in figure 2.4 which is in excellent agreement with experimental data by Zhang *et al* (X. Zhang et al. (2012)). When we plug these constants into the expression for the electrostatic part of the free energy of S-DNA we find that it decreases as the ion concentration increases as shown in Figure 2.4 (inset). Hence, the double-stranded helical charged phosphate backbone is stabilized by the cations in the solution and the stability of S-DNA increases with increasing ion concentration (X. Zhang et al. (2012); King et al. (2016)).

Now we want to determine the consequences of the above electrostatic model on overstretching in torsionally constrained DNA. For doing so we need q_P and a_P for P-DNA which are estimated as follows. As the molecule is twisted the helix angle of the right handed B-DNA increases which leads to reduction of the radius. As the helix angle reaches 70° the base pairing gets disrupted and the corresponding helix radius at that point is 0.6 nm (Allemand et al. (1998)). We, hence, estimate $a_P = 0.65$ nm. Since, the

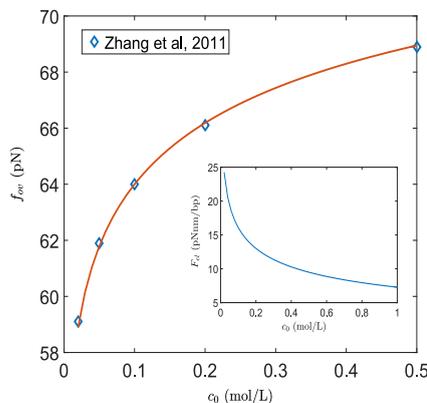


FIGURE 2.4: Plot of overstretching force for a torsionally unconstrained DNA as a function of ion concentration. We use $q_S = 2.8$ and $a_S = 0.8$ nm. In the figure blue diamonds are the experimental data from Zhang et al (X. Zhang et al. (2012)). In the inset, the electrostatic free energy for S-DNA decreases as the concentration increases.

contour length of P-DNA is 1.7 times that of B-DNA the distance between successive base pairs is $b_P = 0.34 \times 1.7 = 0.58$ nm (Bryant et al. (2003); Allemand et al. (1998)) for P-DNA. Using the definition of dimensionless charge q , we obtain $q_P = l_B/(b/2) \approx 2.5$.

We use these constants to examine the effect of electrostatic interactions on the structural transitions in a torsionally constrained DNA. We deploy our analog of the Zimm-Bragg model developed in section (3) to calculate the fractions of the various phases. The Zimm-Bragg parameters s_S and s_P entering the transfer matrix given in eqn. (2.6), now have contributions from the electrostatic free energy for each phase. Our goal is to compute the force f and torque M for prescribed extension λ_{ext} . Note that the additive decomposition posited in eqn. (2.13) implies that the constitutive equations relating the stretch and twist to force and moment given in eqn. (2.4) remain unchanged, but the relative fractions of the three phases calculated using the transfer matrix changes. The force-extension curves and the corresponding trajectories on the $f - M$ plane for a torsionally constrained DNA molecule are presented in figure 2.5a and figure 2.5b, respectively. Experimental data from King *et al.* (King et al. (2016)) for two different ion concentrations are also shown in figure 2.5a and are in excellent agreement with our model. The overall form of the curves in these two figures is similar to those in figure 2.3a and figure 2.3b.

In the most general scenario, the interfacial energies, and consequently σ_{BS} , σ_{PS} and σ_{BP} , are functions of the ion concentration of the surrounding media. We deduce, based upon the data presented in (Argudo & Purohit (2014a)), that σ_{BS} increases with increase in ion concentration but the change is small and the effect of that change is minimal and is indistinguishable in our plots (such as, figure 2.5a). As pointed out, the

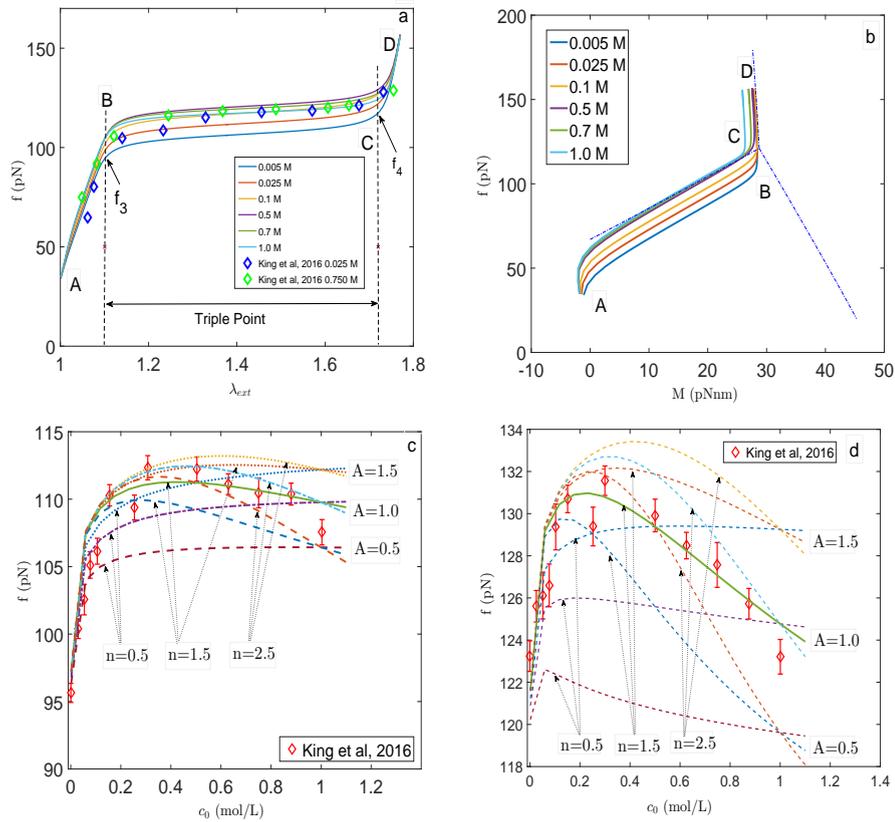


FIGURE 2.5: Effect of ion concentration on the structural transitions in torsionally constrained DNA. (a) force-extension curves, (b) $f - M$ plane, (c) f_3 versus ion concentration, and (d) f_4 versus ion concentration. These results assume $\sigma_{BS} = 0.044$ (Argudo & Purohit (2014a)), $\sigma_{BP} = 0.18$ (Bryant et al. (2003)) and $\sigma_{PS} = Ac_0^n$. Experimental data from King *et al.* (King et al. (2016)) is also shown for comparison. The force corresponding two phase transitions T_3 and T_4 is marked by f_3 and f_4 on $f - \lambda_{ext}$.

value of $\sigma_{BS} \sim 0.044$, itself is an order of magnitude smaller than the experimentally inferred value of $\sigma_{BP} \sim 0.18$. We do not have any information on the effect of ion concentration on σ_{BP} , which is why we take it to be a constant, $\sigma_{BP} = 0.18$, in the treatment henceforth. The values for σ_{SP} are not available in literature. But, we assume the interfacial energy between S- and P-DNA is related to the logarithm of the concentration $\Delta\gamma_{PS}/k_B T = a + n \log c_0$, which upon exponentiating gives $\sigma_{SP} = Ac_0^n$ for some constants A and n that must be fitted to establish agreement with experimental data. This is done as follows by focusing on the force f_3 at which the molecule enters the triple point and force f_4 at which it exits the triple point (see figure 2.5a).

In their experiments King *et al.* (King et al. (2016)) label the two critical structural transitions of interest in a torsionally constrained DNA as:

1. T_3 : Where the mixture of B- and S-DNA enters the triple point marked by [B] on force-extension curve given in figure 2.5a, the corresponding force being f_3 . In the

region A→B, only B- and S-DNA are present with the majority of the molecule ($r_B \sim 0.9$) being in the B-DNA phase.

2. T_4 : Where the mixture of B-, S- and P-DNA exits the triple point marked by [C] on the force-extension curve given in figure 2.5a, the corresponding force being f_4 . Conservation of linking number dictates that in the region C→D the fraction of P-DNA should be approximately 20% ($r_P \sim 0.2 - 0.3$) while that of S-DNA should be close to 80% ($r_S \sim 0.7 - 0.8$). These fractions are exactly the same as what have been experimentally observed in (King et al. (2016)).

The sigmoidal plateau B→C in the force-extension curve sandwiched between T_3 and T_4 denotes the triple point of B-,S- and P-DNA. We plot the variation of f_3 and f_4 with concentration for various values of σ_{SP} in figure 2.5c and figure 2.5d, respectively. As can be deduced from the plots, the parameter A can be related to the overstretching force while the parameter n controls the non-monotonicity of the resulting trend. The values $A = 1.0$ and $n = 1.5$ ($\sigma_{SP} = 1.0c_0^{1.5}$) give good agreement with the experimental data of (King et al. (2016)) for both f_3 and f_4 . As the concentration increases there is a sharp increase in f_3 and f_4 which eventually tapers off as the concentration goes beyond 0.4 mol/L. While both B-DNA and S-DNA get stabilized at higher concentration, this is not so for P-DNA whose stability has a minimum at concentration equal to 0.3 mol/L (King et al. (2016, 2013)). The non-monotonic trend observed for f_3 and f_4 can be attributed to the minima in the stability of P-DNA. Notice how the trajectories for various concentrations collapse closely onto the P-S coexistence line in figure 2.5b while a large deviation from B-S coexistence line is observed. This is because the magnitude of fluctuations in torque on the P-S coexistence line are much smaller than those on the B-S coexistence line as shown in figure 2.1.

We have found that A and n have a profound impact on the nature of the force-extension curves and even lead to non-monotonic variation of force with the ion concentration. This is counter-intuitive, since for a torsionally unconstrained extension the overstretching force increases monotonically with increase in ion concentration (X. Zhang et al. (2012, 2013)). Nonetheless, such non-monotonic trends have been experimentally observed by King et al (King et al. (2016)) and our model produces results which are in excellent agreement with experimental data . Furthermore, we find that such trends are not observed if instead of σ_{SP} , we vary σ_{BS} and σ_{BP} with ion concentration. Thus, the value of σ_{PS} plays a critical role in determining how the overstretching forces corresponding to various structural transitions vary with ion concentration.

We can further calculate the work done by the external agency in pulling a torsionally constrained DNA molecule for various values of A and n . Since the twist of the optical bead is held fixed, the work done per unit reference contour length is the area under the

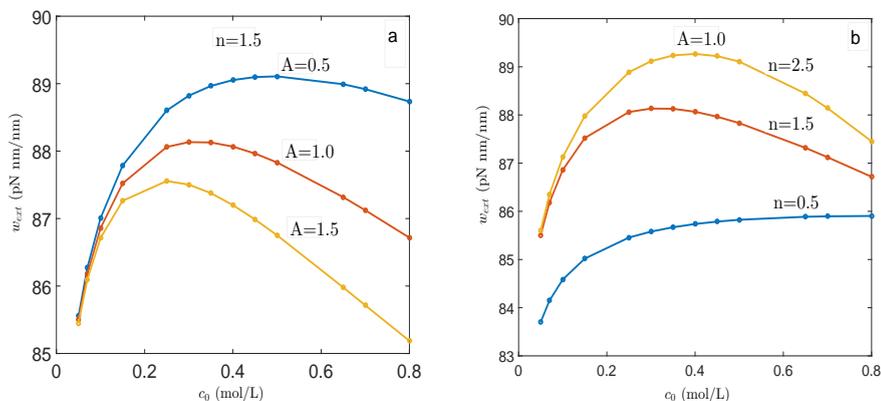


FIGURE 2.6: Non-monotonic trend in external work done per unit reference contour length. Note that the length of the DNA molecule is of the order of thousands of nm which is why a small difference in the work done will get amplified. We take $\sigma_{BS} = 0.04$, $\sigma_{BP} = 0.18$ and $\sigma_{PS} = Ac_0^n$.

force-extension curve ($W_{ext}(c_0) = \int_0^{\lambda_{ext}} f d\lambda$), and is a function of concentration only. The results of this exercise are plotted in figure 2.6. We find again that n determines the non-monotonicity of the trend. For low values of n , which implies a weak dependence of σ_{SP} on concentration, the trend is monotonic. The trend is strongly non-monotonic for higher values of n , and in particular for $n = 1.5$ has a maximum at 0.3 mol/L, which is indeed what King *et al.* (King *et al.* (2016)) have reported in their experiments. This observation can also be attributed to the low stability of P-DNA at intermediate concentrations (~ 0.3 mol/L).

2.4 Discussion

Our goal in this chapter is to quantitatively understand structural transitions in torsionally constrained DNA molecules. In a typical experiment under consideration, two strands of DNA are attached to an optical/ magnetic bead whose rotation is held fixed while it is pulled using optical/ magnetic tweezers. This leads to various structural and conformational changes in the molecule. The slender shape of the molecule allows us to model it as a fluctuating elastic rod. For such a system, the free energy can be decomposed additively into an elastic component and an entropic component. The elastic energy is quadratic in force f and torque M and takes into account stretching, torsion and the interplay between these. Since the forces in the experiment are large (> 10 pN), the bending energy can be neglected and the entropic contribution due to bending fluctuations is small. We equate the free energy per base pair for various phases and compute the phase diagram for DNA on a force-torque ($f - M$) plane. The constitutive relations relating the stretch and twist to force and torque are obtained using the Gibbs

free energy functional.

The kinematic conditions of a torsionally constrained DNA molecule dictate that the sum of the changes in the length of three phases is equal to the external displacement of the bead and the sum of the respective twists is zero. In order to calculate the fractions of various phases, we develop an analog of Zimm Bragg model for three phases. This treatment allows us to examine the effect of interfacial energy on the sigmoidal nature of force-extension curves and the respective trajectories on the force-torque plane. Our model correctly predicts the overstretching force (~ 115 pN) corresponding to the triple point and the fractions of the molecule before ($r_B = 0.9$ and $r_S = 0.1$) and after ($r_P = 0.2$ and $r_S = 0.8$) the transition (King *et al.* (2016)). We further note that the fractions of different phases vary linearly as the molecule passes through the triple point. We employ our model to study the transition of B-DNA to P-DNA and find that for $f = 45$ pN, overstretching torque is 32 pNnm and that almost 3.7×10^3 turns are required to convert a 5100 nm long B-DNA completely into P-DNA. These results are in agreement with the experimental data reported by Bryant *et al.* (Bryant *et al.* (2003)). In order to study the effect of electrostatic interactions, we assume the molecule to be an infinite rod of radius a carrying a uniform charge q per unit length. We use the Poisson-Boltzmann equation to calculate the resultant axisymmetric dimensionless electrostatic potential which is integrated to compute the electrostatic contribution to the free energy functional. We find that when the magnitude of the dimensionless potential is small, the free energy computed using Poisson-Boltzmann varies linearly with logarithm of concentration. We use the data available from previous work (Argudo & Purohit (2014a); X. Zhang *et al.* (2012); Bryant *et al.* (2003)) to get the values for σ_{BS} and σ_{BP} . We start by testing our model by computing the overstretching force for a torsionally unconstrained DNA molecule and observe that our results are in good agreement with the experimental data (X. Zhang *et al.* (2012)). Due to lack of any experimental data, we assume that the energy of the interface between S- and P-DNA is related linearly to the logarithm of the concentration ($\Delta\gamma_{SP} = a + n \log c_0$), which implies $\sigma_{SP} = Ac_0^n$ where c_0 is the ion concentration. We obtain the force-extension curves for a torsionally constrained DNA for various ion concentrations and show that our results are in good agreement with the experimental data reported by King *et al.* (King *et al.* (2016)). The effect of constants A and n is archived in figure 2.5 and we conjecture that such dependence of σ_{SP} on concentration could be a reason for the non-monotonic trends observed in the overstretching force with increasing ion concentration. We also calculate the external work done by integrating the force-extension relation and observe that its dependence on the ion concentration is monotonic for small values of n and non-monotonic for larger ones. As pointed out by King *et al.* (King *et al.* (2016)), this non-monotonic dependence in work done is related to a minimum in stability of P-DNA

at intermediate concentrations (0.3-0.4 mol/L). Our expression $\sigma_{SP} = Ac_0^n$ and its consequences in figure 2.5 and figure 2.6 are predictions that can be tested by performing new experiments.

2.5 Conclusion

The overarching objective of this chapter was to explore the implications of variation in ion concentration on the structural transitions driven by external forces in a torsionally constrained DNA molecule. Specifically, the mechanics of the overstretching transition—characterized by a 70% jump in contour length converting a mixture of B- and S-DNA converts into a mixture of S- and P-DNA—is examined. This transition proceeds via the triple point of B-, S- and P-DNA whose comprehensive understanding was the central concern in this chapter. Despite the fact that the results produced using the methodology are corroborated by experimental data at every step, it suffers from two drawbacks. The homogeneous continuum model that forms the bedrock of the framework precludes its utility in studying sequence dependent behaviors. Our methodology for electrostatic interactions works only for monovalent cations such as Na^+ and K^+ . Nonetheless, it can be used to study many aspects of structural transitions in DNA and makes specific testable predictions.

Chapter 3

Elasticity as the basis of allostery in DNA

3.1 Introduction

When a ligand binds to DNA it induces conformational changes at the binding site which could propagate to regions tens of base-pairs away, thereby encouraging or inhibiting the binding of a second ligand in those places. Such interactions between two binding agents are called *allosteric* interactions. Our focus here is on a mechanism for allostery based on elasticity of long molecules. Although we will illustrate our theory using DNA as an example, long range allosteric interactions have been documented in actin, microtubules and helical peptide chains. For example, myosin binds to actin filaments leading to suppression of the formation of cofilin clusters via allosteric signalling (Ngo et al. (2016)). Long range structural changes induced by taxol binding to microtubules inside a cell prevents cell division thus making it a potent anti-tumor agent (Mitra & Sept (2008)). The transfer of chiral stimulus triggered by a binding agent across a helical peptide chain gives the molecule an overall chiral character and is yet another instance of allostery (Ousaka & Inai (2009)). Instances of allostery in DNA have been known for decades (Ridge et al. (1994); Krugh & Young (1977)). Anti tumor drug actinomycin D binds to DNA by intercalating between the adjacent base pairs (Krugh & Young (1977)) containing a guanine base. However, in the presence of daunomycin, another anti-tumor drug, actinomycin is observed to bind to poly(dAT) DNA oligomers too. This is due to the allosteric stabilizing influence exerted by the already bound daunomycin molecule near its binding site. We will analyze allostery in dsDNA because detailed experimental and simulation results are available for it (Drsata et al. (2014); Kim et al. (2013); Koslover & Spakowitz (2009)), thus allowing quantitative comparisons with our theory.

We define the allosteric interaction energy $\Delta G = E_{12}^{\{0,p\}} - E_1^0 - E_2^p$, where $E_{12}^{\{0,p\}}$ is the free energy of the protein-DNA complex consisting of two proteins separated by distance p , and E_1^0 and E_2^p are the free energies of the protein-DNA complexes consisting of one protein. Kim *et al.* (Kim *et al.* (2013)) have conducted single molecule experiments using fluorescence techniques to measure allosteric interaction energy ΔG between two proteins on a DNA oligomer. In their paper (Kim *et al.* (2013)), DNA binding proteins are categorized as (a) proteins that bend DNA such as LacR and T7-RNAP, and (b) proteins that bind to straight DNA such as GRDBD and BamHI. Here we deal with the latter category.

We use the theory of elastic birods (Moakher & Maddocks (2005)) to develop a mechanical model for investigating protein-DNA interactions. A birod consists of two elastic strands joined by an elastic web. We represent the sugar-phosphate backbone of DNA using the outer strands and the complimentary base pairing is modeled using the elastic web. A birod model of DNA has different properties than a homogeneous rod model at short length scales; however, as the length of the birod increases the elastic properties of both models become indistinguishable (Wolfe *et al.* (2012)).

We discuss key features that distinguish our model from the state of art (Koslover & Spakowitz (2009)) worm like chain model for DNA allostery.

1. *Helical geometry*: Kim *et al.* (Kim *et al.* (2013)) discovered that the interaction energy ΔG between two proteins on DNA decays exponentially while oscillating with the periodicity of the DNA double helix. It is thereby imperative that we account for the double helical geometry of DNA which is conveniently incorporated in a birod model (Moakher & Maddocks (2005)), but is absent in a worm like chain model of DNA.
2. *Elasticity of base-pairs*: Proteins interact with DNA by altering the geometry of the double helix, such as changing the width of major/minor groove (Kopka *et al.* (1985); Drsata *et al.* (2014)). The elasticity of the basepairs, represented by the elastic web in a birod model is essential to accurately model these local deformations.
3. *Stacking energy*: Stacking energy penalizes the change in orientations of the basepairs with respect to each other. We use a stacking energy quadratic in the twist and stretch of the DNA-double helix.

In an existing model (Koslover & Spakowitz (2009)) of allostery, tension in the worm-like polymer chain to which the two proteins are bound, plays an important role in the decaying oscillatory behavior of the interaction energy ΔG . However, in the experiments of Kim *et al.* (Kim *et al.* (2013)) and simulations of Drsata *et al.* (Drsata *et al.* (2014))

the oscillatory exponentially decaying allosteric interactions on DNA are present even in the absence of tension. Here we show through an analytical model that decaying oscillatory behavior of the interaction energy can arise from the interplay between the double helical geometry and the elasticity of the base pairs.

Our model provides key insights into the structural deformations of the DNA helix, changes in the groove width when a ligand binds to DNA, and the allosteric interaction energy ΔG between two proteins on DNA. We compute the correlations between the displacement variables at the two sites of protein binding and establish their connection to the interaction energy between the two proteins. Our results are in excellent quantitative agreement with the experimental data in Kim *et al.* (Kim *et al.* (2013)) and Kopka *et al.* (Kopka *et al.* (1985)).

Numerical simulations (Drsata *et al.* (2014)) have been used to propose mechanical models for DNA allostery. These papers describe DNA using three sets of coordinates: intra-basepair coordinates buckle, propeller, opening, shear, stretch and stagger, inter-basepair coordinates tilt, roll, twist, shift, slide and rise, and major and minor groove widths. The DNA binding protein fixes some (or all) of the degrees of freedom at the site of binding resulting in deformations away from it. The energy of binding can therefore be computed. The approach is comprehensive, but computationally expensive. Drawing upon the know-how from Drsata *et al.* (Drsata *et al.* (2014)), we allow for the bending, twisting, stretching and shearing of the basepairs. Furthermore, we go beyond their numerical models by considering the mechanics of the outer strands, which as pointed out later, is crucial to getting the correct twist-stretch coupling for double-stranded DNA.

3.2 Theory

DNA comprises of two helical strands held together via complementary base-pairing. When a ligand, such as a protein or a drug, binds to DNA it exerts forces and moments on the double helix (Wiggins *et al.* (2009)) causing deformations at the base-pair level. We use the theory of birods (Moakher & Maddocks (2005)) to investigate these deformations. A birod consists of two elastic strands which interact through an elastic web. This construction makes it suitable for investigating the deformations at the base-pair level in a DNA molecule which a homogeneous rod model cannot capture (Drsata *et al.* (2014)). The latter ignores the double helical structure and the elasticity of the base pairs, both of which are crucial to the problem under consideration. In the following $(\cdot)_x$ denotes $\frac{\partial(\cdot)}{\partial x}$. Lower case letters such as a, r, β^\pm are scalars, bold lower case letters such as $\mathbf{t}^+, \mathbf{b}^-$ are vectors while bold upper case letters such as $\mathbf{R}^-, \mathbf{R}_0^+, \mathbf{Z}$ are 3×3 tensors.

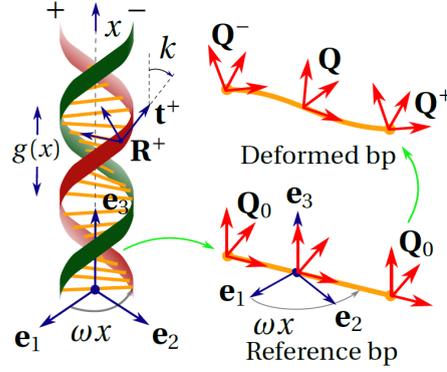


FIGURE 3.1: Birod model of DNA. The angle between the tangent \mathbf{t}^+ and \mathbf{e}_3 is k . A base pair in reference and deformed state is shown. The director frames attached to \pm ends of the base pair change from \mathbf{Q}_0 to \mathbf{Q}^\pm , respectively. The rigid rotation of the strand $\mathbf{Q} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{\frac{1}{2}} \mathbf{Q}^-$ and micro-rotation $\mathbf{P} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{\frac{1}{2}}$.

We assume the phosphate backbones comprising of phosphodiester bonds to be inextensible and unshearable elastic strands. Since these backbones consist of consecutive single bonds which allow for free rotation about the bond, we assume that they can not resist twisting moments. The base pairing is represented by the elastic web which is capable of extending, shearing, bending and twisting. In addition to the elastic energy, we consider contributions from the stacking energy which is associated with the change in orientations of the successive base pairs.

We denote the helical strands as \pm ; their positions in the reference state are denoted by \mathbf{r}_0^\pm . We use arclength parameter x to parametrize the double helix (fig. (3.1)). Thus,

$$\begin{aligned} \mathbf{r}_0^+ &= a(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) + x \mathbf{e}_3, \\ \mathbf{r}_0^- &= a(\cos(\omega x + \alpha) \mathbf{e}_1 + \sin(\omega x + \alpha) \mathbf{e}_2) + x \mathbf{e}_3, \end{aligned} \quad (3.1)$$

where $a = 1$ nm is the radius of the DNA helix, $p = 3.4$ nm is the pitch, $\omega = \frac{2\pi}{p}$ and α is the phase difference between the helices. Here we assume $\alpha = \pi$ to make the computations analytically tractable. We consider a deformed configuration where the double helix extends and twists about \mathbf{e}_3 , and its radius and phase angle also change due to binding of ligands. The deformed state of the \pm strands is denoted by $\mathbf{r}^\pm(x)$, where

$$\begin{aligned} \mathbf{r}^+(x) &= (a + r) \left(\cos(\omega x + \beta^+) \mathbf{d}_1 + \sin(\omega x + \beta^+) \mathbf{d}_2 \right) + \left(x + \int_{-\infty}^x a \xi^+ dx \right) \mathbf{e}_3, \\ \mathbf{r}^-(x) &= -(a + r) \left(\cos(\omega x + \beta^-) \mathbf{d}_1 + \sin(\omega x + \beta^-) \mathbf{d}_2 \right) + \left(x + \int_{-\infty}^x a \xi^- dx \right) \mathbf{e}_3, \end{aligned} \quad (3.2)$$

such that $\mathbf{d}_{1x} = k_3 \mathbf{d}_2$ and $\mathbf{d}_{2x} = -k_3 \mathbf{d}_1$. To gain a better physical insight, we give a visual representation of the deformation described in the above equation in fig.(3.2). We only show the deformation for the $+$ strand for clarity and indicate the strain variables

r (change in radius), β (change in phase angle) and ξ (stretch of the center-line). We assume all the displacement and strain parameters r, β^\pm and ξ^\pm vanish at $x = \pm\infty$ because the deformations caused by the proteins are local. The change in radius r , change in the phase angle β^\pm , stretches ξ^\pm , and the twist k_3 are assumed to be small ($\sim O(\varepsilon)$) such that second order terms such as r^2 and $\xi\beta^-$ are negligible. However, there could be finite rotations resulting from k_3 .

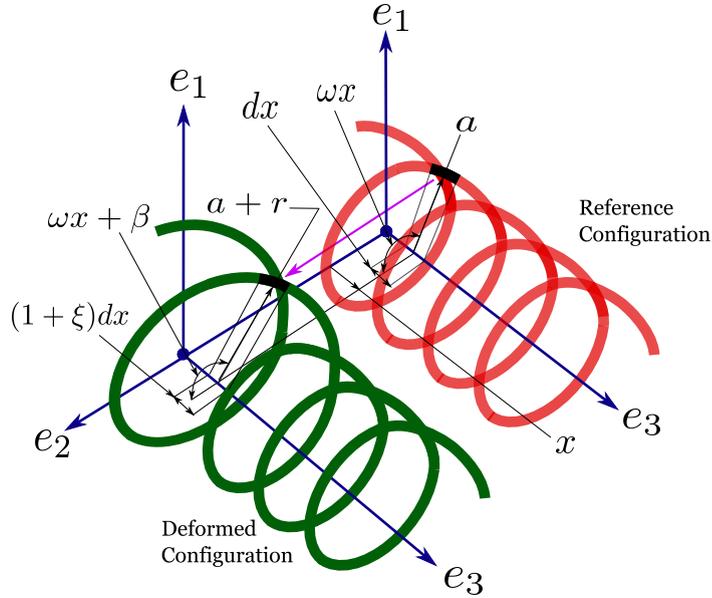


FIGURE 3.2: Distortions to the helical geometry of the + strand. We assume that the DNA remains straight after the binding of proteins. Therefore, the axis \mathbf{e}_3 remains undisturbed. The radius of the helix a changes to $a + r$, the phase angle ωx changes to $\omega x + \beta$, and the length of the infinitesimal element dx , shown in black, changes to $(1 + \xi)dx$.

As shown in fig.3.3, the birod consists of two elastic strands joined by an elastic web. We show a straight elastic ladder for easy visualization of the key forces and moments. The + strand exerts a body force \mathbf{l} and a body moment \mathbf{c} on the - strand via the elastic web. The balance laws for the two outer strands constitute the governing equations for the birod (Moakher & Maddocks (2005)). The deformation of the elastic web can be calculated once the deformation of the outer strands is known. We need to solve the following balance equations for a helical birod,

$$\begin{aligned} \mathbf{n}_x^\pm \mp \mathbf{f} + \mathbf{l} &= 0, \\ \mathbf{m}_x^\pm + \mathbf{r}_x^\pm \times \mathbf{n}^\pm + \frac{1}{2}(\mathbf{r}^+ - \mathbf{r}^-) \times \mathbf{f} \mp \mathbf{c} + \mathbf{h} &= 0, \end{aligned} \quad (3.3)$$

where \mathbf{m}^\pm and \mathbf{n}^\pm are the contact moment and contact force respectively in \pm strands. \mathbf{f} and \mathbf{c} are the distributed force and distributed moment exerted by the + strand on the - strand. \mathbf{l} and \mathbf{h} are the body force and body moment exerted by the base pairs

onto both \pm strands. In what follows, we use the position vectors for the deformed helix $\mathbf{r}^\pm(x)$ (eqn. 3.2) to compute these quantities. The constitutive relations for the forces \mathbf{n}^\pm , moments \mathbf{m}^\pm , and the force \mathbf{l} and moment \mathbf{c} transferred by the web, are discussed in the relevant subsections.

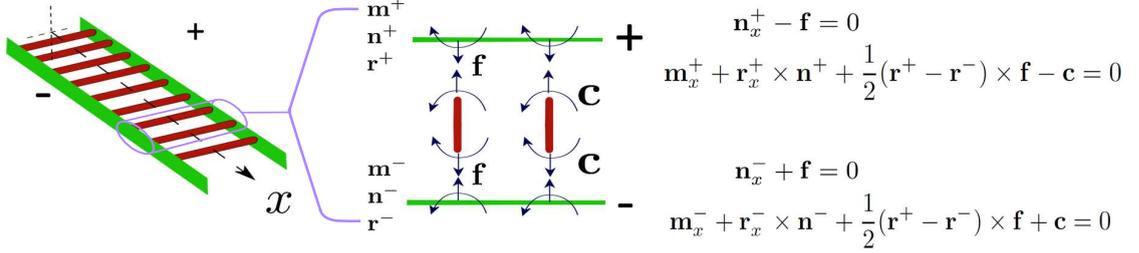


FIGURE 3.3: Free-body diagrams which establish the connection between an elastic rod and an elastic birod. We deliberately show a straight ladder instead of helical birod to illustrate the mechanics. An elastic birod comprises two elastic rods + and -. The + strand exerts a moment \mathbf{c} and force \mathbf{f} on the - strand through an elastic web. This transfer of moment and force leads to deformation of the web. In the figure \mathbf{r}^\pm denotes the position vector for \pm strands, and \mathbf{n}^\pm and \mathbf{m}^\pm denote the contact forces and contact moments in \pm strands, respectively. The force and moment balance for + and - strand constitute the governing equations (eqn. (3.3)) for the elastic birod. For further discussion see Moakher and Maddocks (Moakher & Maddocks (2005)).

3.2.1 Contact forces in the outer strands (\mathbf{n}^\pm)

The outer strands are inextensible, which means $|\mathbf{r}_x^\pm| = |\mathbf{r}_{0x}^\pm|$ yielding,

$$\omega^2 r + a\omega(k_3 + \beta_x^\pm) + \xi^\pm = 0. \quad (3.4)$$

We use the above equation to eliminate ξ^\pm from eqn. 3.2. Due to the constraint in eqn. 3.4, the contact forces \mathbf{n}^\pm enter as Lagrange multipliers.

3.2.2 Contact moments in the outer strands (\mathbf{m}^\pm)

We attach a director frame $\mathbf{R}^\pm = [\mathbf{n}_0^\pm \quad \mathbf{b}_0^\pm \quad \mathbf{t}_0^\pm]$ to each cross-section of the \pm strands, where \mathbf{n}_0^\pm , \mathbf{b}_0^\pm , and \mathbf{t}_0^\pm are the normal, binormal, and tangent in the reference state to \pm strand, respectively. $\mathbf{n}^\pm, \mathbf{b}_0^\pm, \mathbf{t}_0^\pm$ and the curvature in the reference configuration Ω_0^\pm

are computed using eqn. 3.1 as follows,

$$\begin{aligned}
\mathbf{t}_0^\pm &= \frac{\mathbf{r}_{0x}^\pm}{|\mathbf{r}_{0x}^\pm|} = \pm \sin k (-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + \cos k \mathbf{e}_3, \\
\mathbf{n}_0^\pm &= \frac{\mathbf{t}_x^\pm}{|\mathbf{t}_x^\pm|} = \mp (\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2), \\
\mathbf{b}_0^\pm &= \mathbf{t}_0^\pm \times \mathbf{n}_0^\pm = \mp \cos k (-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + \sin k \mathbf{e}_3, \\
\Omega_0^\pm &= \Omega_0 = (\mathbf{t}_0^\pm \cdot \mathbf{t}_0^\pm)^{\frac{1}{2}} = \omega \sin k.
\end{aligned} \tag{3.5}$$

Similarly, we use eqn. 3.2 to compute the Frenet-Serret frame $\mathbf{R}^\pm = [\mathbf{n}^\pm \ \mathbf{b}^\pm \ \mathbf{t}^\pm]$ and curvature Ω^\pm in the deformed state. We neglect terms higher than first order, such as $r\beta^+, \xi^-r \sim O(\varepsilon^2)$, and get

$$\begin{aligned}
\mathbf{R}^\pm &= [\mathbf{n}^\pm \ \mathbf{b}^\pm \ \mathbf{t}^\pm] = \mathbf{Z}\mathbf{R}_0^\pm(\mathbf{1} + \Theta^\pm), \\
\mathbf{Z} &= \mathbf{d}_1 \otimes \mathbf{e}_1 + \mathbf{d}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3, \\
\Theta^\pm &= \begin{bmatrix} 0 & -\theta_3^\pm & \theta_2^\pm \\ \theta_3^\pm & 0 & -\theta_1^\pm \\ -\theta_2^\pm & \theta_1^\pm & 0 \end{bmatrix} \\
\theta_1^\pm &= r\omega + a(\beta_x^\pm + k_3), \quad \theta_2^\pm = -r_x \cos k + \beta^\pm \sin k, \\
\theta_3^\pm &= \frac{-\omega r_x - a(\beta_{xx}^\pm + k_{3x})}{\omega \sin k} - \frac{(r_x \cos k - \beta^\pm \sin k) \cos k}{\sin k}.
\end{aligned} \tag{3.6}$$

The bending moment in the outer strands \mathbf{m}^\pm is proportional to the change in curvature $\kappa^\pm = \Omega^\pm - \Omega_0^\pm$ and is directed along the binormal \mathbf{b}^\pm such that $\mathbf{m}^\pm = EI\kappa^\pm\mathbf{b}^\pm$ where EI is the bending modulus of the strand. Note that the twisting moment is zero since the phosphate backbone consists of single bonds which permit free rotations.

3.2.3 Force and moment transferred by the web (\mathbf{f}, \mathbf{c})

Now, we compute the bending and twisting of the web which represents base-pairing. We attach a director frame \mathbf{Q}_0 to both + and - end of the base pair (fig.3.1).

$$\mathbf{Q}_0 = [\mathbf{e}_r \ \mathbf{e}_\theta \ \mathbf{e}_3], \tag{3.7}$$

where $\mathbf{e}_r = \cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2$ and $\mathbf{e}_\theta = -\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2$. As the birod deforms, these frames respectively get mapped to \mathbf{Q}^\pm . We compute \mathbf{Q}^\pm using the deformation of \mathbf{R}^\pm from eqn. (3.6) keeping in mind that the angles between the columns of \mathbf{R}_0^\pm and

\mathbf{Q}_0^\pm should remain constant during deformation implying $(\mathbf{R}_0^{\pm T} \mathbf{Q}_0 = \mathbf{R}^{\pm T} \mathbf{Q})$, thus,

$$\mathbf{Q}^\pm = \mathbf{Z} \mathbf{R}_0^\pm (\mathbf{I} + \Theta^\pm) \mathbf{R}_0^{\pm T} \mathbf{Q}_0, \quad \Theta^\pm \sim O(\varepsilon). \quad (3.8)$$

Now, we can compute the rigid rotation \mathbf{Q} and micro-rotation \mathbf{P} for each base pair. The micro-rotation contains information about the ‘difference’ between the rotations \mathbf{Q}^\pm (Moakher & Maddocks (2005)). This is related to the moment transferred by the base pair \mathbf{c} via an elastic constitutive relation for the web,

$$\mathbf{P} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{\frac{1}{2}} = \mathbf{Z} (\mathbf{I} + \Phi^c) \mathbf{Z}^T. \quad (3.9)$$

Here, $\Phi^c = \frac{\mathbf{R}^+ \Theta^+ \mathbf{R}^{+T} - \mathbf{R}^- \Theta^- \mathbf{R}^{-T}}{2}$ is a skew symmetric tensor. The moment transferred by the base pair is directly proportional to the Gibbs vector of \mathbf{P} . $\boldsymbol{\eta} = \tan \frac{\lambda}{2} \hat{\mathbf{k}}$ is a Gibbs rotation vector for a rotation matrix \mathbf{T} if $\mathbf{T} \hat{\mathbf{k}} = \hat{\mathbf{k}}$ and $1 + 2 \cos \lambda = \text{tr} \mathbf{T}$. In our case, the Gibbs vector of \mathbf{P} is $2\boldsymbol{\eta} = 2\mathbf{Z}\bar{\boldsymbol{\eta}} = \mathbf{Z}\phi^c$, where ϕ^c is the axial vector of skew symmetric tensor Φ^c . Note that in the reference state, $\boldsymbol{\eta}_0 = 0$ since $\mathbf{P}_0 = (\mathbf{Q}_0 \mathbf{Q}_0^T)^{1/2} = \mathbf{I}$. The rigid rotation of the base pair $\mathbf{Q} = \mathbf{P} \mathbf{Q}^-$. Here

$$\mathbf{Q} = \mathbf{Z} (1 + \Phi) \mathbf{Q}_0, \quad (3.10)$$

and $\Phi = \frac{\mathbf{R}^+ \Theta^+ \mathbf{R}^{+T} + \mathbf{R}^- \Theta^- \mathbf{R}^{-T}}{2}$ is a skew symmetric matrix. The moment exerted by + strand on the – strand by means of the elastic web, \mathbf{c} , is computed using $\mathbf{c} = \mathbf{Q} \mathbf{H} \mathbf{Q}^T \boldsymbol{\eta}$ where $\mathbf{H} = \text{diag}[H_1, H_2, H_3]$ are the elastic moduli (Moakher & Maddocks (2005)). Now, we shift our focus to the extension and shear of the web. In the reference configuration, the displacement between the two strands $\mathbf{w}_0 = \frac{\mathbf{r}_0^+ - \mathbf{r}_0^-}{2} = a \mathbf{e}_r$ which, in deformed configuration changes to $\mathbf{w} = \frac{\mathbf{r}^+ - \mathbf{r}^-}{2}$. The force \mathbf{f} exerted by + strand on the – strand is computed using $\mathbf{f} = \mathbf{Q} \mathbf{L} (\mathbf{Q}^T \mathbf{w} - \mathbf{Q}_0^T \mathbf{w}_0)$ where $\mathbf{L} = \text{diag}[L_1, L_2, L_3]$ are the elastic moduli, $\beta = \frac{\beta^+ + \beta^-}{2}$ and $\beta^c = \frac{\beta^+ - \beta^-}{2}$.

$$\begin{aligned} \mathbf{c} &= \mathbf{Q} \mathbf{H} \mathbf{Q}^T \boldsymbol{\eta} \\ &= H_1 (-ak_3 - \omega r_x - a\beta_x) \mathbf{f}_1 + H_2 \frac{(-ak_{3x} - \omega r_{xx} - a\beta_{xx})}{\omega} \mathbf{f}_2 + H_3 \left(\beta^c - \frac{a \cot k}{\omega} \beta_{xx}^c \right) \mathbf{e}_3, \\ \mathbf{f} &= \mathbf{Q} \mathbf{L} (\mathbf{Q}^T \mathbf{w} - \mathbf{Q}_0^T \mathbf{w}_0) \\ &= L_1 r \mathbf{f}_1 + a L_2 \cot k \frac{ak_{3x} + 2\omega r_x + a\beta_{xx}}{\omega} \mathbf{f}_2 - a^2 L_3 \frac{\omega^2 \beta^c + \beta_{xx}^c}{\omega} \mathbf{e}_3, \end{aligned} \quad (3.11)$$

3.2.4 Contributions from the stacking energy (\mathbf{l}, \mathbf{h})

We now consider the contributions from the stacking energy. The center line of the double helix \mathbf{e}_3 undergoes both twist k_3 and extension $\xi = \frac{\xi^+ + \xi^-}{2}$. We associate a quadratic stacking energy $E_s = K_c k_3^2 + K_e \left(\frac{\xi^+ + \xi^-}{2}\right)^2$ to penalize this change in the orientation of successive base pairs. Due to this energy, the base pairs exert a body force \mathbf{l} and a body moment \mathbf{h} on both \pm strands which are given by

$$\mathbf{l} = K_e \left(\frac{\xi^+ + \xi^-}{2}\right) \mathbf{e}_3, \quad \mathbf{h} = K_c k_3 \mathbf{e}_3. \quad (3.12)$$

3.3 Results

Now we have all the ingredients for solving the governing differential equations of a birod. Substituting these quantities in the balance laws (eqn.3.3) gives us a set of 12 differential equations. The complete procedure for solving those equations is in the supplement, however we highlight crucial points here. It follows from the governing equations that $\beta^+ = \beta^- (= \beta, \text{ say}), n_3^c = n_1 = n_2 = 0$. $\beta^+ = \beta^-$ implies $\xi^+ = \xi^- (= \xi, \text{ say})$ thereby reducing 12 equations to 6 equations in 6 unknowns $r, \beta, k_3, n_{1,2}^c, n_3$. We look for solutions of the form,

$$\begin{aligned} r(x) &= r_0 e^{-\lambda x}, \beta(x) = \beta_0 e^{-\lambda x}, \xi(x) = \xi_0 e^{-\lambda x}, \\ n_1^c(x) &= n_{10}^c e^{-\lambda x}, n_2^c(x) = n_{20}^c e^{-\lambda x}, n_3(x) = n_{30} e^{-\lambda x}. \end{aligned} \quad (3.13)$$

We substitute this form into the governing equations (eqn. 3.3) and obtain an eigenvalue problem in λ . In order to make further progress, we need the values of the elastic constants. We use $K_c = 80 \text{ pNnm}^2$, $K_e = 600 \text{ pN}$, $L_1 = L_2 = L_3 = H_1 = H_2 = H_3 = 10 \text{ pN}$. In the supplement, we show that these values yield the correct twist, stretch and twist-stretch coupling moduli for double stranded B-DNA (Singh & Purohit (2017)). Solving for the eigenvalues λ we get

$$\lambda = \pm \zeta \pm i\omega, \quad \zeta = 0.32 \text{ nm}^{-1}, \quad (3.14)$$

and the solution for the strain parameters $y_1 = r, y_2 = k_3$ and $y_3 = \beta$ is of the form:

$$y_i(x) = A_1 \mathbf{V}_1(i) e^{(-\zeta - i\omega)x} + A_2 \mathbf{V}_2(i) e^{(-\zeta + i\omega)x} + B_1 \mathbf{V}_3(i) e^{(\zeta - i\omega)x} + B_2 \mathbf{V}_4(i) e^{(\zeta + i\omega)x}. \quad (3.15)$$

where $\mathbf{V}_j(i)$ is the i^{th} component of the eigenvector corresponding to the eigenvalue j in the exponent. Clearly, the decay length ζ is only a function of the elastic parameters of

dsDNA, in agreement with the conclusion of Kim *et al* (Kim *et al.* (2013)). Note that the strain parameters are exponentially decaying while oscillating with the period ω of the double helix. We impose the boundary conditions on r and β remembering that the displacements of the strands must be continuous. For a protein binding at $x = p$,

$$\begin{aligned} \text{as } x \rightarrow \pm\infty \quad r(x), \beta(x) &\rightarrow 0, \\ \text{at } x = p \quad r(0) = r_0, \beta(0) &= \beta_0. \end{aligned} \quad (3.16)$$

We present the variation of r , k_3 and β for a protein binding at $x = 0$ for two different sets of boundary conditions in fig. 4.2. Notice the sinusoidal correlation between the local deformation of base-pairs which is in agreement with earlier work which used Monte Carlo simulations(Gu *et al.* (2015)).

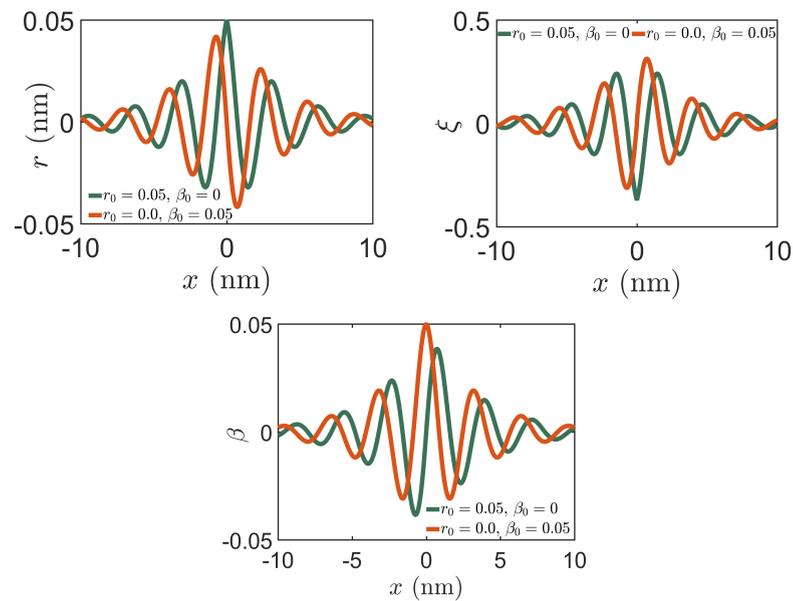


FIGURE 3.4: Variation of r , k_3 , ξ and $\beta^+ = \beta^- = \beta$ for a single protein. The red curve corresponds to the boundary conditions $\beta_0 = 0, r_0 = 0.05$ nm and the green curve to $r_0 = 0, \beta_0 = 0.05$. The decay length is $l_d = \zeta^{-1} \approx 10$ bp which is close to that documented in literature (Kim *et al.* (2013); Gu *et al.* (2015)).

We show the deformed shapes of the helices in fig. 3.5 for three cases: first when one protein binds at $x = 0$, second when two proteins bind at $x = \pm 1.5$ nm, and third when two proteins bind at $x = \pm 3.5$ nm. The boundary condition for each protein is $r_0 = 0.2$ nm, $\beta_0 = 0$. We deliberately choose large values for r_0 and β_0 to distinguish the deformed shape from the reference shape. The large configuration changes near the site of protein binding ($x = 0$) decay exponentially with distance. Note the strong overlap in the deformation fields when the distance between two proteins is 3 nm compared to 7 nm. This overlap results in an interaction energy between the two proteins which we

subsequently quantify using eqn. (4.40).

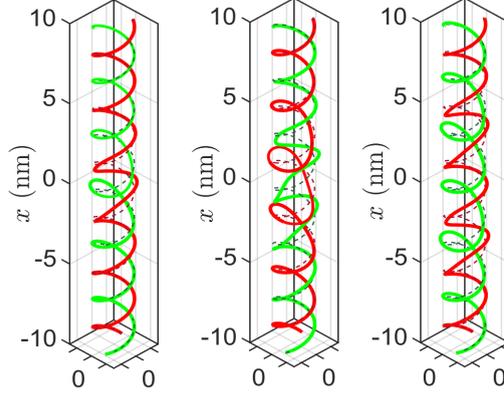


FIGURE 3.5: We show the deformed configuration of the double helix, red and green colors correspond to + and – strand, respectively. In the first figure, one protein binds at $x = 0$ with $r_0 = 0.2$ nm and $\beta_0 = 0$. In the second figure, two proteins bind at $x = \pm 1.5$ nm. In the third figure, two proteins bind at $x = \pm 3.5$ nm. Notice the overlap of deformations in the second figure which is absent in the third one. This overlap manifests itself as interaction energy between the two proteins. The dotted lines denote the corresponding undeformed configuration.

We now compute the interaction energy ΔG for two proteins. The energy functional of the double helical rod is

$$E[r, \beta, k_3] = \frac{1}{2}EI(\kappa^+)^2 + \frac{1}{2}EI(\kappa^-)^2 + \sum_{i=1}^3 \frac{1}{2}(L_i \Delta \mathbf{w}_i^2 + H_i \hat{\eta}_i^2) + K_c k_3^2 + K_e \xi^2, \quad (3.17)$$

where $\hat{\eta} = \mathbf{Q}^T \eta$ and $\Delta \mathbf{w} = \mathbf{Q}^T \mathbf{w} - \mathbf{Q}_0^T \mathbf{w}_0$. Consider two proteins, P_1 and P_2 binding at $x = 0$ and $x = p$. The interaction energy ΔG is defined as,

$$\Delta G(p) = E_{12}^{\{0,p\}} - E_1^0 - E_2^p, \quad (3.18)$$

where $E_{12}^{\{0,p\}} = E[r_{12}, \beta_{12}, (k_3)_{12}]$ is the energy of two proteins binding onto DNA at $x = 0$ and $x = p$, while $E_1^0 = E[r_1, \beta_1, (k_3)_1]$ and $E_2^p = E[r_2, \beta_2, (k_3)_2]$ are the energies of a single protein binding at $x = 0$ and $x = p$, respectively. We linearly superimpose the strain fields from each protein (r_1 and r_2 , etc) to get the resultant strain field (r_{12} , etc.) caused by two proteins simultaneously binding to DNA.

$$r_{12}(x) = r_1(x) + r_2(x - p). \quad (3.19)$$

We obtain β_{12} and $(k_3)_{12}$ similarly. We compute the interaction energy $\Delta G(p)$ as a function of the distance between two proteins p and plot it in fig.3.6 together with

experimental data from Kim *et al.* (Kim *et al.* (2013)). In excellent agreement with experiment (Kim *et al.* (2013)) and numerical simulations (Gu *et al.* (2015)), ΔG decays exponentially while oscillating with the period of the double helix (~ 10 bp).

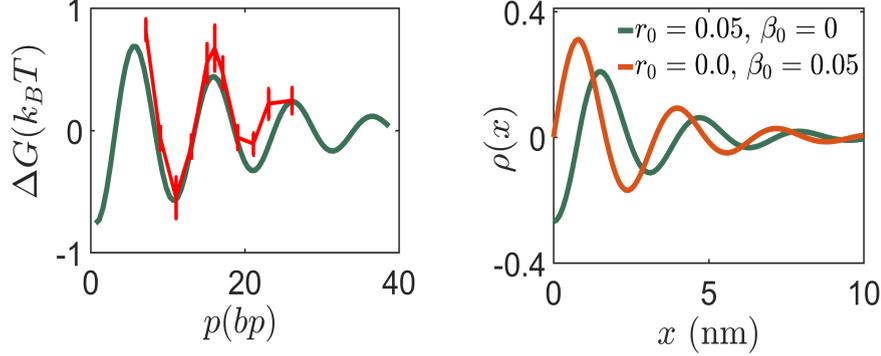


FIGURE 3.6: The first figure shows the variation of interaction energy ΔG with distance p between the two proteins P_1 and P_2 . The boundary conditions $r_1 = 0.001$ nm, $\beta_1 = 0.0045$ for P_1 and $r_2 = 0.001$ nm, $\beta_2 = -0.0045$ for P_2 give the best fit to the experimental data for ΔG (Kim *et al.* (2013)). In the second figure, we show the variation of change in groove width $\rho(x) = g(x) - \frac{r}{2}$ when a protein with boundary conditions r_0, β_0 binds at $x = 0$. The decaying sinusoidal character is documented in previous work (Kim *et al.* (2013))(Gu *et al.* (2015)). The magnitude of the change in groove width (~ 3 Å) is consistent with estimates in (Kopka *et al.* (1985)).

We justify this variation of interaction energy for a simple case as follows. Consider a strain parameter $\delta(x)$ and the associated quadratic energy potential $\mathcal{E}[\delta(x)] = \int_{-\infty}^{\infty} \frac{\delta^2(x)}{2} dx$. Similar to our strain parameters in eqn. 3.15 let us assume $\delta(x) = Ae^{-bx} \cos(\mu x)$, then

$$\mathcal{E}[\delta(x)] = \int_{-\infty}^{\infty} \frac{\delta^2(x)}{2} dx = \frac{A^2(2b^2 + \mu^2)}{4b(b^2 + \mu^2)} \quad (3.20)$$

$\mathcal{E}[\delta(x-p)] = \mathcal{E}[\delta(x)]$. Now the strain obtained by superposing two strain sources a distance p apart are $\delta_2(x) = \delta(x) + \delta(x-p)$. The energy functional corresponding to $\delta_2(x)$ is

$$\begin{aligned} \mathcal{E}[\delta_2(x)] &= \frac{A^2(2b^2 + \mu^2)}{2b(b^2 + \mu^2)} + A^2 c_1 e^{-bp} \sin(\mu p) + A^2 c_2 e^{-bp} \cos(\mu p) \\ &= \mathcal{E}[\delta(x)] + \mathcal{E}[\delta(x-p)] + \Delta G \end{aligned}$$

where $c_1 = \frac{b^3}{2b\mu(b^2 + \mu^2)}$ and $c_2 = \frac{\mu(\mu^2 + 2b^2 + pb^3 + pb\mu^2)}{2b\mu(b^2 + \mu^2)}$. It is notable how the decaying sinusoidal behavior of the interaction energy ΔG follows naturally from the functional form of the strain parameters and their eventual superposition. A cartoon illustrating this key point is presented in fig. 3.7

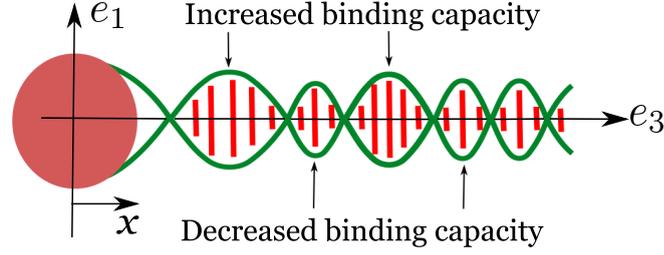


FIGURE 3.7: Eqn. 3.15 shows that the strain parameters r , β , and ξ decay exponentially while oscillating with the periodicity of the double helix. Let us assume that the protein binding at $x = 0$ increases the radius of the double helix from a to $a + r_0$. This change in radius at $x = 0$ decays exponentially while oscillating with the periodicity of the double helix, away from the binding site. Similar behavior is observed for other strain parameters, β and ξ . Due to this sinusoidal modulation of the geometry, the binding of the second protein is facilitated at some locations, while inhibited at others; this manifests as an exponentially decaying oscillatory behavior observed in the allosteric interaction energy (ΔG).

Next, we focus on the width of the groove since many proteins are known to change the width of the major/minor groove of DNA (Kopka et al. (1985); Kim et al. (2013); Hancock et al. (2013)). We define the width of the groove, $g(x)$, as follows (we do not have a major/minor groove because $\alpha = \pi$, for simplicity):

$$g(x) = \mathbf{r}^- \cdot \mathbf{e}_3(x + \frac{\pi}{2\omega}) - \mathbf{r}^+ \cdot \mathbf{e}_3(x - \frac{\pi}{2\omega}). \quad (3.21)$$

Note that in the reference configuration the groove width $g_0 = \frac{\pi}{\omega} = \frac{p}{2}$. We consider a protein binding at $x = 0$ and compute the change in groove width $\rho(x) = g(x) - g_0$ for two sets of boundary conditions, $r_0 = 0, \beta_0 = 0.02$ and $r_0 = 0.02$ nm, $\beta_0 = 0$ (see fig.3.6). The groove width ρ decays exponentially with increasing distance from the binding site while oscillating with the periodicity of the double helix. This characteristic decaying sinusoidal oscillation is documented in (Gu et al. (2015)) and is also observed experimentally (Kim et al. (2013)). It has been proposed that this change in groove width could explain the sinusoidally decaying interaction energy (notice the similarity of the two panels in fig.3.6) between two proteins bound to DNA because the binding energy of a protein binding to DNA could potentially depend on the groove width. However, we have arrived at the decaying sinusoidal variation of the interaction energy by computing the elastic energy stored in the birod without assuming any connection to the groove width. Thus, we argue that the characteristic variation in groove width and the characteristic variation of the interaction energy have the same underlying cause – the geometry and elasticity of helical DNA at the base-pair level.

To make the above point more concrete we give another analytical argument. Consider two proteins P1 and P2 binding at $x = 0$ and $x = p$, respectively. Kim et al. (Kim et al. (2013)) argue that when a protein binds to DNA it alters the groove width, which leads

to $\Delta G \propto \rho_0^1 \rho_0^2 + \rho_p^1 \rho_p^2$, where ρ_p^1 is the change in the groove width caused by protein P1 at $x = p$, and so on. On the other hand, we assume that the protein binds to DNA by fixing the radius r and phase angle β at the binding site. Let us examine whether $\Delta G \propto \alpha_1 r_{12}(p) + \alpha_2 \beta_{12}(p)$, where $r_{12}(p) = r_0^1 r_0^2 + r_p^1 r_p^2$, $\beta_{12}(p) = \beta_0^1 \beta_0^2 + \beta_p^1 \beta_p^2$, for some constants α_1 and α_2 . Here, r_0^1 is the change in radius caused by protein P1 at $x = 0$. Other quantities ($r_0^2, r_p^1, r_p^2, \beta_0^1, \beta_0^2, \beta_p^1, \beta_p^2$) are defined similarly. For simplicity, assume $\alpha_1 = \alpha_2$ and define $q(p) = r_{12}(p) + \beta_{12}(p)$. We plot $\Delta G_n = \frac{\Delta G}{|\Delta G|_{\max}}$ and $q_n = \frac{q(p)}{|q(p)|_{\max}}$ versus the distance between the two proteins p in fig. 3.8(a). We observe that while the location of peaks and valleys for ΔG_n and $q(p)_n$ coincides, the magnitudes are not identical.

Our next step is to relate the magnitudes of two quantities ΔG_n and q_n . Assume an empirical relation $\Delta G_n = y(q_n)$. We plot ΔG_n versus q_n in fig.3.8 and find that the resultant profile looks akin to $y(x) = \tanh(ax)$, $a \approx 3.0$ gives the best fit. Thus, $\Delta G_n \approx \tanh(3q_n)$. Note that for large values of p ($p > 10$ nm), the correlation function q_n is small, thence $\tanh(3q_n) \approx 3q_n$, and we recover the form similar to that used in Kim *et al.* (Kim *et al.* (2013)) (but with different strain variables) $\Delta G_n \propto q_n$. Note that we used a particular set of boundary conditions to extract the relation $\Delta G_n \approx \tanh(3q_n)$. Now, we test this relation to compute interaction energies for other sets of proteins which apply different boundary conditions in fig. 3.8(b). We observe a remarkable agreement with the free energies computed using eqn. 4.40. Thus, we have shown that the correlation function $q(p)$ can be used as a surrogate for the interaction energy as Kim *et al.* (Kim *et al.* (2013)) did. Evaluating the correlation function $q(p)$ involves measuring displacement variables at two binding sites which can in turn be related to the free energy using the above scheme.

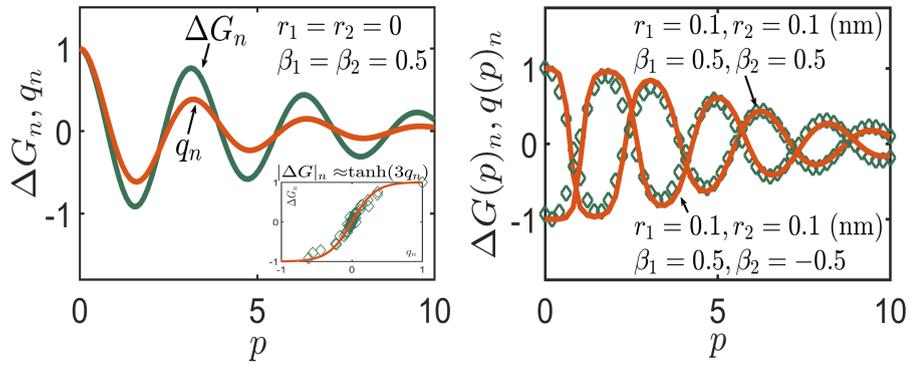


FIGURE 3.8: Consider two proteins P1 and P2 binding at $x = 0$ and $x = p$, respectively. The first figure shows the variation of normalized free energy $\Delta G(p)_n = \frac{\Delta G}{|\Delta G|_{\max}}$ and correlation function $q(p)_n = \frac{q(p)}{|q|_{\max}}$ with the distance between the two proteins p . The correlation function $q(p) = (r_0^1 r_0^2 + r_p^1 r_p^2) + (\beta_0^1 \beta_0^2 + \beta_p^1 \beta_p^2)$ where r_p^1 is the change in radius caused by protein P1 at $x = p$. The boundary conditions for the two proteins are given in the figure. We find that the peaks and valleys of ΔG_n and q_n coincide, however the magnitudes are not identical. We find that the magnitudes are related as $\Delta G_n \approx \tanh(3q_n)$, as shown in the inset. We test this empirical relation for two different sets of boundary conditions and find a remarkable match. The diamonds denote the free energies computed using eqn. 4.40 and the solid line denotes the free energy computed using the normalized correlation function $q(p)_n$. This exercise shows that the correlation functions can be used as a surrogate for free energies.

To conclude, the analysis presented in this chapter ties together the continuum theory (Moakher & Maddocks (2005)), experiments (Kim et al. (2013)) and numerical simulations (Drsata et al. (2014); Gu et al. (2015)) and provides useful insights into the allosteric interactions between two proteins binding to a DNA molecule. The results from the model predict that the interaction energy (eqn. (4.40)) for two proteins bound to DNA decays exponentially while oscillating with the period of the DNA double-helix. The decay length depends only on the elastic characteristics of the web while the oscillatory behavior is inherited from the underlying double-helical geometry. It has been shown that the strong correlation of interaction energy with the changes in groove widths caused by the proteins is rooted in elasticity and geometry of DNA. However, the model suffers from some shortcomings. Existing models such as (Drsata et al. (2014)) relying on numerical simulations can account for the stacking energy in a more comprehensive way compared to our approach which assumes it is quadratic in the twist and stretch of the centerline. Also, these models (Drsata et al. (2014)) can account for a wider variety of boundary conditions applied by a protein, owing to more variables describing the DNA structure. While the results from our model agree with the experimental observations, the outcomes from existing models in literature such as (Gu et al. (2015)) agree as well. The main strength of the birod model compared to the existing models is twofold: i) it accounts for the mechanics of the outer strands, and ii) computationally efficient. The elastic helical birod based techniques developed in this chapter could potentially be

applied to other molecules which have a double helical geometry such as dsRNA, and coiled-coil intermediate filaments.

Chapter 4

Allosteric interactions in a birod model of DNA

In this chapter, the allosteric interactions between two proteins on DNA are studied using an elastic birod (Moakher & Maddocks (2005)). The objective is to compute the allosteric interactions energy ΔG given by,

$$\Delta G(a) = E_a^{12} - E^1 - E^2, \quad (4.1)$$

where E_a^{12} is the free energy of two proteins P1 and P2, separated by distance a , and E^1 and E^2 is the free energy when P1 and P2 respectively, bind to DNA individually. Experimental measurements by Kim *et al* (Kim *et al.* (2013)) show that the interaction energy ΔG decays exponentially with the separation between the proteins a while oscillating with the periodicity of double helix. The characteristic length of the aforementioned exponential is referred to as decay length. The elastic birod model is used to compute the interaction energy and results are shown to quantitatively match with the experimental data by Kim *et al* (Kim *et al.* (2013)) after fitting certain parameters. Furthermore, we compute the dependence of the decay length on the GC content of the DNA molecule.

4.1 Strategy to compute interaction energy

In this section, we give a concise blueprint of our strategy to compute the interaction energy for two proteins binding to DNA. We assume elastic deformations throughout. When a protein binds to DNA it causes local bending and twisting. We assume that the resulting twist and curvatures are small. These curvatures could possibly add up

to produce large displacements and rotations. The two phosphate backbones of DNA constitute the helical outer strands which are out of phase by a phase angle $\alpha = 2.1$ radians. We assume these backbones to be inextensible. These outer strands consist of sugar phosphate single bonds. Thus, we assume that they can not support twisting moments. The inextensibility of the outer strands is a strong geometrical constraint which induces a change in the radius and phase angle between the two helices when a protein causes local deformations. We assume that these changes are small and of the same order as the curvatures.

We give a stepwise procedure to do the calculation and in the following sections we label each step. We give a pictorial representation for this procedure in fig. B.1 (in Appendix B).

1. We begin by assuming a form of displacement for each of the outer strands which are assumed to be inextensible and unshearable.
2. This displacement is used to calculate the tangent, normal and binormal to the deformed configuration of the outer strands thereby obtaining the rotation matrix attached to the deformed configuration of the outer strands.
3. Then, the deformation and rotation of the outer strands are used to calculate the extension, shear and rotation of the web.
4. At this point, these quantities are substituted into the balance laws for the birod. We, then, seek non-zero solutions to the resulting system of differential equations. This leads to an eigenvalue problem.
5. In the next step, the boundary conditions are applied to evaluate the constants in the solution to the eigenvalue problem.
6. This process is carried out to compute the energy of a one-protein-DNA complex. The displacement fields for two proteins can be computed using superpositions, which are then used to compute the energy for a two-protein-DNA complex.
7. Finally, we subtract the two energies obtained in the previous step to get the interaction energy. We find that it takes the form of a decaying exponential oscillating with the periodicity of the underlying DNA helix.

We demonstrate the above procedure for a straight non-helical birod (called a ladder) in sec. 1 of the appendix. Our calculations show that interaction energy for two defects on a ladder decays exponentially with the distance between them. In the next section, we focus on how the geometry of the DNA helix leads to an interaction energy which decays exponentially while oscillating with the periodicity of the helix.

4.2 Interaction energy for two DNA binding proteins

4.2.1 Step 1: Deformation of the outer strands

DNA consists of two helical strands with radius $b = 1$ nm and pitch $p = 3.4$ nm, out of phase by $\alpha = 2.1$ radians, wrapped around a common axis as shown in fig. 4.1. We follow the notation used by Moakher and Maddocks (Moakher & Maddocks (2005)) and refer to the two strands as \pm . The undeformed state of the outer strands denoted by $\mathbf{r}_0^\pm(x)$ is a helix with a constant radius and pitch. We choose to parametrize both the curves by arclength parameter x . Here, $\omega = \frac{2\pi}{p}$ and k is the characteristic angle of the helix such that $\tan k = \frac{2\pi b}{p} = \omega b$.

$$\mathbf{r}_0^+ = b \left(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2 \right) + x \mathbf{e}_3,$$

$$\mathbf{r}_0^- = b \left(\cos(\omega x + \alpha) \mathbf{e}_1 + \sin(\omega x + \alpha) \mathbf{e}_2 \right) + x \mathbf{e}_3.$$

Let us now focus on the two strands separately. The calculations for the $+$ strand are given in this section while the results for the $-$ strand are given in the appendix. We posit a form of displacement wherein the radius of the helix changes and its axis is allowed to take arbitrary shapes within the ambit of the assumptions specified in section 4.1. Here $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ denotes the standard spatial reference frame and \mathbf{e}_3 is along the common axis of the two helices \pm in the *reference* configuration. This common axis in the *deformed* configuration is defined by the set of orthogonal directors $[\mathbf{d}_1(x), \mathbf{d}_2(x), \mathbf{d}_3(x)]$. The displacement fields which define the undeformed and deformed configuration are,

$$\begin{aligned} \mathbf{r}_0^+ &= b \left(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2 \right) + x \mathbf{e}_3, \\ \mathbf{r}^+(x) &= (b+r) \left(\cos(\omega x + \beta^+) \mathbf{d}_1 + \sin(\omega x + \beta^+) \mathbf{d}_2 \right) + \int_0^x dx (1 + b\xi) \mathbf{d}_3. \end{aligned} \quad (4.2)$$

where

$$r = r(x), \quad \beta^+ = \beta^+(x), \quad \xi = \xi(x).$$

Here r is the change in the radius of the helix, β^+ is the change in the phase of the $+$ strand, and ξ can be considered as a stretching of the axis of the helix.

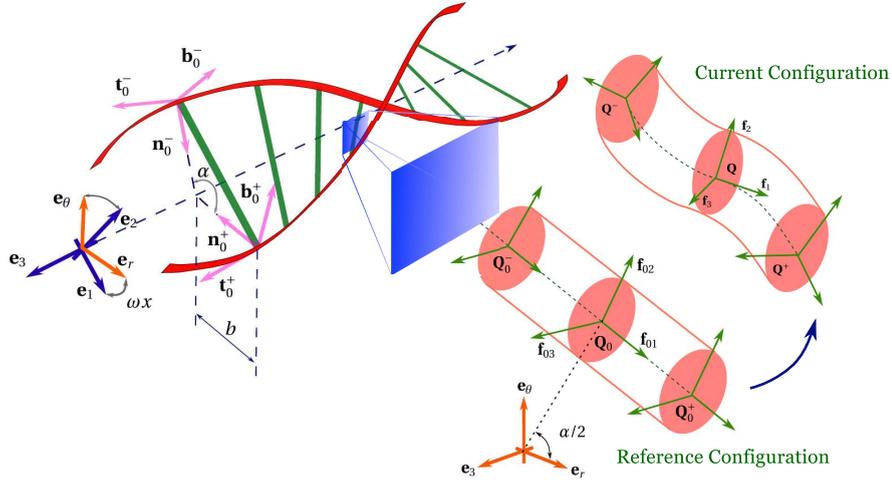


FIGURE 4.1: A DNA molecule as a double helical elastic birod is shown on the left. The phosphate backbones are represented by outer strands while the complimentary base-pairing is represented by the elastic web. The phase angle between the two helices is $\alpha = 2.1$ radians. Here $\mathbf{R}^+ = [n_0^+ \ b_0^+ \ t_0^+]$ and $\mathbf{R}^- = [n_0^- \ b_0^- \ t_0^-]$ are the Frenet-Serret frames attached to the + and - strands, respectively. Base-pairs in reference and current configuration are shown to the right. $\mathbf{Q}_0^+ = \mathbf{Q}_0^- = \mathbf{Q}_0$ in the reference configuration. In the current configuration, the rigid rotation of the base-pair is quantified by $\mathbf{Q} = \mathbf{Z}(\mathbf{1} + \Phi)\mathbf{Q}_0$ (eqn. 4.23) and the elastic moment \mathbf{c} is related to the Gibbs rotation vector of $\mathbf{P} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{\frac{1}{2}}$ (eqn. 4.19).

Let \mathbf{Z} be a second order orthogonal tensor which relates the directors of the deformed centerline \mathbf{d}_i to those of the undeformed one \mathbf{e}_i , $i = 1, 2, 3$. As stated in section 2, the curvatures (k_1, k_2, k_3) associated with the deformation of the centerline are assumed to be small, nonetheless these could aggregate to potentially produce large rotations. The orthogonal tensor \mathbf{Z} operates as follows.

$$\mathbf{d}_i = \mathbf{Z} \mathbf{e}_i, \quad \mathbf{Z} = \sum_{i=1}^3 \mathbf{d}_i \otimes \mathbf{e}_i, \quad i = 1, 2, 3 \quad (4.3)$$

and

$$\begin{aligned} \mathbf{d}_{ix} &= \boldsymbol{\kappa} \times \mathbf{d}_i, & \text{where } \boldsymbol{\kappa} &= k_1 \mathbf{d}_1 + k_2 \mathbf{d}_2 + k_3 \mathbf{d}_3. \\ \mathbf{d}_{1x} &= k_3 \mathbf{d}_2 - k_2 \mathbf{d}_3, & \mathbf{d}_{2x} &= k_1 \mathbf{d}_3 - k_3 \mathbf{d}_2, & \mathbf{d}_{3x} &= k_2 \mathbf{d}_1 - k_1 \mathbf{d}_2. \end{aligned} \quad (4.4)$$

In the above equations, we assume that

$$r^+(x), k_1(x), k_2(x), k_3(x), \zeta(x), \beta^+(x) \sim O(\varepsilon).$$

Thus, in the treatment henceforth, any product terms such as r^2 or ξk_3 are $O(\varepsilon^2)$ and are neglected.

4.2.2 Step 2: Rotation of strands

A Frenet-Serret director frame consisting of normal, binormal and tangent is attached to each cross-section of the strand as shown in fig. 4.1. We denote it by $\mathbf{R}_0^+(x)$ in the reference configuration.

$$\begin{aligned}\mathbf{R}_0^+ &= [\mathbf{n}_0^+ \quad \mathbf{b}_0^+ \quad \mathbf{t}^+], \\ \mathbf{n}_0^+ &= -\cos \omega x \mathbf{e}_1 - \sin \omega x \mathbf{e}_2, \\ \mathbf{b}_0^+ &= -\cos k(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + \sin k \mathbf{e}_3, \\ \mathbf{t}_0^+ &= \sin k(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + \cos k \mathbf{e}_3,\end{aligned}\tag{4.5}$$

For the sake of brevity, we use

$$(\cos \omega x \mathbf{d}_1 + \sin \omega x \mathbf{d}_2) = \mathbf{f}_1^+, \quad (-\sin \omega x \mathbf{d}_1 + \cos \omega x \mathbf{d}_2) = \mathbf{f}_2^+, \quad \mathbf{d}_3 = \mathbf{f}_3^+.$$

As the strand deforms, the frame \mathbf{R}_0^+ evolves into $\mathbf{R}^+(x)$ which consists of normal, binormal and tangent to the deformed configuration of the strand. Our next step is to calculate the tangent vector to the deformed configuration. We differentiate eqn. (4.2) to obtain,

$$\begin{aligned}\mathbf{r}_x^+ &= (r_x - b\omega\beta^+)(\cos \omega x \mathbf{d}_1 + \sin \omega x \mathbf{d}_2) + (b\omega + \omega r + b\beta_x^+ + bk_3)(-\sin \omega x \mathbf{d}_1 + \cos \omega x \mathbf{d}_2) + \\ &\quad (1 + b\xi - bk_2 \cos \omega x + bk_1 \sin \omega x) \mathbf{d}_3.\end{aligned}\tag{4.6}$$

The strand is assumed to be inextensible and unsharable. This means,

$$|\mathbf{r}_x^+|^2 = 1 + \omega^2 b^2 + 2b(\omega^2 r + b\omega\beta_x^+ + b\omega k_3 + \xi - k_2 \cos \omega x + k_1 \sin \omega x) + O(\varepsilon^2) = |\mathbf{r}_{0x}^+|^2 = 1 + b^2 \omega^2,$$

which leads us to the inextensibility condition:

$$\xi - k_2 \cos \omega x + k_1 \sin \omega x = -\omega^2 r - b\omega(k_3 + \beta_x^+).\tag{4.7}$$

This equation will be used in the following sections to impose boundary conditions. Substituting eqn. (4.7) into eqn. (4.6) to get,

$$\mathbf{r}_x^+ = (r_x - b\omega\beta^+) \mathbf{f}_1^+ + (b\omega + \omega r + b\beta_x^+ + bk_3) \mathbf{f}_2^+ + (1 - b\omega^2 r - b^2\omega(k_3 + \beta_x^+)) \mathbf{f}_3^+.\tag{4.8}$$

The director frame in the deformed configuration can be computed as follows. The tangent vector \mathbf{t}^+ is,

$$\begin{aligned}\mathbf{t}^+ &= \frac{\mathbf{r}_x^+}{|\mathbf{r}_{0x}^+|} \\ &= (r_x \cos k - \beta^+ \sin k) \mathbf{f}_1^+ + (\sin k + \omega r \cos k + b(\beta_x^+ + k_3) \cos k) \mathbf{f}_2^+ \\ &\quad + (\cos k - \omega r \sin k - b(\beta_x^+ + k_3) \sin k) \mathbf{f}_3^+ \\ &= \mathbf{Z}(\mathbf{t}_0^+ - (r_x \cos k - \beta^+ \sin k) \mathbf{n}_0^+ - (\omega r + b(\beta_x^+ + k_3)) \mathbf{b}_0^+).\end{aligned}\quad (4.9)$$

The tangent vector is differentiated to calculate the normal vector in the deformed configuration

$$\begin{aligned}\mathbf{t}_x^+ &= (-\omega \sin k + (r_{xx} + \xi) \cos k - (\beta_x + k_3) \sin k) \mathbf{f}_1^+ \\ &\quad + (2\omega \cos k r_x + -\omega \beta^+ \sin k + b \cos k (\beta_{xx}^+ + k_{3x}) - f \cos k) \mathbf{f}_2^+ \\ &\quad + (f - \omega r_x - b(\beta_{xx}^+ + k_{3x})) \sin k \mathbf{f}_3^+ + O(\varepsilon^2).\end{aligned}\quad (4.10)$$

The above expression is used to calculate the curvature Ω^+ for the strand. This curvature is equal to the sum of the original curvature ($\omega \sin k$) and the one induced by the deformation (κ^+). Hence,

$$\begin{aligned}\Omega^+ &= (\mathbf{t}_x^+ \cdot \mathbf{t}_x^+)^{1/2} = \omega \sin k - (r_{xx} + \xi) \cos k + (\beta_x^+ + k_3) \sin k, \\ \kappa^+ &= \Omega^+ - \omega \sin k = -(r_{xx} + \xi) \cos k + (\beta_x^+ + k_3) \sin k.\end{aligned}\quad (4.11)$$

The bending moment \mathbf{m}^+ in the strand is proportional to κ^+ .

$$\mathbf{m}^+ = EI \kappa^+ \mathbf{b}^+ = EI \kappa^+ (-\cos k \mathbf{f}_2^+ + \sin k \mathbf{f}_3^+).\quad (4.12)$$

Also, the normal is

$$\begin{aligned}\mathbf{n}^+ &= \frac{1}{\Omega^+} \mathbf{t}^+ = -\mathbf{f}_1^+ + \frac{1}{\sin k} (r_x \sin k - \beta^+ \sin k) \mathbf{f}_2^+ + \frac{f - \omega r_x - b(\beta_{xx}^+ + k_{3x})}{\omega \sin k} (-\cos k \mathbf{f}_2 + \sin k \mathbf{f}_3^+), \\ &= \mathbf{Z}(\mathbf{n}_0^+ + (r_x \cos k - \beta^+ \sin k) \mathbf{t}_0^+ + (-\frac{(r_x \cos k - \beta^+ \sin k) \cos k}{\sin k} + \frac{g}{\omega \sin k}) \mathbf{b}_0^+).\end{aligned}\quad (4.13)$$

where

$$g(x) = f(x) - \omega r_x - b(\beta_{xx}^+ + k_{3x}), \quad f(x) = k_1 \cos \omega x + k_2 \sin \omega x.$$

Using the above deformed orthogonal frame attached to each cross section

$$\mathbf{R}^+ = [\mathbf{n}^+ \quad \mathbf{b}^+ \quad \mathbf{t}^+] = \mathbf{Z} \mathbf{R}_0^+ (\mathbf{1} + \mathbf{\Theta}^+),\quad (4.14)$$

where Θ^+ is a skew symmetric tensor and $\mathbf{Z} = \sum_{i=1}^3 \mathbf{d}_i \times \mathbf{e}_i$ as defined in eqn. 4.3,

$$\Theta^+ = \begin{bmatrix} 0 & -\theta_3^+ & \theta_2^+ \\ \theta_3^+ & 0 & -\theta_1^+ \\ -\theta_2^+ & \theta_1^+ & 0 \end{bmatrix}, \quad (4.15)$$

$$\text{in which } \theta_1^+ = r\omega + b(\beta_x^+ + k_3), \quad \theta_2^+ = -r_x \cos k + \beta^+ \sin k, \\ \theta_3^+ = \frac{g}{\omega \sin k} - \frac{(r_x \cos k - \beta^+ \sin k) \cos k}{\sin k}.$$

The above quantities \mathbf{r}^- , \mathbf{R}^- and κ^- etc. can be derived for the $-$ strand too; the relevant expressions are given in appendix B.

4.2.3 Step 3: Mechanics of base-pairing

The sugar-phosphate backbones of the DNA molecule are tied together by means of complimentary base-pairing. The base-pairing is modeled by elastic rods capable of extension, shear, bending and twisting. An orthogonal frame $\mathbf{Q}_0 = [\mathbf{f}_{01} \quad \mathbf{f}_{02} \quad \mathbf{f}_{03}]$ is attached to the strands such that \mathbf{f}_{01} is a unit vector pointing from the $-$ strand to the $+$ strand in the reference configuration as shown in fig. 4.1. Thus,

$$\begin{aligned} \mathbf{Q}_0 &= [\mathbf{f}_{01} \quad \mathbf{f}_{02} \quad \mathbf{f}_{03}], \\ \mathbf{f}_{01} &= \sin(\omega x + \frac{\alpha}{2}) \mathbf{e}_1 - \cos(\omega x + \frac{\alpha}{2}) \mathbf{e}_2, \\ \mathbf{f}_{02} &= \cos(\omega x + \frac{\alpha}{2}) \mathbf{e}_1 + \sin(\omega x + \frac{\alpha}{2}) \mathbf{e}_2, \\ \mathbf{f}_{03} &= \mathbf{e}_3. \end{aligned} \quad (4.16)$$

The two ends of the rod in the web are denoted as \pm such that the $+$ end lies on the $+$ strand and the $-$ end lies on the $-$ strand. The deformation of the web is completely determined by the displacement $(\mathbf{r}^+(x), \mathbf{r}^-(x))$ and rotation $(\mathbf{R}^+(x), \mathbf{R}^-(x))$ of its ends. As the outer strands undergo the deformation prescribed by eqn. (4.2), the web themselves undergo various kinds of deformation. We describe the rotation of the web via a rigid rotation and a micro-rotation (Moakher & Maddocks (2005)). The micro-rotation encapsulates the information about the difference in rotation of the two ends of the web. We calculate the mechanical quantities associated with the extension and bending of the web in two separate sections below.

4.2.3.1 Bending and twisting of the web

Our objective in this section is to calculate the micro-rotation tensor \mathbf{P} . A copy of \mathbf{Q}_0 say \mathbf{Q}_0^\pm is attached on the + and - end of every spoke in the reference configuration. \mathbf{Q}_0^\pm change to \mathbf{Q}^\pm in the current configuration. The 'difference' between \mathbf{Q}_0^+ and \mathbf{Q}_0^- gives the bending and torsion of the web while the 'average' of \mathbf{Q}_0^+ and \mathbf{Q}_0^- gives the rigid rotation of the web. \mathbf{Q}^\pm can be computed from $\mathbf{R}(x)^\pm$ —already known. The angles between the columns of \mathbf{Q}_0^+ and \mathbf{R}_0^+ should remain same during the deformation which translates into the following condition.

$$\begin{aligned}\mathbf{R}_0^{+T} \mathbf{Q}_0 &= \mathbf{R}^{+T} \mathbf{Q}^+, \\ \mathbf{Q}^+ &= \mathbf{R}^+ \mathbf{R}_0^{+T} \mathbf{Q}_0 = \mathbf{Z} \mathbf{R}_0^+ (\mathbf{1} + \boldsymbol{\Theta}^+) \mathbf{R}_0^{+T} \mathbf{Q}_0, \\ \mathbf{Q}^- &= \mathbf{Z} \mathbf{R}_0^- (\mathbf{1} + \boldsymbol{\Theta}^-) \mathbf{R}_0^{-T} \mathbf{Q}_0.\end{aligned}\tag{4.17}$$

Let the micro-rotation tensor in the reference configuration be \mathbf{P}_0 which changes to \mathbf{P} during deformation. We use an expression for \mathbf{P}/\mathbf{P}_0 given in Moakher and Maddocks (Moakher & Maddocks (2005)).

$$\begin{aligned}\mathbf{P}_0^2 &= \mathbf{Q}_0^+ \mathbf{Q}_0^{-T} = \mathbf{I}, \\ \mathbf{P}^2 &= \mathbf{Q}^+ \mathbf{Q}^{-T} = \mathbf{Z} \mathbf{R}_0^+ (\mathbf{1} + \boldsymbol{\Theta}^+) \mathbf{R}_0^{+T} \mathbf{Q}_0 \mathbf{Q}_0^T \mathbf{R}_0^- (\mathbf{1} - \boldsymbol{\Theta}^-) \mathbf{R}_0^{-T} \mathbf{Z}^T, \\ &= \mathbf{Z} (\mathbf{1} + \mathbf{R}^+ \boldsymbol{\Theta}^+ \mathbf{R}^{+T} - \mathbf{R}^- \boldsymbol{\Theta}^- \mathbf{R}^{-T}) \mathbf{Z}^T.\end{aligned}\tag{4.18}$$

This gives

$$\mathbf{P}_0 = \mathbf{I}, \quad \mathbf{P} \approx \mathbf{Z} \left(\mathbf{1} + \frac{\mathbf{R}^+ \boldsymbol{\Theta}^+ \mathbf{R}^{+T} - \mathbf{R}^- \boldsymbol{\Theta}^- \mathbf{R}^{-T}}{2} \right) \mathbf{Z}^T = \mathbf{Z} (\mathbf{1} + \boldsymbol{\Phi}^c) \mathbf{Z}^T.\tag{4.19}$$

Note that $\boldsymbol{\Phi}^c$ is a skew symmetric tensor. The next step is to calculate the Gibbs rotation vector of \mathbf{P} (Moakher & Maddocks (2005)). The Gibbs rotation vector \bar{t} of a rotation matrix \mathbf{T} is defined as $\bar{t} = \tan \frac{\theta}{2} \mathbf{k}$ such that $\text{tr} \mathbf{P} = 1 + 2 \cos \theta$ and \mathbf{k} is a unit vector such that $\mathbf{T} \mathbf{k} = \mathbf{k}$. Consider $\bar{\mathbf{P}} = \mathbf{1} + \boldsymbol{\Phi}^c$ where $\boldsymbol{\Phi}^c \sim O(\varepsilon)$. The axis of the infinitesimal rotation $\bar{\mathbf{P}}$ is the axial vector of $\boldsymbol{\Phi}^c$. Hence,

$$\bar{\mathbf{P}} \boldsymbol{\phi}^c = (\mathbf{1} + \boldsymbol{\Phi}^c) \boldsymbol{\phi}^c = \boldsymbol{\phi}^c, \quad \text{which gives } \mathbf{k} = \frac{\boldsymbol{\phi}^c}{|\boldsymbol{\phi}^c|}.\tag{4.20}$$

The magnitude of the rotation can not be calculated by taking $\text{tr} \bar{\mathbf{P}}$, since it gives $1 + 2 \cos \theta = 3$ which implies $\theta = 0$. We consider the following limit.

$$\mathbf{1} + \boldsymbol{\Phi}^c = \lim_{\phi_1^c \rightarrow 0} \lim_{\phi_2^c \rightarrow 0} \lim_{\phi_3^c \rightarrow 0} \mathbf{R}_1(\phi_1^c) \mathbf{R}_2(\phi_2^c) \mathbf{R}_3(\phi_3^c).\tag{4.21}$$

Now we take the trace of the RHS and get $\theta = |\phi^c|$. Hence, the Gibbs rotation vector of $\bar{\mathbf{P}}$, $\bar{\boldsymbol{\eta}}$ is given as

$$2\bar{\boldsymbol{\eta}} = 2 \tan \frac{\theta}{2} \mathbf{k} \approx |\phi^c| \frac{\phi^c}{|\phi^c|} = \phi^c. \quad (4.22)$$

The Gibbs rotation vector of \mathbf{P} is simply $\boldsymbol{\eta} = \mathbf{Z}\bar{\boldsymbol{\eta}}$. Note that in the undeformed state $\boldsymbol{\eta}_0 = \bar{\boldsymbol{\eta}}_0 = 0$. We now proceed to calculate the rigid rotation of the spoke \mathbf{Q} .

$$\mathbf{Q} = \mathbf{P}\mathbf{Q}^{-1} = \mathbf{Z}\left(\mathbf{1} + \frac{\mathbf{R}^+\boldsymbol{\Theta}^+\mathbf{R}^{+T} + \mathbf{R}^-\boldsymbol{\Theta}^-\mathbf{R}^{-T}}{2}\right)\mathbf{Q}_0 = \mathbf{Z}(\mathbf{1} + \boldsymbol{\Phi})\mathbf{Q}_0 = \mathbf{Z}(\mathbf{1} + \boldsymbol{\Phi})\mathbf{Q}_0. \quad (4.23)$$

Here $\boldsymbol{\eta} \sim O(\varepsilon)$. Now, the micro-moment \mathbf{c} is related linearly to the $\boldsymbol{\eta}$ via an elastic tensor \mathbf{H} .

$$\mathbf{c} = \mathbf{Q}\bar{\mathbf{H}}[\mathbf{Q}^T\boldsymbol{\eta} - \mathbf{Q}_0^T\boldsymbol{\eta}_0] + O(\varepsilon^2) \approx \mathbf{Z}\mathbf{Q}_0\bar{\mathbf{H}}\mathbf{Q}_0^T\bar{\boldsymbol{\eta}}. \quad (4.24)$$

For further reference, let

$$\hat{\boldsymbol{\zeta}} = \mathbf{Q}_0^T\bar{\boldsymbol{\eta}}. \quad (4.25)$$

4.2.3.2 Extension of the web

The distance between the two strands is $\mathbf{w} = \frac{\mathbf{r}^+ - \mathbf{r}^-}{2}$ and in the undeformed configuration $\mathbf{w}_0 = \frac{\mathbf{r}_0^+ - \mathbf{r}_0^-}{2}$. By direct calculation we observe

$$\begin{aligned} \mathbf{w}_0 &= b \sin \frac{\alpha}{2} \left(\sin(\omega x + \frac{\alpha}{2}) \mathbf{e}_1 - \cos(\omega x + \frac{\alpha}{2}) \mathbf{e}_2 \right), \\ \mathbf{w} &= (b \sin \frac{\alpha}{2} + w_1) \left(\sin(\omega x + \frac{\alpha}{2}) \mathbf{d}_1 - \cos(\omega x + \frac{\alpha}{2}) \mathbf{d}_2 \right) + \\ & \quad w_2 \left(\cos(\omega x + \frac{\alpha}{2}) \mathbf{d}_1 + \sin(\omega x + \frac{\alpha}{2}) \mathbf{d}_2 \right), \end{aligned} \quad (4.26)$$

where

$$w_1 = \frac{r + r^-}{2} \sin \frac{\alpha}{2} - b \frac{\beta^+ - \beta^-}{2} \cos \frac{\alpha}{2}, \quad \text{and} \quad w_2 = \frac{r - r^-}{2} \cos \frac{\alpha}{2} + b \frac{\beta^+ + \beta^-}{2} \sin \frac{\alpha}{2}.$$

The force exerted by the + strand on the - strand \mathbf{f} is given by,

$$\mathbf{f} = \mathbf{Q}\bar{\mathbf{L}}[\mathbf{Q}^T\mathbf{w} - \mathbf{Q}_0^T\mathbf{w}_0], \quad (4.27)$$

where $\bar{\mathbf{L}}$ is a tensor of mechanical properties of the web. This force \mathbf{f} causes the web to extend and shear. For further reference let,

$$\hat{\mathbf{w}} = \mathbf{Q}^T\mathbf{w} - \mathbf{Q}_0^T\mathbf{w}_0 \quad (4.28)$$

4.2.3.3 Stacking energy

DNA consists of consecutive base-pairs stacked on top of each other in a regular fashion. The resistance to external forces and moments not only comes from the elastic deformation of the strands and the webbing but also from the change in alignment of the base-pairs. We call the energy associated with this change in bases' position and spatial orientation 'stacking energy'. Stacking energy plays a critical role in various phenomena such as melting of DNA (Dauxois et al. (1993); Peyrard & Bishop (1989)). We prescribe a form of free energy which is quadratic in the twist k_3 and stretch ξ .

$$F_{int} = K_c k_3^2 + K_e \xi^2. \quad (4.29)$$

There are other sophisticated expressions for the stacking energy (Dauxois et al. (1993)), but we use the quadratic form for two reasons: one, the non-quadratic terms in the energy of (Dauxois et al. (1993)) account for effects such as base-pair severing which are crucial to DNA melting which does not occur in our problem, two, a quadratic energy keeps our problem linear. This interaction energy results in a distributed body force \mathbf{l} and distributed body moment \mathbf{h} on the strands.

$$\mathbf{h} = K_c k_{3x} \mathbf{d}_3, \quad \mathbf{l} = K_e \xi_x \mathbf{d}_3. \quad (4.30)$$

4.2.4 Step 4: Governing equations

We are now in a position to solve the governing equations for the mechanics of our helical birod. These equations consist of balance of linear momentum and angular momentum for both the strands. In the balance equations eqn. (4.31) and eqn. (4.32):

- $\mathbf{m}^\pm = EI\kappa^\pm$ (eqn. 4.12) denotes the elastic moment in the \pm strand. \mathbf{n}^\pm are the contact forces for which there is no constitutive relation since the outer strands are assumed to be inextensible and unshearable.
- \mathbf{f} and \mathbf{c} are the distributed force and moment, respectively, exerted by the + strand on the - strand.
- \mathbf{l} and \mathbf{h} are the distributed force and moment exerted by base-pairs on the + and - strand.

The balance equations are:

$$\mathbf{n}_x^+ - \mathbf{f} + \mathbf{l} = 0, \quad (4.31a)$$

$$\mathbf{n}_x^- + \mathbf{f} + \mathbf{l} = 0, \quad (4.31b)$$

$$\mathbf{m}_x^+ + \mathbf{r}_x^+ \times \mathbf{n}^+ + \frac{1}{2}(\mathbf{r}^+ - \mathbf{r}^-) \times \mathbf{f} - \mathbf{c} + \mathbf{h} = 0, \quad (4.32a)$$

$$\mathbf{m}_x^- + \mathbf{r}_x^- \times \mathbf{n}^- + \frac{1}{2}(\mathbf{r}^+ - \mathbf{r}^-) \times \mathbf{f} + \mathbf{c} + \mathbf{h} = 0, \quad (4.32b)$$

Let $[\mathbf{f}_1 \quad \mathbf{f}_2 \quad \mathbf{f}_3] = \mathbf{ZQ}_0$. This gives

$$\begin{aligned} \mathbf{f}_1 &= \left(\sin(\omega x + \frac{\alpha}{2}) \mathbf{d}_1 - \cos(\omega x + \frac{\alpha}{2}) \mathbf{d}_2 \right), \\ \mathbf{f}_2 &= \left(\cos(\omega x + \frac{\alpha}{2}) \mathbf{d}_1 + \sin(\omega x + \frac{\alpha}{2}) \mathbf{d}_2 \right), \\ \mathbf{f}_3 &= \mathbf{d}_3. \end{aligned} \quad (4.33)$$

We decompose the forces, $\mathbf{n}^+ = (\mathbf{n} + \mathbf{n}^c) \sim O(\varepsilon)$ and $\mathbf{n}^- = (\mathbf{n} - \mathbf{n}^c) \sim O(\varepsilon)$. $\mathbf{n} = n_1 \mathbf{f}_1 + n_2 \mathbf{f}_2 + n_3 \mathbf{f}_3$ and $\mathbf{n}^c = n_1^c \mathbf{f}_1 + n_2^c \mathbf{f}_2 + n_3^c \mathbf{f}_3$. Now, $\mathbf{n}_x = (n_{1x} - \omega n_2) \mathbf{f}_1 + (n_{2x} + \omega n_1) \mathbf{f}_2 + n_{3x} \mathbf{f}_3 + O(\varepsilon^2)$. Similarly for \mathbf{n}_x^c . We use $\mathbf{c} = c_1 \mathbf{f}_1 + c_2 \mathbf{f}_2 + c_3 \mathbf{f}_3$ and $\mathbf{f} = f_1 \mathbf{f}_1 + f_2 \mathbf{f}_2 + f_3 \mathbf{f}_3$ from eqn. (4.27) and eqn. (4.24). Then, the balance equations become:

$$\begin{aligned} n_{1x} - \omega n_2 &= 0, \\ n_{2x} + \omega n_1 &= 0, \\ n_{3x} + K_e \xi_x &= 0, \\ n_{1x}^c - \omega n_2^c - f_1 &= 0, \\ n_{2x}^c + \omega n_1^c - f_2 &= 0, \\ n_{3x}^c - f_3 &= 0, \\ EI \cos k[(\kappa_x^+ + \kappa_x^-) \cos \frac{\alpha}{2} + (\kappa^+ - \kappa^-) \omega \sin \frac{\alpha}{2}] - 2n_2 + 2a\omega n_3^c \sin \frac{\alpha}{2} &= 0, \\ EI \cos k[(\kappa_x^- - \kappa_x^+) \sin \frac{\alpha}{2} + (\kappa^- + \kappa^+) \omega \cos \frac{\alpha}{2}] + 2n_1 + 2a\omega n_3 \cos \frac{\alpha}{2} - 2af_3 \sin \frac{\alpha}{2} &= 0, \\ EI \sin k(\kappa_x^+ + \kappa_x^-) + 2af_2 \sin \frac{\alpha}{2} - 2sn_2 \cos \frac{\alpha}{2} - 2a\omega n_1^c \sin \frac{\alpha}{2} + 2K_c k_{3x} &= 0, \\ EI \cos k[(\kappa_x^+ - \kappa_x^-) \cos \frac{\alpha}{2} + (\kappa^+ + \kappa^-) \omega \sin \frac{\alpha}{2}] + 2a\omega n_3 \sin \frac{\alpha}{2} - 2n_2^c - 2c_1 &= 0, \\ EI \cos k[-(\kappa_x^- + \kappa_x^+) \sin \frac{\alpha}{2} + (\kappa^+ - \kappa^-) \omega \cos \frac{\alpha}{2}] + 2\omega n_3^c \cos \frac{\alpha}{2} + 2n_1^c - 2c_2 &= 0, \\ EI \sin k(\kappa_x^+ - \kappa_x^-) - 2a\omega n_2^c \cos \frac{\alpha}{2} - 2a\omega n_1 \sin \frac{\alpha}{2} - 2c_3 &= 0, \end{aligned} \quad (4.34)$$

We have 12 differential equations in the 12 unknowns $(r, f, \xi, k_3, \beta^+, \beta^-, n_1^c, n_2^c, n_3^c, n_1, n_2, n_3)$.

We substitute the following ansatz into the equations.

$$y = y_0 e^{-\lambda x} \quad \text{where } y \text{ could be } r(x), f(x), \xi(x), k_3(x), \beta^+(x), \beta^-(x), n_1^c, n_2^c, n_3^c, n_1, n_2, n_3. \quad (4.35)$$

This results in an eigenvalue problem. We find 23 eigenvalues, but retain only 6 for reasons explained in the appendix. Let those 6 eigenvalues be $\pm\lambda, \pm\mu, \pm\delta$ and the corresponding eigenvectors $\mathbf{v}_{\pm\lambda}$ and $\mathbf{v}_{\pm\mu}$. Let

$$\mathbf{v}(x) = [r(x) \quad f(x) \quad \xi(x) \quad k_3(x) \quad \beta^+(x) \quad \beta^-(x) \\ n_1^c(x) \quad n_2^c(x) \quad n_3^c(x) \quad n_3(x) \quad n_1(x) \quad n_2(x)]^T.$$

Hence,

$$\mathbf{v}(x) = p_1 e^{-\lambda x} \mathbf{v}_\lambda + p_2 e^{\lambda x} \mathbf{v}_{-\lambda} + p_3 e^{-\mu x} \mathbf{v}_\mu + p_4 e^{\mu x} \mathbf{v}_{-\mu} + p_5 e^{-\delta x} \mathbf{v}_\delta + p_6 e^{\delta x} \mathbf{v}_{-\delta}. \quad (4.36)$$

Here, p_1, p_2, p_3, p_4, p_5 and p_6 are the constants which are determined using boundary conditions.

4.2.5 Step 5: Boundary conditions

We assume that the impact of a protein binding to DNA is two fold: a) the protein fixes the curvatures at the binding site as in (Liang & Purohit (2018a)), and b) the protein causes a change in the radius of the DNA helix (Kim et al. (2013)) as shown in the inset of fig. 4.5 (b). Thus, we apply boundary conditions on the curvatures k_1, k_2 and the change in radius r of the DNA helix. We discuss two cases, first, when one protein binds to the DNA, and second, when two proteins bind to it.

1. *One protein:* Let us assume that the protein binds at $x = 0$. The boundary conditions for this case are:

$$\begin{aligned} \text{At } x = 0, \quad k_1(x) = k_{10}, \quad k_2(x) = k_{20}, \quad r(x) = r_0. \\ \text{As } x \rightarrow \pm\infty, \quad k_1(x), k_2(x), r(x) \rightarrow 0. \end{aligned} \quad (4.37)$$

The second boundary condition says that the DNA is straight far away from the protein and that the perturbation in DNA radius occurs only in the vicinity of the bound protein.

2. *Two proteins:* Let us assume that the two proteins bind at $x = 0$ and $x = a$, respectively. We divide our domain into three parts $-\infty < x < 0$, $0 < x < a$ and

$a < x < \infty$ each of which has different boundary conditions attached to it.

Region 1 : $x \in (-\infty, 0)$

as $x \rightarrow -\infty$, $k_1(x), k_2(x), r(x) \rightarrow 0$, at $x = 0$, $k_1(x) = k_{11}$, $k_2(x) = k_{12}$, $r(x) = r_1$.

Region 2 : $x \in (0, a)$

at $x = 0$, $k_1(x) = k_{11}$, $k_2(x) = k_{12}$, $r(x) = r_1$,

at $x = a$, $k_1(x) = k_{21}$, $k_2(x) = k_{22}$, $r(x) = r_2$.

Region 3 : $x \in (a, \infty)$

at $x = a$, $k_1(x) = k_{21}$, $k_2(x) = k_{22}$, $r(x) = r_2$, as $x \rightarrow \infty$, $k_1(x), k_2(x), r(x) \rightarrow 0$.

(4.38)

4.2.6 Step 6: Energy of the birod

We assume small elastic deformations throughout, hence the resulting energy is quadratic in the strain variables. The elastic energy has contributions from the bending of the outer strands eqn. (4.12), the extension, bending and twisting of the web eqn. (4.25), (4.26) and the stacking energy eqn. (4.30).

$$E = \int_{-\infty}^{\infty} \left[\frac{1}{2} EI \kappa^{+2} + \frac{1}{2} EI \kappa^{-2} + \frac{1}{2} \hat{\mathbf{w}} \cdot \mathbf{L} \hat{\mathbf{w}} + \frac{1}{2} \hat{\boldsymbol{\zeta}} \cdot \mathbf{H} \hat{\boldsymbol{\zeta}} + K_e \xi^2 + K_c k_3^2 \right] dx. \quad (4.39)$$

We are especially interested in the interaction energy ΔG which is the elastic energy of interactions between the two proteins.

$$\Delta G = E_a^2 - E_a^1 - E_0^1, \quad (4.40)$$

where E_a^2 is the energy of two proteins bound to DNA, one at $x = 0$ and other at $x = a$, and E_a^1 and E_0^1 are the elastic energies corresponding to a single protein binding at $x = a$ and $x = 0$, respectively.

4.3 Elastic constants

Our model has 9 elastic constants $L_1, L_2, L_3, H_1, H_2, H_3, K_c, K_e, EI$. The experimental values for these constants are not known. In order to get some idea about the magnitude of the elastic constants we calculate the extensional modulus, torsional modulus and twist-stretch coupling modulus for a double-stranded DNA within our birod model.

The explicit calculation is presented in the appendix. We choose

$$\begin{aligned} K_c &= 80\text{pNnm}^2, & K_e &= 600\text{pN}, & H_1 &= 15\text{pN}, & H_2 &= 10\text{pN}, & H_3 &= 20\text{pN}, \\ L_1 &= 50\text{pN nm}^{-1}, & L_2 &= 250\text{pN nm}^{-1}, & L_3 &= 30\text{pN nm}^{-1}, & EI &= 65\text{pN nm}^2. \end{aligned} \quad (4.41)$$

This choice of elastic constants gives the extensional modulus $S \approx 1245$ pN, torsional modulus $C \approx 490$ pNnm² and twist-stretch coupling modulus $g \approx -90$ pNnm which are close to actual values for ds-DNA (Singh & Purohit (2017)) measured in experiments. We point out that this choice of elastic constants is not unique, nonetheless we use them to make further calculations.

When we substitute these constants into the governing equations (eqn. (4.34)) and solve the eigenvalue problem involving λ , we get the following eigenvalues

$$\lambda_1 = -0.68, \lambda_2 = -0.42, \lambda_3 = -0.36, \lambda_4 = 0.36, \lambda_5 = 0.42, \lambda_6 = 0.68. \quad \text{Units: nm}^{-1} \quad (4.42)$$

Other eigenvalues are either very large ($\rightarrow \pm\infty$), very small (~ 0) or purely imaginary. Purely imaginary and zero eigenvalues when substituted in $e^{\lambda x}$ give a sinusoidal and a constant function, respectively, which do not decay to zero as $x \rightarrow \pm\infty$. As mentioned in section 4.2.5, the curvatures k_1, k_2 and change in radius r must go to zero at $\pm\infty$. Thus, zero or purely imaginary eigenvalues cannot satisfy our boundary conditions, and are, therefore, not useful. We refer the reader to the appendix for further discussion on the choice of eigenvalues.

Consider a situation in which two proteins bind DNA, one at $x = 0$ and the other at $x = a$. In the region $a < x < \infty$ the solution eqn. (4.36) consists of only negative eigenvalues. There are three negative eigenvalues $\lambda_{2,3,4}$ and consequently three unknown constants. We have three boundary conditions on k_1, k_2 and r at $x = a$ to determine those constants. Similarly in the region $-\infty < x < 0$, the solution consists of only positive eigenvalues $\lambda_{7,8,9}$, so the constants can again be evaluated from three boundary conditions. We use this scheme to evaluate the strain parameters which we substitute into the expression for the elastic energy functional eqn. (4.39). Notice that the dominant eigenvalue ± 0.36 nm⁻¹ corresponds to a decay length of 2.8 nm (≈ 10 bp) which is what Kim *et al.* (Kim *et al.* (2013)) report in their experiments.

4.4 Results

The experimental evidence for allosteric interactions when two proteins bind to DNA is documented in Kim *et al.* (Kim *et al.* (2013)). Many earlier papers have also described allostery in DNA, but Kim *et al.* present exquisite quantitative details which call for a quantitative explanation.

To unravel the physics behind these allosteric interactions, we begin by examining the case when one protein binds to DNA. As discussed in section 4.2.4, the strain variables $(r, \zeta, \beta^\pm, k_{1,2,3})$ are linear combinations of decaying exponentials. For instance, consider $k_3(x)$ for a protein binding at $x = 0$:

$$\begin{aligned} k_3(x) &= p_1 \mathbf{v}_{-\lambda}(4) e^{\lambda x} + p_2 \mathbf{v}_{-\mu}(4) e^{\mu x} + p_3 \mathbf{v}_{-\delta}(4) e^{\delta x} & x < 0, \\ k_3(x) &= q_1 \mathbf{v}_\lambda(4) e^{-\lambda x} + q_2 \mathbf{v}_\mu(4) e^{-\mu x} + q_3 \mathbf{v}_\delta(4) e^{-\delta x} & x > 0, \end{aligned} \quad (4.43)$$

where $\lambda = 0.36 \text{ nm}^{-1}$, $\mu = 0.42 \text{ nm}^{-1}$, and $\delta = 0.68 \text{ nm}^{-1}$. $\mathbf{v}_{\pm\lambda}$, $\mathbf{v}_{\pm\mu}$, and $\mathbf{v}_{\pm\delta}$ are the eigenvectors associated with eigenvalues $\pm\lambda$, $\pm\mu$, and $\pm\delta$, respectively. The constants p_i and q_i ($i = 1, 2, 3$) are evaluated using the boundary conditions at $x = 0$. It is not difficult to see that the strain variables decay to zero as $x \rightarrow \pm\infty$. We can replace k_3 in the above equation by other strain variables (r, ξ, β^\pm) and recover similar behavior. We discuss a few characteristics of the variation of the strain parameters as functions of position. The results are plotted in fig. 4.2 and fig. 4.3. The strain parameters (r, k_3, β^\pm) decay exponentially with distance from the site of protein binding. The curvatures exhibit an exponentially decaying sinusoidal character with a period of 11 bp. This periodic decay of the curvatures manifests itself as sinusoidal variations in the interaction energy. We find that these plots are slightly asymmetric about $x = 0$. We attribute this to the structural asymmetry in the right-handed double-helix with phase angle $\alpha = 2.1$ radians. If we choose the phase angle $\alpha = \pi$ radians instead, we find that the plots are exactly symmetric about the site of protein binding as shown in the appendix. We now consider the case when two proteins bind to DNA, one at $x = 0$ and the other at $x = a$. We proceed in a similar manner as above and express the strain profiles as linear combinations of exponentials:

$$\begin{aligned} \text{Case 1} \quad k_3(x) &= p_1 \mathbf{v}_{-\lambda}(4) e^{\lambda x} + p_2 \mathbf{v}_{-\mu}(4) e^{\mu x} + p_3 \mathbf{v}_{-\delta}(4) e^{\delta x} & x < 0, \\ \text{Case 2} \quad k_3(x) &= m_1 \mathbf{v}_\lambda(4) e^{-\lambda x} + m_2 \mathbf{v}_\mu(4) e^{-\mu x} + m_3 \mathbf{v}_{-\lambda}(4) e^{\lambda x} + \\ & m_4 \mathbf{v}_{-\mu}(4) e^{\mu x} + m_5 \mathbf{v}_\delta(4) e^{-\delta x} + m_6 \mathbf{v}_{-\delta}(4) e^{\delta x} & 0 < x < a, \\ \text{Case 3} \quad k_3(x) &= q_1 \mathbf{v}_\lambda(4) e^{-\lambda x} + q_2 \mathbf{v}_\mu(4) e^{-\mu x} + q_3 \mathbf{v}_\delta(4) e^{-\delta x} & x > a. \end{aligned} \quad (4.44)$$

The constants p_i and q_i ($i = 1, 2, 3$) are determined by three boundary conditions (on k_1, k_2 and r) at $x = 0$ and $x = a$, respectively. The constants m_j , ($j = 1, 2, 3, 4, 5, 6$)

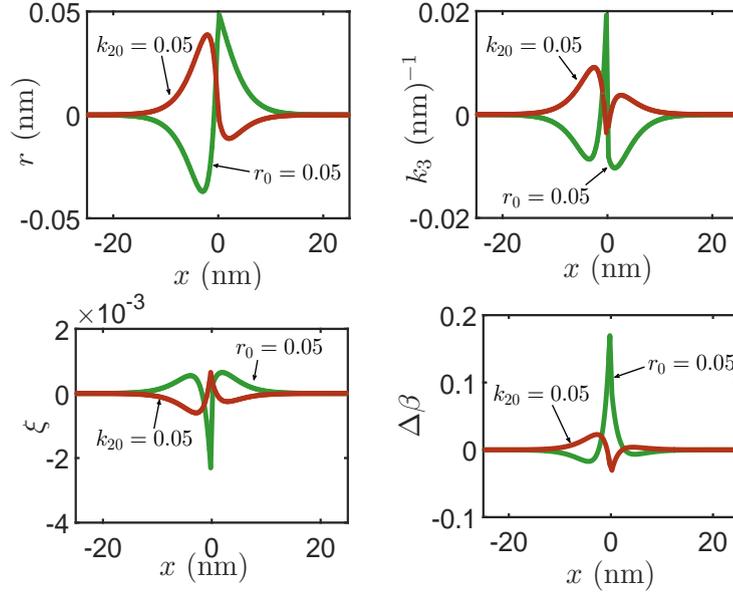


FIGURE 4.2: Variation of strain variables for a single protein. We plot the change in radius r , twist k_3 , stretch of centerline ξ and change in phase angle $\Delta\beta = \beta^+ - \beta^-$ for the double-helix. The red curve corresponds to the boundary conditions $k_{10} = r_0 = 0$ and $k_{20} = 0.1 \text{ nm}^{-1}$ at $x = 0$ and the green curve corresponds to $k_{10} = k_{20} = 0$ and $r_0 = 0.05 \text{ nm}$ at $x = 0$. The asymmetry of the double-helix (there is a major and minor groove in DNA) arising from the phase angle $\alpha = 2.1$ radian gives the curves a slight asymmetry about the site of protein binding. The curves are exactly symmetric about the site of protein binding if we choose phase angle $\alpha = \pi$ radians (which results in no major and minor groove) as shown in the appendix.

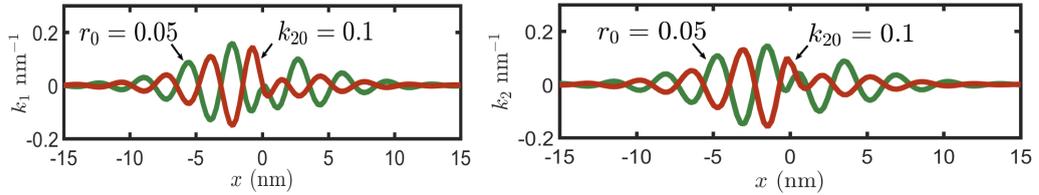


FIGURE 4.3: Variation of curvatures k_1 and k_2 for a single protein. The red curve corresponds to the boundary conditions $k_{10} = r_0 = 0$ and $k_{20} = 0.1 \text{ nm}^{-1}$ at $x = 0$ and the green curve corresponds to $k_{10} = k_{20} = 0$ and $r_0 = 0.05 \text{ nm}$ at $x = 0$. We find that the curvature decays exponentially and oscillates with a period $\approx 11 \text{ bp}$.

are determined by six boundary conditions at $x = 0$ and $x = a$. The behavior of the strain variables for two proteins is similar to that for one protein as shown in fig. 4.4. When two proteins are separated by a large distance $a > 10 \times 3.4 \text{ nm}$ (i.e., more than 10 helical turns of DNA), the strain profile looks like a concatenation of the profiles of two proteins binding separately. Their strain fields do not interact at such distances, thus there is little interaction energy. When the distance decreases, the strain fields of the two proteins overlap, and this is responsible for the interaction energy. As discussed in the appendix, two defects on a straight ladder interact via an interaction energy that decays exponentially with the distance between them. Now, we focus on the double-helical

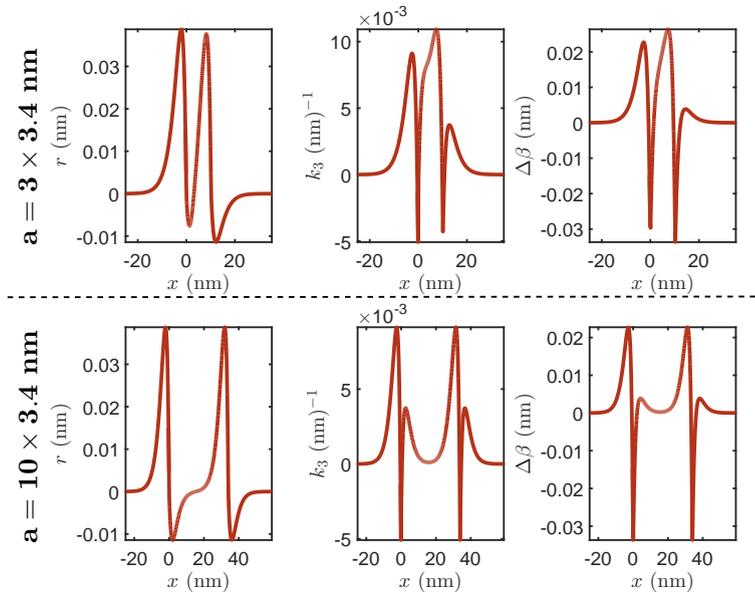


FIGURE 4.4: Variation of r , k_3 and $\Delta\beta$ for two proteins. Here a is the distance between the sites of protein binding. The strain variables decay exponentially away from the site of protein binding. When the distance between the proteins is large 10×3.4 nm, the profile looks like a concatenation of two solutions for a single protein.

birod and examine the behavior of different boundary conditions on the interaction energy ΔG in fig. 4.5. We assume for simplicity that both proteins apply the same boundary conditions on the DNA, the exact numerical values are given in the figure. If we choose the change in radius $r_0 = 0$ and apply the boundary conditions only on the two curvatures k_1, k_2 , the interaction energy decays exponentially while varying sinusoidally with a period of $5.5 \approx 11/2$ bp. This case corresponds to proteins that bend DNA as shown in the inset of fig. 4.5(b). On the other hand, if the curvatures k_1, k_2 are zero while the change in radius r_0 is non-zero, we get an exponentially decaying profile devoid of any oscillatory character, which is similar to the results for the ladder. The exponentially decaying component originates from the elasticity of the web, and the sinusoidal behavior comes from the double-helical structure of DNA. From this exercise we conclude that in order to get a sinusoidally varying interaction energy a protein must change the local curvature in the DNA, a mere change in radius of the DNA is not sufficient to give rise to the interaction energy profiles observed in experiments.

In our model the magnitude of the interaction energy increases monotonically with increase in the magnitude of the changes in curvatures or radius caused by the two proteins. Thus, by systematically varying the boundary conditions imposed by the proteins we can establish agreement of our theoretical results for ΔG with the experimental values documented by Kim *et al.* (Kim *et al.* (2013)). This is done in fig. 4.5(b). The values of the curvatures that give the best fit to the experimental data are $k_{11} = k_{21} = 0.02$

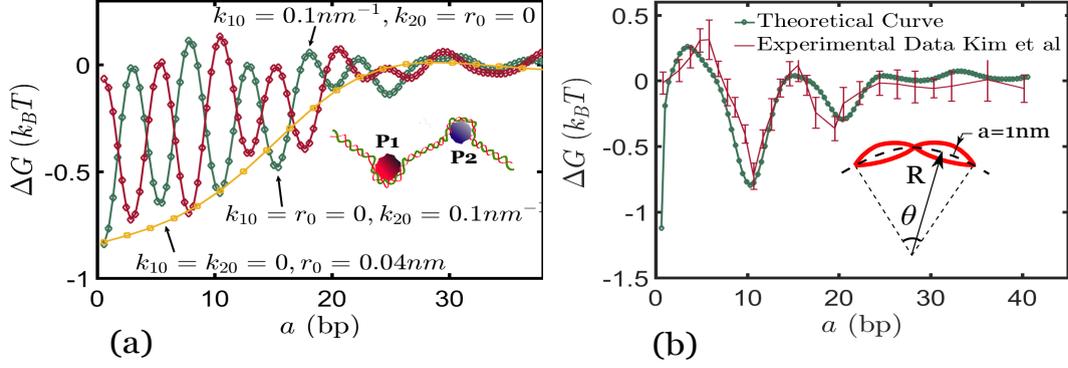


FIGURE 4.5: We plot the interaction energy between two proteins eqn. (4.40). In (a) we plot the behavior of ΔG for various boundary conditions. If the boundary conditions are specified on the curvatures we get an exponentially decaying profile oscillating with 5–6 bp ($\approx 11/2$ bp). The oscillatory behavior arises from the periodic geometry of DNA. In (b) the experimental data reproduced for comparison are from Kim *et al* (Kim *et al.* (2013)). We use $k_{11} = k_{21} \approx 0.02 \text{ nm}^{-1}$, $k_{12} = k_{22} = 0.05 \text{ nm}^{-1}$, $r_1 = -r_2 = 0.02 \text{ nm}$. The inset in (a) shows a protein DNA complex in which the proteins locally bend DNA. The inset in (b) shows that bending a DNA oligomer leads to widening of the groove on one side while narrowing it on the other. We find that the change in the groove width is approximately $\approx 2 \text{ \AA}$ which is close to the values reported by Kopka *et al* (Kopka *et al.* (1985)) (0.5 – 2 \AA)

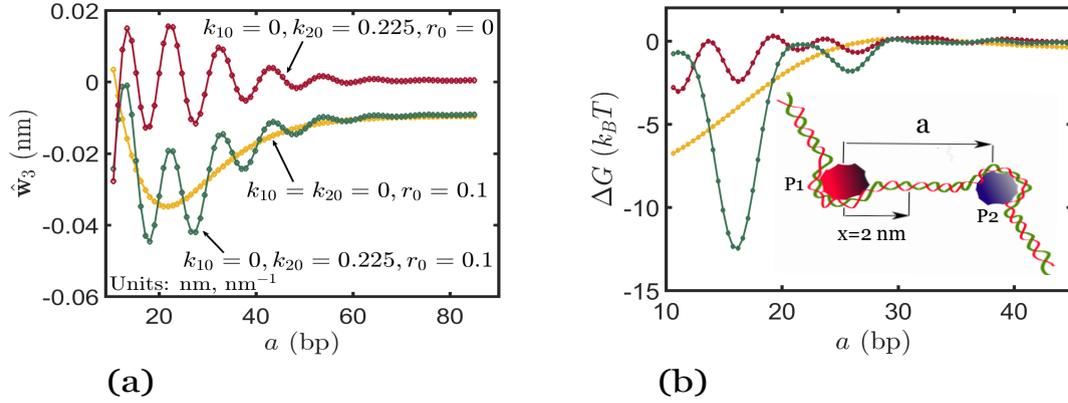


FIGURE 4.6: The inset in (b) shows a two protein complex. The boundary conditions are identical for both the proteins $k_{11} = k_{21} = k_{10}$, $k_{12} = k_{22} = k_{20}$, $r_1 = r_2 = r_0$; the legend in (a) contains the exact numerical values. For (b) the legend is the same as in (a). We examine behavior of $\hat{w}_3(x = 2 \text{ nm}, a)$ (eqn. 4.39) as a function of distance between the two proteins a for these boundary conditions. The strain variables oscillate with a period of 11 bp. We observe that in case of $r_0 = 0$, the strain parameter $\hat{w}_3(x = 2 \text{ nm}, a)$ decays as $e^{-\Gamma a} \psi(\omega a)$ where $\psi(\omega a)$ is a sinusoidal function, hence the combined energy of a two protein complex which is proportional to $(e^{-\Gamma a} \psi(\omega a))^2$ oscillates with a period of 5.5 bp (period of $\sin^2 x$ is half that of $\sin x$). If $k_{10} = k_{20} = 0$ the decay is exponential. If $r_0 \neq 0$ and k_{10} or $k_{20} \neq 0$, $\hat{w}_3(x = 2 \text{ nm}, a) \sim (e^{-\Gamma_1 a} \psi(\omega a) + e^{-\Gamma_2 a})$ and the energy of the two protein complex, which is proportional to $(e^{-\Gamma_1 a} \psi(\omega a) + e^{-\Gamma_2 a})^2$, oscillates with a period of 11 bp. The behavior of the other strain variables in eqn (4.39) is similar. We plot the interaction energy $\Delta G(a)$ in (b) for the boundary conditions indicated in the legend of (a) and use it to verify the period we predict using this argument.

nm^{-1} , $k_{12} = k_{22} = 0.05 nm^{-1}$ and $r_1 = -r_2 = 0.02 nm$. The magnitude of the curvature $|k| = \sqrt{k_1^2 + k_2^2} \approx 0.055 nm^{-1}$ which gives a radius of curvature in fig. (4.5)b (inset) $R = 1/k \approx 18 nm$. Assuming the centerline does not extend and the pitch of the DNA helix is 3.4 nm, we get $\theta = 3.4/18 \approx 0.18 rad$. The radius of the DNA molecule is $a = 1 nm$, thus the change in the groove width is approximately given by $\Delta = (R + a)\theta - R\theta = 1 \times 0.18 nm \approx 2 \text{ \AA}$. This value is close to the one reported by Kopka *et al* (Kopka *et al.* (1985)) (0.5 – 2 \AA). Hence, our choice of curvature boundary conditions is reasonable; it is, however, not unique and it is coupled with the choice of stiffnesses of the webbing in our birod model. Be that as it may, our exercise above demonstrates that a birod model can capture the dependence of interaction energy on the distance between proteins bound to DNA. Calibration of the model and faithfully connecting it to experiment will require deeper analysis, and perhaps also, computation. It is important to note here that an elastic model of DNA allostery, consistently parametrized from atomic-resolution molecular dynamics simulations, has already been proposed by Dršata *et al.* (Drsata *et al.* (2014)) which was later extended to describe the experiment of Kim *et al.* (Dršata *et al.* (Drsata *et al.* (2016))). Our model is similar to the one presented by Dršata *et al.* in certain aspects:

1. Quadratic energy: Dršata *et al.* (Drsata *et al.* (2014)) use a quadratic energy to penalize deformations which is similar to our approach. This is appropriate for small deformations as assumed in our work and that of Dršata *et al.* The expression for the interaction energy in Dršata *et al.* (Drsata *et al.* (2016)) (eqn. (14) in their appendical information) is identical to our's (eqn. 4.40).
2. Boundary conditions: Here a protein interacts with the DNA helix by imposing boundary conditions on the curvatures and the radius of the double helix at the binding site. In Dršata *et al.*, a protein interacts with DNA by changing the width of the minor groove. These approaches are equivalent since extending the minor/major groove at the binding site causes bending of DNA.

Our model differs from Dršata *et al.* (Drsata *et al.* (2014)) on the following key points:

1. Effect of outer strands: The base-pair centric model of Dršata *et al.* (Drsata *et al.* (2014)) ignores the effect of the mechanics of the phosphate backbones. We model the backbones as inextensible rods (or worm-like chains of polymer elasticity (J. F. Marko & Siggia (1995))) capable of bending. Elasticity of the outer strands is crucial for twist-stretch coupling observed in DNA (refer sec. 3 in the appendix).
2. Stacking energy: We use a stacking energy quadratic in the twist and stretch of the centerline to account for the change in the orientations of successive base pairs

relative to each other. Our approach is simplistic compared to the holistic approach of Dršata *et al.* (Drsata et al. (2014)) where stacking energy has quadratic contributions from the six inter-basepair degrees of freedom.

3. Boundary conditions: One of the drawbacks of our model is its inability to account for the nature of protein-DNA contacts. Dršata *et al.* (Drsata et al. (2016)) point out that while some proteins such as *Bam*HI show tight protein-DNA contacts thereby constraining all inter- and intra-basepair degrees of freedom, others such as GRDBD interact rather loosely by altering only the major groove width. This flexibility is absent from our current model.

The period the interaction energy in fig. 4.5(a) is approximately 5.5 bp while that in fig. 4.5(b) is 11 bp as in the experiment. Why? Note that the strain variables in a two protein complex shown in fig. 4.6 (b) are a function of both the parameter x and the distance between the two proteins a . We fix x ($= 2$ nm from protein P_1) and focus on the dependence on a . We assume that both the proteins apply identical boundary conditions. If the proteins do not cause any change in the radius such that $r_0 = 0$, then the strain parameters involved in the elastic energy (eqn. (4.39)) $\propto e^{-\Gamma x} \psi(\omega a)$, where $\psi(\omega a)$ is a sinusoidal function oscillating with a period 11 bp, and the elastic energy of the two protein complex $\propto (e^{-\Gamma x} \psi(\omega a))^2$ oscillates with a period 5.5 bp. On the other hand, when the protein causes both a change in radius r_0 and a change in curvature k_{20} , the strain variables are $\propto (e^{-\Gamma_1 a} \psi(\omega a) + e^{-\Gamma_2 x})$ and the elastic energy of the two protein complex $\propto (e^{-\Gamma_1 a} \psi(\omega a) + e^{-\Gamma_2 a})^2$ oscillates with a period of 11 bp due to the cross term $e^{-(\Gamma_1 + \Gamma_2)a} \psi(\omega a)$. We plot the interaction energy $\Delta G(a)$ between the two proteins constituting the protein complex in fig. 4.6(b) and verify the periods for respective boundary conditions which resolves the apparent discrepancy in the periods in fig.4.5 (a) and (b). As a final application of our birod model we examine the sequence dependence of the allosteric interaction energy ΔG . While there is overwhelming qualitative evidence, both experimental (Kim et al. (2013)) and numerical (Gu et al. (2015)), showing that AT-rich sequences exhibit stronger allosteric interactions compared to GC-rich ones, a theoretical explanation is still lacking. Stronger interactions are associated with longer decay lengths. Using our theory we can find the dependence of the decay length on the elastic constants of the web. Since, AT base-pairs consist of two hydrogen bonds, the corresponding elastic constants for the web are expected to be lower than GC base-pairs which comprise of three hydrogen bonds. In an attempt to simulate such a scenario we replace the elastic constants for the web (K_c, K_e, L_i, H_i $i = 1, 2, 3$) in eqn. (4.41) with $(\chi K_c, \chi K_e, \chi L_i, \chi H_i$ $i = 1, 2, 3)$ while keeping EI fixed, and vary the parameter χ in the range $0.5 \leq \chi \leq 1$. We define a measure of the decay length l_d to be the inverse of the eigenvalue having the least non-zero magnitude, obtained in eqn. (4.42). For instance,

if $\chi = 1$, decay length $l_d = 1/0.34 \text{ nm} \approx 10 \text{ bp}$. We plot the variation of l_d with χ in fig. 4.7. We find that the decay length increases with the decrease in elastic constants of the web. We plot $\log l_d$ versus $\log \chi$ and deduce that $l_d \sim \frac{1}{\chi^{2/3}}$.

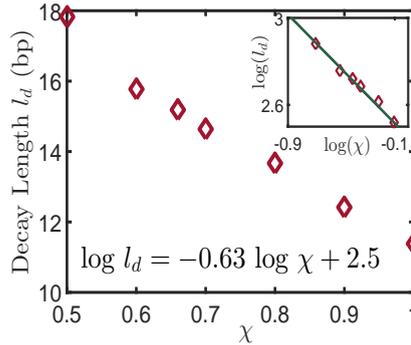


FIGURE 4.7: Decay length l_d is defined as the inverse of the eigenvalue with the least non-zero magnitude, for $\chi = 1$, $l_d = 1/|\lambda|_{min} = 1/0.34 \approx 9 \text{ bp}$. χ is meant to account for the reduction in the elastic constants for AT base-pairs compared to GC base-pairs. The elastic constants for the web are $(\chi K_c, \chi K_e, \chi H_1, \chi H_2, \chi H_3, \chi L_1, \chi L_2, \chi L_3)$, eqn. (4.41) gives the numerical values for $\chi = 1$. We find that the decay length increases with a decrease in elastic constants for the web, thus AT-rich DNA sequences are expected to have higher decay lengths. Qualitative experimental and numerical evidence in support of the above conclusion is documented in (Kim *et al.* (2013)) and (Gu *et al.* (2015)), respectively. The inset shows how we extracted the the power law $l_d \sim \chi^{-2/3}$.

4.5 Conclusion

Kim *et al.* (Kim *et al.* (2013)) have presented compelling quantitative evidence for allosteric interactions between two proteins bound to DNA at distant locations. They showed that the interaction energy for two proteins separated by distance a on DNA is a decaying exponential oscillating with period of 11 bp. We approach the problem from a purely mechanical standpoint and conjecture that the local deformation field in DNA caused by a bound protein is similar to that produced by a defect in an elastic solid. First, the interaction energy for two defects on a ladder is computed; it decays exponentially with the distance between them. Then, the same calculation is replicated for DNA by modelling it as a double-helical birod (Moakher & Maddocks (2005)). In the birod model for DNA, the outer phosphate backbones are represented by \pm strands and are assumed to be inextensible and unsharable; while the base-pairs are capable of elastic extension, shear, bending and, twisting. A general form of displacement for these strands (eqn. 4.2) is assumed which is used to calculate the micro-displacement and micro-rotation for the base-pairs. These expressions are then substituted into the balance laws. A crucial factor in our treatment is the boundary conditions. We follow Kwiecinski *et al.* (Kwiecinski *et al.* (2017)), Kim *et al.* (Kim *et al.* (2013)) and Liang

and Purohit (Liang & Purohit (2018b)) and impose boundary conditions on the curvatures and the radius of the DNA double-helix. The question, “what kind of boundary conditions a protein could possibly apply”, is not yet comprehensively addressed in the literature and is not the central issue of this study either. Rather the message of this chapter is that after solving the governing equations and plugging in boundary conditions, we recover the exponentially decaying profile that oscillates with a period of 11 bp. The last result of the chapter involves examining the sequence dependence of allosteric interactions—the model predicts that AT-rich sequences exhibit stronger interactions than GC-rich sequences.

Even though the birod model developed here does surprisingly well by capturing the dependence of interaction energy on distance there are many important caveats that we must point out. First, the birod model is not accurate near the site of protein binding. The deformations near the binding site could be large enough that a linear elastic theory may not be applicable. The assumptions that the outer strands are inextensible and the web is elastic could also break down in the vicinity of the binding site. Second, we have little knowledge of the elastic constants of the web. We have assumed some stiffness parameters for the web that gave the right experimentally verified moduli for the DNA, but there could have been another set of parameters that would have given similar results. One may have to appeal to molecular simulations (Lankaš et al. (2009); Olson (1996); Petkeviciute (2012); Drsata et al. (2014)) to get these parameters. Third, the boundary conditions applied by the proteins on the DNA are not clear. One may have to look for guidance from molecular simulations or protein-DNA co-crystal structures to get a clearer picture (Drsata et al. (2016)). Finally, we have not accounted for fluctuations or entropic interactions in our model. This is partly justifiable because the length of DNA between two protein binding sites for which significant allosteric interactions are observed is often much smaller than the persistence length of the DNA. However, a rigorous calculation should be done to verify this assumption. In spite of these shortcomings, the birod model could provide a starting point for analyzing allosteric interactions in DNA within the broad framework of configurational forces in elastic solids.

Chapter 5

Statistical mechanics of an elastic birod

5.1 Introduction

We combine ideas from statistical mechanics and continuum mechanics to study the temperature driven strand separation in DNA and explore the effect of tensile loads on the melting temperature as well. From the continuum mechanics side, we closely follow the birod framework presented in Moakher and Maddocks ([Moakher & Maddocks \(2005\)](#)). We import ideas from statistical mechanics to show that the average distance between the two strands in a birod increases steeply in an unbounded fashion as the temperature reaches the melting point. A non-quadratic interaction between the strands is essential to achieve this effect. This is nontrivial because for a quadratic interaction used hitherto in the thesis, the average change in the distance between the two strands is zero even as the temperature increases. We improve upon the previous work (([Dauxois et al. \(1993\)](#)),([Peyrard & Bishop \(1989\)](#))) by including the elasticity of the outer strands and shearing and bending rigidity of basepairs. Our work suggests that the elasticity of the outer strands is responsible for the cooperativity observed in the melting transition ([Gibbs-Davis et al. \(2007\)](#); [Nishigaki et al. \(1984\)](#)).

5.1.1 Kinematics

For studying the melting of DNA, we envision the following scenario, both strands are acted upon by identical force F . We assume small displacements throughout and confine ourselves to deformations in a plane.

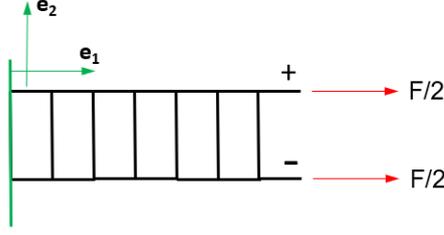


FIGURE 5.1: The cartoon of a straight birod acted upon by identical forces on both the strands. The two strands are referred to as \pm strands.

We begin by discussing the kinematics. The planar birod consists of two outer strands held together by a web resembling a ladder like shape. Both the outer strands and the web are elastic. The birod lies in $\mathbf{e}_1 - \mathbf{e}_2$ plane as shown in fig.5.1. The axial coordinate x is along \mathbf{e}_1 : $x \in [0, L]$ where L is the contour length. The reference configuration of the \pm outer strands denoted by \mathbf{r}^\pm is

$$\begin{aligned}\mathbf{r}_0^+ &= x\mathbf{e}_1 + a\mathbf{e}_2, \\ \mathbf{r}_0^- &= x\mathbf{e}_1 - a\mathbf{e}_2.\end{aligned}\tag{5.1}$$

For a general 2-D deformation in the $\mathbf{e}_1 - \mathbf{e}_2$ plane, the deformed configuration of the strands is denoted by:

$$\begin{aligned}\mathbf{r}^+ &= \int_0^x (1 + \zeta)\mathbf{d}_1 dx + (a + v)\mathbf{d}_2 + u\mathbf{d}_1, \\ \mathbf{r}^- &= \int_0^x (1 + \zeta)\mathbf{d}_1 dx - (a + v)\mathbf{d}_2 - u\mathbf{d}_1.\end{aligned}\tag{5.2}$$

where, $\mathbf{d}_1 = \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2$, and $\mathbf{d}_2 = -\sin\theta\mathbf{e}_1 + \cos\theta\mathbf{e}_2$. Note that $\mathbf{r} = \frac{\mathbf{r}^+ + \mathbf{r}^-}{2} = \int_0^x (1 + \zeta)\mathbf{d}_1 dx$ denotes the deformed centerline (global/macro displacement), while u and v denote the displacements of the \pm strands w.r.t. to the centerline (micro-displacements). Heretofore, $()_x$ denotes derivative w.r.t x . Hence, $\mathbf{d}_{1x} = \theta_x\mathbf{d}_2$ and $\mathbf{d}_{2x} = -\theta_x\mathbf{d}_1$. Now, $\mathbf{r}_x^\pm = (1 + \zeta \pm u_x \mp a\theta_x)\mathbf{d}_1 \pm v_x\mathbf{d}_2$. We assume that the outer strands are inextensible which implies $|\mathbf{r}_x^\pm| = 1$. Hence $\zeta \pm u_x \mp a\theta_x = 0$, which gives $\zeta = 0$, and $u_x = a\theta_x$. Hence $u = a\theta$. The tangent to the \pm strands is $\mathbf{t}^\pm = \mathbf{d}_1 \pm v_x\mathbf{d}_2$. The curvature for \pm strands $\kappa^\pm = |\mathbf{t}_x^\pm| = \theta_x \pm v_{xx}$.

We now focus on the stretching, shearing, twisting and bending of the base pairs. The tangent vectors for the \pm strands are $\mathbf{t}^\pm = \mathbf{d}_1 \pm v_x\mathbf{d}_2$, and the respective normal vectors are $\mathbf{n}^\pm = \mp v_x\mathbf{d}_1 + \mathbf{d}_2$. The rotation matrices $\mathbf{Q}^\pm = [\mathbf{t}^\pm, \mathbf{n}^\pm]$ can be decomposed as,

$$\mathbf{Q}^\pm = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & \mp v_x \\ \pm v_x & 1 \end{bmatrix}.\tag{5.3}$$

$\mathbf{Z} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ depends on the deformation of the center-line, while $\begin{bmatrix} 1 & \mp v_x \\ \pm v_x & 1 \end{bmatrix}$ depends only on the displacements of the strands about the center-line. \mathbf{Q}^\pm can be used to compute the the micro-rotation tensor \mathbf{P} and macro-rotation tensor \mathbf{Q} (for details see (Moakher & Maddocks (2005))).

$$\mathbf{P}^2 = \mathbf{Q}^+ \mathbf{Q}^{-T} = \mathbf{Z} \left(\mathbf{I}_{2 \times 2} + \begin{bmatrix} 0 & -2v_x \\ 2v_x & 0 \end{bmatrix} \right) \mathbf{Z}^T, \quad (5.4)$$

where $\mathbf{I}_{2 \times 2}$ is the identity tensor. Hence,

$$\mathbf{P} = \begin{bmatrix} 1 & -v_x \\ v_x & 1 \end{bmatrix}. \quad (5.5)$$

Also,

$$\mathbf{Q} = \mathbf{P} \mathbf{Q}^- = \mathbf{Z} = \mathbf{d}_1 \otimes \mathbf{e}_1 + \mathbf{d}_2 \otimes \mathbf{e}_2 \quad (5.6)$$

The stretching and shearing of the base-pairs are proportional to $\xi = \mathbf{Q}^T \mathbf{w} - \mathbf{w}_0$ (Moakher & Maddocks (2005)), where $\mathbf{w} = \frac{\mathbf{r}^+ - \mathbf{r}^-}{2}$, $\mathbf{w}_0 = \frac{\mathbf{r}_0^+ - \mathbf{r}_0^-}{2}$, and $\xi = a\theta \mathbf{d}_1 + v \mathbf{d}_2$.

Similarly the bending and twisting of the base pairs is equal to the difference between the rotations on the + and - strand. In fact, the moment transferred by the web is proportional to the Gibbs vector of the micro-rotation tensor \mathbf{P} (Moakher & Maddocks (2005)). In a 2D-setting, this moment can be easily computed; since the directions of the rotations are fixed, there is no twisting, and the bending is proportional to v_x .

We now discuss the energy associated with each kinematic deviation from the reference configuration. The outer strands are inextensible and unshearable. The bending energy per unit length for the outer \pm strands is $E_s = \frac{EI}{2}(\kappa^{+2} + \kappa^{-2}) = EI(\theta_x^2 + v_{xx}^2)$. The energy associated with shearing the base-pairs is $E_{sh} = L_1(u_0 + a\theta)^2$, where L_1 is the associated elastic constant. To account for the stretching of the base pairs, we ascribe an energy which penalizes the steric hindrance between the two strands using an assymmetric energy profile of the form $f(x) = (e^{-\lambda x} - 1)^2$ instead of the symmetric x^2 (Peyrard & Bishop (1989); Dauxois et al. (1993)). Note that for small x , both are identical. Hence, the energy per unit length required to stretch the base-pairs $E_{st} = L_2(e^{-\lambda v} - 1)^2$. The energy associated with the bending of the base-pairs is $E_{bb} = H_1 v_x^2$. Altogether, the

energy of the birod is,

$$\begin{aligned} E_e &= \int_0^L dx (E_b + E_{sh} + E_{st} + E_{bb}) \\ &= \int_0^L dx (EI(\theta_x^2 + v_{xx}^2) + L_1 a^2 \theta^2 + L_2 (e^{-\lambda v} - 1)^2 + H_1 v_x^2), \end{aligned} \quad (5.7)$$

where both v and θ are functions of x . This energy will enter the statistical mechanical model for the birod.

5.2 Force-displacement curve for birod

We start by discussing the force-displacement curve for the ladder-like birod. We already have the elastic energy of the birod given by Eq. 5.7. Next we need to compute the work done by external force F distributed equally on both the strands. For the $+$ strand the displacement at the free end is,

$$\Delta^+ = \int_0^L (\mathbf{t}^+ \cdot \mathbf{e}_1 - 1) dx = \int_0^L (\cos \theta - \sin \theta v_x - 1) dx = \int_0^L \left(-\frac{\theta^2}{2} - \theta v_x\right) dx. \quad (5.8)$$

Similarly, for the $-$ strand, $\Delta^- = \int_0^L \mathbf{t}^- \cdot \mathbf{e}_1 dx - L$, which upon simplification yields $\Delta^- = \int_0^L \left(-\frac{\theta^2}{2} + \theta v_x\right) dx$. Summing up the individual contributions from the strands yields,

$$W_e = \frac{F}{2}(\Delta^+ + \Delta^-) = F \int_0^L -\frac{\theta^2}{2} dx. \quad (5.9)$$

Hence, the potential energy of the birod is,

$$E = E_e - W_e = \int_0^L dx (EI(\theta_x^2 + v_{xx}^2) + L_1 a^2 \theta^2 + L_2 (e^{-\lambda v} - 1)^2 + H_1 v_x^2 + \frac{F}{2} \theta^2) = E(\theta(x), v(x)). \quad (5.10)$$

The average end-to-end extension is,

$$y = \int_0^L \cos \theta dx \approx \int_0^L \left(1 - \frac{\theta^2}{2}\right) dx, \quad (5.11)$$

where we assumed that θ is small. To compute the average end-to-end distance $\langle y \rangle$,

$$\langle y \rangle = L - \left\langle \int_0^L \frac{\theta^2}{2} dx \right\rangle, \quad (5.12)$$

where $\langle \rangle$ denotes the ensemble average. We need to evaluate the partition function to compute the above average. The partition function Z is a path-integral given as follows,

$$Z = \int \mathcal{D}\theta(x) \int \mathcal{D}v(x) \exp\left(-\frac{E(\theta(x), v(x))}{k_B T}\right). \quad (5.13)$$

The above partition function Z can be multiplicatively decomposed: $Z = Z_v Z_\theta$, where Z_v comprises the path-integral over the function $v(x)$, while Z_θ over $\theta(x)$.

$$Z_\theta = \int \mathcal{D}\theta(x) \exp\left(-\frac{E_\theta}{k_B T}\right), \quad \text{and} \quad Z_v = \int \mathcal{D}v(x) \exp\left(-\frac{E_v}{k_B T}\right), \quad (5.14)$$

where,

$$\begin{aligned} E_\theta &= \int_0^L dx \left(EI\theta_x^2 + L_1 a^2 \theta^2 + \frac{F}{2} \theta^2 \right), \quad \text{and} \\ E_v &= \int_0^L dx \left(EIv_{xx}^2 + L_2 (e^{-\lambda v} - 1)^2 + H_1 v_x^2 \right), \end{aligned} \quad (5.15)$$

Now,

$$\left\langle \int_0^L \frac{\theta^2}{2} dx \right\rangle = \frac{1}{Z} \int \mathcal{D}\theta(x) \int \mathcal{D}v(x) \left(\int_0^L \frac{\theta^2}{2} dx \right) \exp\left(-\frac{E(\theta(x), v(x))}{k_B T}\right). \quad (5.16)$$

Summing over all the admissible functions $v(x)$ and canceling the common factor Z_v yields,

$$\left\langle \int_0^L \frac{\theta^2}{2} dx \right\rangle = \frac{1}{Z_\theta} \int \mathcal{D}\theta(x) \left(\int_0^L \frac{\theta^2}{2} dx \right) \exp\left(-\frac{E_\theta(\theta(x))}{k_B T}\right), \quad (5.17)$$

The above expression can be evaluated by differentiating the logarithm of the partition function.

$$\left\langle \int_0^L \frac{\theta^2}{2} dx \right\rangle = -k_B T \frac{\partial \ln Z_\theta}{\partial F}, \quad (5.18)$$

which gives,

$$\langle y \rangle = L + k_B T \frac{\partial \ln Z_\theta}{\partial F}. \quad (5.19)$$

Notice that the remaining functional is only a function of $\theta(x)$. To evaluate the partition function Z_θ , we discretize the domain $x \in [0, L]$ into n -segments $(x_i, x_{i+1}]$, where $0 \leq i \leq n$, such that $\theta_x = \frac{\theta_i - \theta_{i-1}}{\delta}$ where $\delta = \frac{L}{n}$. For the energy functional E_θ , the integral

over the domain can be expressed as a quadratic expression in terms of θ_i 's:

$$E_\theta^n = \sum_{i=0}^n \delta \left[EI \left(\frac{\theta_i - \theta_{i-1}}{\delta} \right)^2 + L_1 a^2 \theta_i^2 + \frac{F}{2} \theta_i^2 \right] = \boldsymbol{\theta} \cdot \left[\frac{EI}{\delta} \mathbf{A} + \delta (L_1 a^2 + \frac{F}{2}) \mathbf{I} \right] \boldsymbol{\theta} = \boldsymbol{\theta} \cdot \mathbf{K}_\theta \boldsymbol{\theta}, \quad (5.20)$$

where $\boldsymbol{\theta} = [\theta_0, \theta_2, \dots, \theta_n]^T$, \mathbf{I} is an identity matrix, and \mathbf{A} is another matrix as follows:

$$\mathbf{A}_{5 \times 5} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$$

Notice that \mathbf{K}_θ is a constant depending only on the elastic properties of the birod. This enables us to transform the path integral Z_θ into a n -dimensional integral as follows:

$$Z_\theta^n = \int \mathcal{D}\theta(x) \exp\left(-\frac{E_\theta}{k_B T}\right) = \prod_{i=0}^n \left(\int_{-\pi}^{\pi} d\theta_i \right) \exp\left(-\frac{E_\theta^n}{k_B T}\right) = \int d\boldsymbol{\theta} \exp\left(-\frac{\boldsymbol{\theta} \cdot \mathbf{K}_\theta \boldsymbol{\theta}}{k_B T}\right). \quad (5.21)$$

To evaluate the above integral conveniently, we change the limits from $-\pi, \pi$ to $-\infty, \infty$, which transforms the above expression into a n -dimensional Gaussian integral which can be computed analytically as follows,

$$Z_\theta^n = \int_{(-\infty, \infty)^{n+1}} d\boldsymbol{\theta} \exp\left(-\frac{\boldsymbol{\theta} \cdot \mathbf{K}_\theta \boldsymbol{\theta}}{k_B T}\right) = \pi^{n/2} \frac{(k_B T)^{n/2}}{\sqrt{\det \mathbf{K}_\theta}}. \quad (5.22)$$

Substituting it in Eq. 5.19 gives,

$$\langle y \rangle = L + k_B T \frac{\partial \ln Z_\theta}{\partial F} = L - \frac{k_B T}{2} \frac{\partial \log \det \mathbf{K}_\theta}{\partial F}. \quad (5.23)$$

We can now plot the force-extension curve for the birod. The results are documented in Fig. 5.2(a). As is characteristic of entropic elasticity, the tensile force increases sharply as the end-to-end distance $\langle y \rangle$ reaches close to the contour length of the molecule (J. Marko & Siggia (1995)). The next step is to determine the persistence length l_p of the birod model. We do so by fitting the data to the expression for the worm-like-chain picked from (Petrosyan (2017)):

$$\frac{F l_p}{k_B T} = 0.25 \left(1 - \frac{\langle y \rangle}{L} \right)^{-2} - 0.25 + \frac{\langle y \rangle}{L} - 0.8 \left(\frac{\langle y \rangle}{L} \right)^{2.15} \quad (5.24)$$

where F is the applied force and L is the contour length. The best fit to the data shown in Fig.5.2(a) gives $l_p = 36$ nm, which is close to the value for ds-DNA ($l_p = 50$ nm). This ensures that the values of the elastic constants used to obtain the above data ($EI = 0.8$ pNnm², $L_1 = 80$ pN/nm², $L_2 = 1280$ pN/nm², $H = 0.33$ pN, $a = 1$ nm, and $\lambda = 0.5 \text{ \AA}^{-1}$) are biologically relevant. We observe a decrease in the end-to-end distance $\langle y \rangle$ as the temperature increases as shown in Fig.5.2(c) which is a characteristic of entropic elasticity. The detailed microscopic description provided by the model enables us to examine the effect of various parameters therein. For instance, the force-extension exhibits higher stiffness as the stretch moduli of web L_1 increases (see Fig.5.2(b)). This observation agrees with the qualitative experimental observation that the oligomer's stiffness increases as the GC content increases (X. Zhang et al. (2012)). This is because the GC base pair consists of three H bonds and is consequently stiffer compared to the AT base pair which consists of two H bonds. Similar stiffening can be achieved by increasing the stiffness of the outer strands (see Fig.5.2(d)).

5.3 Application to DNA melting

At room temperature under zero tensile loads, the DNA molecule exhibits a double helical structure. However, as the temperature increases and reaches the melting temperature, the complimentary base-pairing is disrupted and the two strands spontaneously disintegrate into two single strands. This transition is highly cooperative, and is known as the melting transition (Rouzina & Bloomfield (1999)), and the temperature at which it occurs is referred to as the melting temperature. Aside from the sequence dependence, the melting temperature is also highly sensitive to the tensile loads and the ionic concentration of the solution (Rouzina & Bloomfield (2001a,b)). The experimental evidence suggests that the melting temperature increases with the increase in ionic concentration and drops with the increase in tensile loads on the molecule. Thermodynamics based studies relying on Clausius-Clayperon equation have led to various empirical relations among these quantities (Rouzina & Bloomfield (2001a,b, 1999)). For the Na⁺ concentration of 0.075 M, the melting temperature is approximately 75°C (Blake & Delcourt (1998)) (The melting temperature is highly sensitive to the base-pair sequence comprising the DNA oligomer, see Blake & Delcourt (1998) for the exact sequence.). In this section, the birod model developed above is used to study DNA melting. We seek a relation between the average inter-strand distance $\langle v \rangle$ and the temperature T . We assume no tensile forces on the molecule, hence the elastic potential energy E is,

$$E = \int_0^L dx (EI(\theta_x^2 + v_{xx}^2) + L_1 a^2 \theta^2 + L_2 (e^{-\lambda v} - 1)^2 + H_1 v_x^2). \quad (5.25)$$

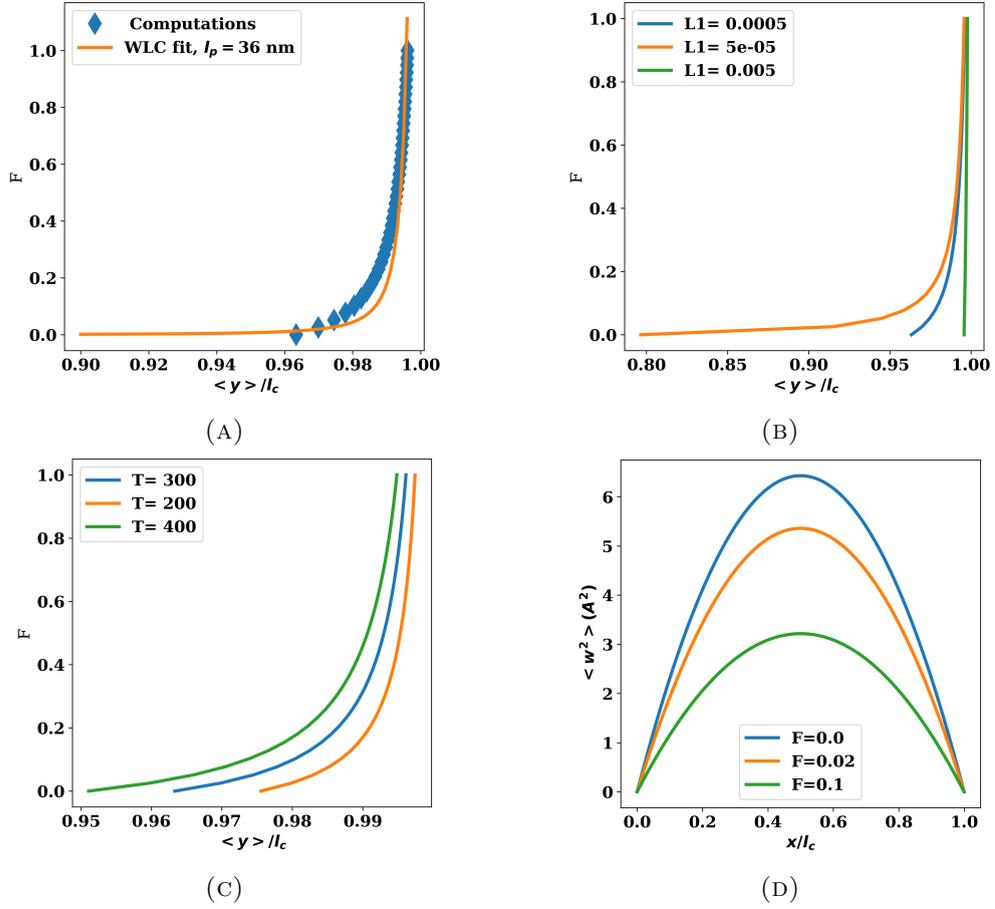


FIGURE 5.2: (a) Force-extension curve for the birod. The parameter values are $EI = 0.8 \text{ pNnm}^2$, $L_1 = 80 \text{ pN/nm}^2$, $L_2 = 1280 \text{ pN/nm}^2$, $H = 0.33 \text{ pN}$, $a = 1 \text{ nm}$, and $\lambda = 0.5 \text{ \AA}^{-1}$. The unit for the force values shown in the graph are $eV/\text{\AA}$ ($1 \text{ eV}/\text{\AA} = 1600 \text{ pN}$). Using the WLC chain formula (Grosberg et al. (1995)), the best fit to the data yields a persistence length of $l_p = 36 \text{ nm}$, which is close to the persistence length of ds-DNA— 50 nm (). (b) Effect of increase in L_1 . The force-displacement curve is independent of stretch modulus L_2 and the bending rigidity H of the web. (c) Effect of increase in temperature (d) Effect of increase in the stiffness of the outer strands. (d) $\langle w^2 \rangle$ vs x for various values of tensile force F with hinged-hinged boundary conditions.

The values of the constants are given in the previous section. The average distance between the strands can be computed as follows:

$$\langle v \rangle = \frac{1}{Z} \int \mathcal{D}\theta(x) \int \mathcal{D}v(x) \left(\frac{1}{L} \int_0^L v(x) dx \right) \exp\left(-\frac{E(\theta(x), v(x))}{k_B T}\right), \quad (5.26)$$

where the expression for the energy E and the partition function Z can be found in Eq. 5.7 and Eq. 5.13, respectively. As done in the previous section, we discretize the domain into n -elements which transforms the integrals into sums and the path integrals into

n -dimensional integrals:

$$\begin{aligned}
 E^n &= \sum_{i=0}^n \delta \left[EI \left(\frac{(\theta_i - \theta_{i-1})^2}{\delta^2} + \frac{(v_{i+1} - 2v_i + v_{i-1})^2}{\delta^2} \right) \right. \\
 &\quad \left. + L_1 a^2 \theta_i^2 + L_2 (e^{-\lambda v_i} - 1)^2 + H \frac{(v_i - v_{i-1})^2}{\delta^2} \right], \\
 Z^n &= \prod_{i=0}^n \left(\int_{(-\infty, \infty)^2} d\theta_i dv_i \right) \exp\left(-\frac{E^n}{k_B T}\right), \\
 \langle v \rangle &= \frac{1}{Z^n} \prod_{i=0}^n \left(\int_{(-\infty, \infty)^2} d\theta_i dv_i \right) \left(\frac{1}{n} \sum v_i \right) \exp\left(-\frac{E^n}{k_B T}\right).
 \end{aligned} \tag{5.27}$$

In contrast to the previous section where the discretization together with quadratic energy functional enabled us to analytically evaluate the partition function, the partition function above can not be evaluated analytically because of the non-quadratic term $(e^{-\lambda v_i} - 1)^2$. Hence, we use Monte-Carlo simulations to compute $\langle v \rangle$ as a function of the bath temperature T . We use the Metropolis algorithm ([Pathria \(1984\)](#)) to perform the MC simulations, the details are available in the appendix. The results are recorded in Fig. 5.3. Each individual \times is one simulation. We find that as the temperature increases the average inter-strand distance increases. The increase is nonlinear, hence can not be alluded to a mere thermal expansion. The nonlinear interaction $((e^{-\lambda v} - 1)^2)$ and the cooperativity terms (v_{xx}^2) are crucial for achieving this effect. For instance, if only quadratic interaction is used the average inter-strand distance is zero even as the temperature increases.

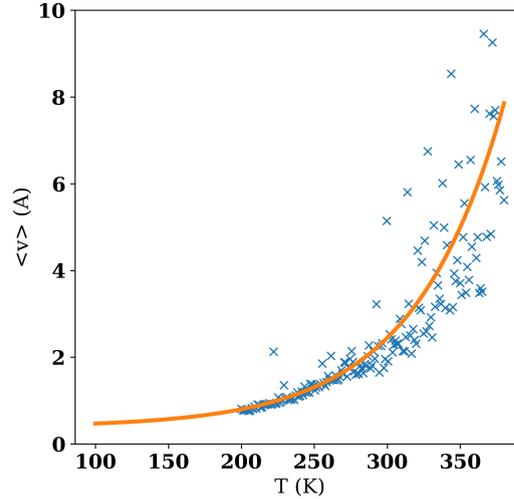


FIGURE 5.3: The average distance between the two strands $\langle v \rangle$ as a function of the temperature T .

To model the above data, we fit a continuous curve of the form $y = a_0 e^{a_1(x-a_2)}$. In the reference configuration, the average inter-strand distance is 10 \AA . We assume that the melting occurs at 50% strain, *i.e.* when the inter-strand distance reaches 15 \AA or when $\langle v \rangle = 5 \text{ \AA}$. Based upon the above criteria, the melting temperature can be read off from the graph: $T_m \approx 354 \text{ K}$, which is quite close to the results for dsDNA documented in literature (Rouzina & Bloomfield (2001a,b); Blake & Delcourt (1998)) for dsDNA. Although, the melting criterion $\langle v \rangle = 5 \text{ \AA}$ was deliberately chosen so that the results from the model agree with the experimental data, our main message is that the birod model has the essential ingredients—the non-quadratic interaction and the cooperativity arising from the elasticity of the outer strands—to account for the cooperative melting transition. The novelty of this work lies in the fact that these crucial factors emerge naturally from the kinematic description of the birod.

5.4 Effect of tensile force on the melting temperature

In this section, the effect of tensile force on the melting temperature is explored. We consider the birod shown in Fig.5.4. Similar geometries have been used earlier to study related problems (de Gennes (2001)). As before, we need to compute the potential energy of the birod shown here.

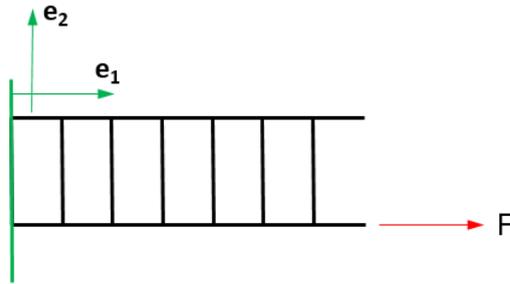


FIGURE 5.4: The birod cartoon for studying DNA melting. Note that the force is applied only on one strand.

The elastic energy of the birod is available in Eq.5.7. The work done by external force F is,

$$W_e = F \int_0^L dx (\mathbf{t}^- \cdot \mathbf{e}_1 - 1). \quad (5.28)$$

Now, $\mathbf{t}^- = \mathbf{d}_1 - v_x \mathbf{d}_2$, hence $\mathbf{t}^- \cdot \mathbf{e}_1 = \cos \theta + v_x \sin \theta \approx 1 - \theta^2/2 + v_x \theta$. This implies,

$$W_e = F \int_0^L dx \left(-\frac{\theta^2}{2} + v_x \theta \right). \quad (5.29)$$

The elastic potential energy E is,

$$E = E_e - W_e = \int_0^L dx \left(EI(\theta_x^2 + v_{xx}^2) + (L_1 a^2 + \frac{F}{2})\theta^2 + L_2(e^{-\lambda v} - 1)^2 + H_1 v_x^2 - F v_x \theta \right). \quad (5.30)$$

The term $Fv_x\theta$ is responsible for coupling the force F and the inter-strand distance v . Now for a given θ , the probability of observing a configuration with interstrand distance v at force F is $e^{Fv_x\theta/k_B T}$ times the probability of observing the same configuration at $F = 0$. As the force F increases the birod straightens out and θ decreases as can be inferred from Fig.5.2(b). However, for high forces (> 15 pN), the DNA molecule with a persistence length 50 nm is mostly straight. In other words, the $\theta(x)$ should not change much as the force increases from 15 pN to 40 pN, however this makes the higher values of $v(x)$ much more likely. Based upon this qualitative argument we expect that the melting temperature should decrease with the increase in the tensile force.

The $\langle v \rangle$ vs T curves are presented in Fig.5.5(a) for various tensile forces F , and we indeed observe that for a given temperature, the inter-strand separation increases with increasing tensile loads. We use the same criteria for computing the melting temperature T_m as in the previous section: the temperature at which increase in the average inter-strand distance $\langle v \rangle$ becomes 5\AA . We plot the dependence of melting temperature on the tensile force in Fig.5.5(b). The experimental data for DNA melting is taken from Zhang *et al* (X. Zhang *et al.* (2012)). Although our model overestimates the drop in melting temperature with force, the trend is correct nonetheless. We made several assumptions such as modeling the DNA using a straight birod instead of a double helical one. Also, we restricted ourselves to deformations on a plane. These assumptions could be responsible for the observed deviation from the experimental data. However, our primary contribution here is to demonstrate that the elastic birod formulation is capable of qualitatively accounting for the effect of force on the melting temperature.

5.5 Conclusion

We have applied the theory of elastic birods to study DNA melting and explore the effect of tensile force on melting temperature. We began by discussing the force-extension curve of a birod, which resembles that of a typical WLC model with a persistence length of approximately 36 nm. Also, the model predicts a decrease in persistence length as the temperature increases. The microscopic insights provided by the birod model helped us explain why the DNA molecules with higher GC content are stiffer. Next, we used

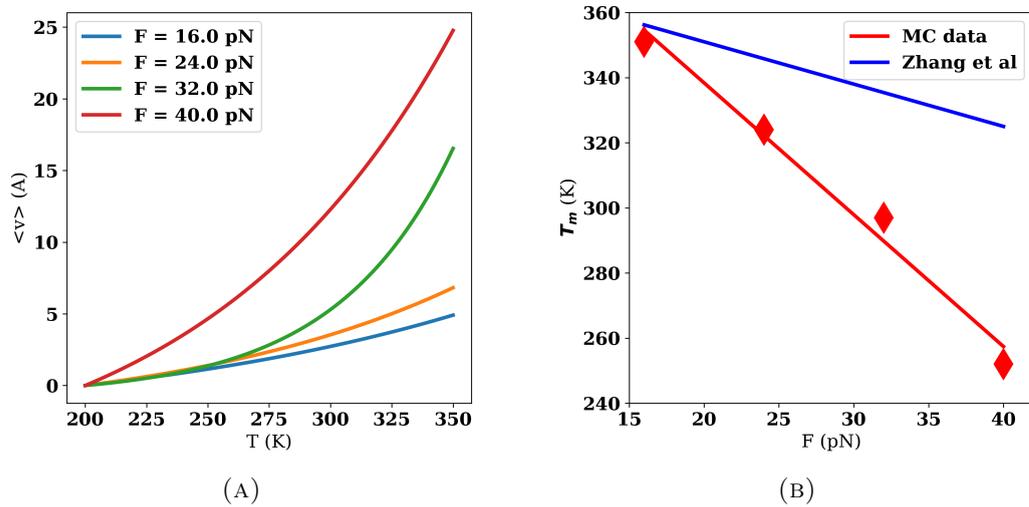


FIGURE 5.5: (a) The effect of force on the the $\langle v \rangle$ vs T curves (b) The effect of the tensile force on the melting temperature. The experimental data is from [X. Zhang et al. \(2012\)](#).

our model to study DNA-melting. We computed the inter-strand distance as a function of temperature and found that it increases monotonically with increasing temperature. The sudden sharp increase distinguishes it from thermal expansion. We note that the nonlinear asymmetric interactions between the strands is crucial to get the melting transition; for a linear interaction, leading to a quadratic energy functional, the average increase in the inter-strand distance is zero-independent of changes in temperature. The elasticity of the outer strands makes this melting transition highly cooperative. Next, we explored the effect of tensile forces on the melting temperature. Our model predicts that the melting temperature decreases with increasing tensile force. Although, we correctly captured the qualitative trends to some extent, the results from our model showed deviations from the experimental data. Various assumptions such as using a straight birod to model double helical DNA and restricting to deformations on a plane could be responsible for the deviations. The work is novel for two reasons: i) we find that the cooperativity in the melting transition, well documented in literature ([Dauxois et al. \(1993\)](#); [Peyrard & Bishop \(1989\)](#); [Rouzina & Bloomfield \(2001b\)](#); [X. Zhang et al. \(2012\)](#)), emerges naturally from the elasticity of the outer strands, and ii) we demonstrated how the birod model can successfully account for the effect of tensile force on the melting temperature.

However, using a straight ladder like birod to model double helical DNA is at best a first order approximation. It can not account for crucial features of DNA arising out of double-helical topology such as twist stretch coupling. Although, it can explain melting transition, it can not explain well documented ([X. Zhang et al. \(2012\)](#); [Sarkar et al. \(2001\)](#)) transitions among various DNA phases such B-DNA \rightarrow S-DNA and B-DNA

→ P-DNA. Hence, a natural extension of the work is to do the statistical mechanics in a double-helical birod. Next, these structural transitions are highly sensitive to ion concentration of the solution ([X. Zhang et al. \(2012\)](#)). A birod model can potentially provide useful insights into microscopic understanding of electrostatic interactions.

Chapter 6

Conclusion

This thesis is devoted to studying various mechanical aspects of DNA using a combination of techniques from elasticity, geometry, thermodynamics, and statistical mechanics. The motivation behind developing these models is to provide a mechanics-based understanding for several biologically relevant phenomena across various length scales. Our studies highlight mechanical deformations and configurational forces as causal agents in allosteric interactions on DNA restricted to tens of nm to structural transitions where the length of DNA oligomer could be as large as a few μm .

In chapter 2, a homogeneous elastic rod model [J. F. Marko & Neukirch \(2013\)](#); [Argudo & Purohit \(2014b\)](#) of DNA is used to probe the structural transitions in DNA. We confine ourselves to 3-phases of DNA: B-DNA (ground state), S-DNA (over-stretched state), and P-DNA (over-twisted state). The energy functional, quadratic in force and torque, combined with Zimm-Bragg helix-coil transition theory resulted in a partition function that could be summed up using the transfer matrix approach. The model was applied to study the over-stretching transition in a torsionally constrained DNA molecule, through the triple point of B-, S-, and P-DNA. The resulting force-extension curve resembles a plateau like sigmoidal transition where majority of the extension that takes place at a constant force goes into changing the phase of DNA. The model incorporated the effect of electrostatic interactions between the negatively charged backbone and the dissolved ions in the solution. The results from the model not only quantitatively matched with the experimental observations, they provided valuable insights into possible causes behind some unconventional experimental observations such as the non-monotonic variation of the over-stretching force with an increase in the ion-concentration.

In chapter 3 and 4, we discuss a mechanical model of DNA that uses the theory of an elastic birod [Moakher & Maddocks \(2005\)](#) to investigate the allosteric interactions between two ligands bound to DNA. These interactions are quantified using an interaction energy. Experimental measurements show that the interaction energy decays exponentially with the distance between the two ligands while oscillating with the periodicity of the DNA double helix. The birod formalism explains this dependence by accounting for two crucial features of DNA: the elasticity of the base-pairs responsible for the exponential decay and the double helical geometry responsible for the sinusoidal oscillations. The detailed microscopic understanding provided by the model distinguishes it from the existing literature on allosteric interactions [Koslover & Spakowitz \(2009\)](#). Following [Kim *et al.* Kim *et al.* \(2013\)](#), the DNA binding proteins are subsumed into two categories: ones that bind to straight DNA (dealt with in chapter 3) and others that bend DNA (dealt with in chapter 4). In addition to the elastic energy of a birod arising from the mechanical distortions of the outer-strands and the base pairs, we also account for the stacking energy of the base-pairs arising out of the local changes in their orientation. The interaction energy computations from the model, after fitting certain parameters, quantitatively match with the experimental data. Additionally, we compute biologically relevant quantities such as changes in the minor/major groove width of DNA which also matched with the data available from the MD simulations.

In chapter 5, we combine the elastic birod model with statistical mechanics and apply it to various problems such as the force-extension curve of dsDNA, temperature driven strand separation in dsDNA, and the effect of tensile loads on the melting temperature. Here, for analytical tractability, we model DNA as a straight ladder-like birod restricted to deformations on a plane. The force-extension curve of the birod shows typical characteristics of entropic elasticity—a sharp increase in force as the end-to-end distance becomes close to the contour length. The response can be approximated by a worm-like-chain model with a persistence length of a few tens of nm which is close to the persistence length of DNA (50 nm). The model captures the stiffening of the DNA oligomers as the GC content increases. We note that to model DNA melting the steric hindrance between the two strands needs to be accounted for. To do so, we penalize the deformations of the web using an asymmetric non-quadratic potential because for a quadratic potential the average distance between the strands is zero even as the temperature increases. Our model exhibits a sudden increase in the inter-strand separation as the temperature increases. Also, we demonstrate how cooperativity in the melting transition emerges from the elasticity of the outer strands. The model predicts that the melting temperature decreases with an increase in the tensile load in agreement with

the experimental observations.

It is clear that DNA elasticity governs the behavior of the molecule at various length scales. The simplifying assumptions made in the models naturally pave the way forward. The electrostatic interactions between the negatively charged phosphate backbone and the solution ions are poorly understood especially in the presence of divalent ions. This is important because experimental evidence [King et al. \(2016\)](#) suggests that the stabilizing effect of divalent ions such as Mg^{+2} and Ca^{+2} is stronger compared to the monovalent ions such as Na^+ and K^+ . We already used the birod model to study the temperature driven cooperative transition from dsDNA to ssDNA. The next step would be to use a double helical birod to model DNA and apply it to other cooperative transitions such as B-DNA \rightarrow S-DNA and B-DNA \rightarrow P-DNA. The mechanics-based understanding from the model can potentially provide useful insights into building artificial molecular motors [Bryant et al. \(2003\)](#).

Appendix A

Supplement to Chapter 2

A.1 Dependence of Gibbs free energy on GC content

In order to incorporate sequence dependence into the physics of structural transitions in DNA we assume that its Gibbs free energy depends on the fraction of GC base-pairs r_{GC} (in addition to force f , moment M and temperature T), *i.e.* $G = G(f, M, T, r_{GC})$. Consider a torsionally unconstrained DNA where $M = 0$ and $c_0 = 150$ mM (the experimental data we use is at 150 mM), hence $G = G(f, T, r_{GC})$. We will obtain the functional dependence of the $G(f, T, r_{GC})$ on r_{GC} by examining the dependence of overstretching force f_{ov} and melting temperature T_m on r_{GC} .

Consider the transition B -DNA \rightarrow ss -DNA at constant temperature $T = T_0$. Define $\Delta G(f, T_0, r_{GC}) = G_B(f, T_0, r_{GC}) - G_{ss}(f, T_0)$. We assume that the free energy for ss -DNA does not depend upon r_{GC} since base-pairing is ripped apart. For such an equilibrium transition

$$\begin{aligned}\Delta G(f_{ov}, T_0, r_{GC}) &= 0, \\ d\Delta G(f_{ov}, T_0, r_{GC}) &= 0.\end{aligned}\tag{A.1}$$

Hence, at the transition point,

$$\left. \frac{\partial \Delta G}{\partial r_{GC}} \right|_{T=T_0} + \left. \frac{\partial \Delta G}{\partial f} \right|_{T=T_0} \left. \frac{\partial f_{ov}}{\partial r_{GC}} \right|_{T=T_0} = 0.\tag{A.2}$$

Now it is known from experiment that the overstretching force f_{ov} is related linearly to r_{GC} *i.e.* $f_{ov} = a_1(T_0) + b_1 r_{GC}$, in such a manner that b_1 (=40pN) is independent of temperature [X. Zhang et al. \(2012\)](#). The plots used for extracting this relation involving f_{ov} and r_{GC} at constant T_0 are not shown here for the sake of brevity, but can be obtained from the phase diagram given in the figure 4 in [X. Zhang et al. \(2012\)](#). Also,

$\left. \frac{\partial \Delta G}{\partial f} \right|_{T=T_0} = -\Delta b$, where $\Delta b = 0.22$ nm is the average extension per bp at the transition point which is also independent of temperature. This enables us to conveniently integrate the resulting first order partial differential equation eqn. (A.2) to get

$$\Delta G = r_{GC} b_1 \Delta b + A(T_0, f), \quad (\text{A.3})$$

where we know $A(T_0, f) = G_B(f, T_0) - G_{ss}(f, T_0)$ (from Raj & Purohit (2011)) and the expressions for $G_B(f, T)$ and $G_{ss}(f, T)$ are available in the literature Raj & Purohit (2011); Rouzina & Bloomfield (2001a,b); Argudo & Purohit (2014b).

We verify our assertion A.3 by obtaining the dependence of melting temperature of DNA on r_{GC} and then cross checking it with the data given in X. Zhang et al. (2012). We consider a temperature driven phase transition from B-DNA to ss-DNA while keeping the force fixed at f_0 . Define $\Delta G(f_0, T, r_{GC}) = G_B(f_0, T, r_{GC}) - G_{ss}(f_0, T)$. Following exactly the same train of thought, $d\Delta G(f_0, T, r_{GC}) = 0$. Hence, at the transition point

$$\left. \frac{\partial \Delta G}{\partial r_{GC}} \right|_{f=f_0} + \left. \frac{\partial \Delta G}{\partial T} \right|_{f=f_0} \left. \frac{\partial T_m}{\partial r_{GC}} \right|_{f=f_0} = 0 \quad (\text{A.4})$$

Now $\left. \frac{\partial \Delta G}{\partial r_{GC}} \right|_{f=f_0} = b_1 \Delta b$ from eqn. (A.3) and $-\left. \frac{\partial \Delta G}{\partial T} \right|_{f=f_0} = \Delta S$ is the entropy change per base during the melting transition Rouzina & Bloomfield (2001b); King et al. (2013). This value is known to be 25 cal/K per mole Rouzina & Bloomfield (2001b) which is equal to 0.17 pNm/K per bp. Thus,

$$\left. \frac{\partial T_m}{\partial r_{GC}} \right|_{f=f_0} = \frac{b_1 \Delta b}{\Delta S} = \frac{40 \times 0.22}{0.17} = 51.76 K. \quad (\text{A.5})$$

Hence,

$$T_m = 51.76 r_{GC} + c(f_0) \quad (\text{A.6})$$

where $c(f_0)$ is a constant of integration that depends on force. This exactly matches with the result we obtain from the phase diagram presented by Zhang *et al.* X. Zhang et al. (2012).

A.2 Poisson-Boltzmann Equation

The objective of this section is to show how the Poisson-Boltzmann equation for solution electrostatics gives the logarithmic dependence of the overstretching force on the ion concentration. The negatively charged infinite cylinder in a monovalent solution naturally leads to a cylindrically symmetrical potential distribution. Let us examine the boundary

value problem posed in eqn. (2.11) when $|u(y)| \ll 1$, which implies $\sinh u \sim u$.

$$\begin{aligned} u''(y) + r^{-1}u'(y) &= \chi^2 u(y), \\ u'(a) = 2q/a, \quad u(R) &= 0, R \rightarrow \infty. \end{aligned} \quad (\text{A.7})$$

Due the above simplification the BVP admits a closed form analytical solution

$$u(y) = -2q \frac{K_0(\chi y)}{\chi a K_1(\chi a)}, \quad (\text{A.8})$$

where K_n is a modified Bessel function. We employ the asymptotic expansions of the Bessel functions given in [Abramowitz & Stegun \(1948\)](#) to gain further insights into the solution. In the limiting case, wherein $\chi a = a/r_D \ll 1$, which represents the situation when Debye length r_D is large compared to the radius of the idealized rod, we get

$$\begin{aligned} \text{For small } z \quad K_0(z) &\sim -\ln z \text{ and } K_1(z) \sim 1/z \\ u(y)|_{y=a} &= 2q \ln \frac{a}{r_D} \end{aligned} \quad (\text{A.9})$$

We can integrate the above expression for the potential to get the electrostatic free energy, so that

$$F^{el} = -2k_B T q \ln \frac{a}{r_D} = -2k_B T q \ln(\sqrt{8\pi l_B c_0} a) = -k_B T q \ln c_0 + k. \quad (\text{A.10})$$

The dependence of free energy on logarithm of concentration is reported in [Rouzina & Bloomfield \(2001a\)](#), [Rouzina & Bloomfield \(2001b\)](#) and [X. Zhang et al. \(2012\)](#). Note that the electrostatic energy calculated above is per base pair. The constant, k , is independent of force and concentration, but it is different for various phases of DNA. Note that the concentration, c_0 , is number of ions per nm^3 for Debye length r_D to be in nm .

$$\begin{aligned} c_0 \text{ mol/L} &= \frac{c_0(\text{mol/L}) \times 6.022 \times 10^{23}(\text{ions/mol}) \times 10^3(\text{L/m}^3)}{10^{27}(\text{nm}^3/\text{m}^3)} = 0.6c_0 \text{ ions/nm}^3 \\ 1\text{mol/L} &= 0.6 \text{ ions/nm}^3 \\ \frac{1}{\chi} = r_D &= \frac{1}{\sqrt{8\pi l_B c_0}} = \frac{1}{\sqrt{8 \times 3.14 \times 0.71 \times 0.6c_0}} = \frac{1}{3.27\sqrt{c_0}} \end{aligned} \quad (\text{A.11})$$

Substituting c_0 in mol/L gives r_D in nm.

A.3 List of symbols used in the text

We present a list of symbols in table [A.1](#).

Symbol	Quantity
T	temperature
β	$1/k_B T$
f	force
M	torque
G	Gibbs free energy per unit contour length in reference state
g	Gibbs free energy per bp ($g = 0.34G$)
g_{ij}	difference in free energy per bp between i and j phase
G^e	elastic Gibbs free energy
G^f	entropic Gibbs free energy
G^{el}	Gibbs free energy
S	Entropy
λ	stretch ($\lambda = -\partial G/\partial f$)
λ_{ext}	external displacement per unit contour length in reference state
κ	twist ($\kappa = -\partial G/\partial M$)
l_B	Bjerrum length
a_j	radius of DNA for j phase
b_j	distance between successive bp for j phase
q_j	dimensionless charge ($q_j = l_B/(b_j/2)$)
c_0	ion concentration
f_{ov}	overstretching force
s_i	$e^{-\beta G_i}$ $i = \text{B, S or P-DNA}$
$\Delta\gamma_{ij}$	interfacial energy between i and j phase
σ_{ij}	$e^{-\beta\Delta\gamma_{ij}}$, here $\Delta\gamma_{ij}$ is the interfacial energy
r_i	fraction of DNA in i^{th} phase
u	non-dimensional potential
A_j	bending modulus for j^{th} phase
S_j	stretch modulus for j^{th} phase
C_j	twist modulus for j^{th} phase
g_j	twist stretch coupling for j^{th} phase
λ_j^0	stretch of j^{th} phase at $f, M = 0$
κ_j^0	twist of j^{th} phase at $f, M = 0$
y	radial distance
L	contour length in reference configuration
r_D	Debye length

TABLE A.1: Symbols used in the text.

Appendix B

Supplement to Chapter 4

Several calculations have been omitted from the main text for the sake of brevity. They are given in this supplement in detail.

B.1 Exponential decay of interaction energy in a ‘ladder’

The calculation of interaction energies in a helical birod is considerably involved, so we first illustrate the main concepts in a simpler birod model which we call a ‘ladder’ because it is not helical. We mimic the binding of a protein by force pairs that tend to widen the ladder as shown in fig. B.1. Our goal in this section is to demonstrate the utility of the apparatus in section 2 and 3 of the main text by computing the interaction energy for two force pairs separated by a distance a as shown in fig. B.1. We work with a planar 2D birod in this section and assume small elastic deformations in the outer strands and web to keep the calculations tractable. We, ultimately, find that the interaction energy between the force pairs decays exponentially with distance a .

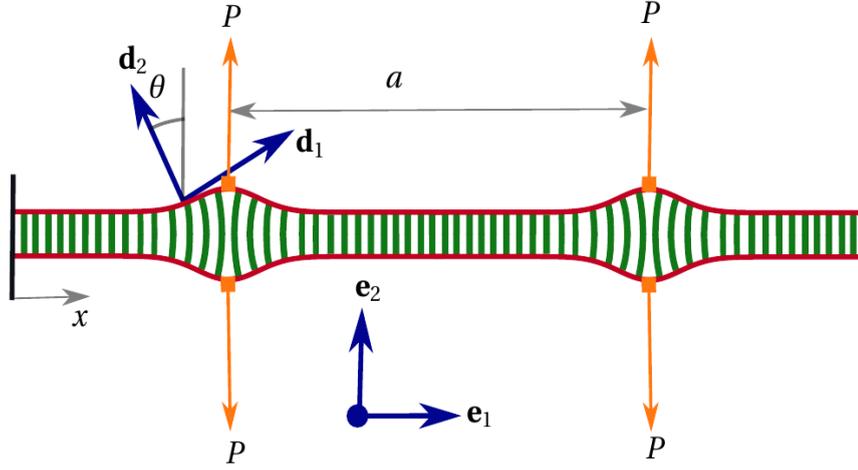


FIGURE B.1: A straight birod, referred to as a ladder, being pulled by two force pairs separated by a distance a . We show that the interaction energy between the two force pairs given by $\Delta G = E_a^2 - E_0^1 - E_a^1$ decreases exponentially with a .

B.1.1 Step 1: Kinematic description of the two strands

We use the arclength parameter x to describe the mechanics of the birod. In the reference configuration, both the strands \pm are straight, $\mathbf{r}_0^\pm = x \mathbf{e}_1 \pm \frac{d}{2} \mathbf{e}_2$, separated by distance d . Here \mathbf{e}_1 is a unit vector along the length of the birod, \mathbf{e}_2 is a unit vector perpendicular to each birod bridging the gap between them and \mathbf{e}_3 is normal to the plane of the birod as shown in fig. B.1. We begin by assuming a general displacement in $\mathbf{e}_1 - \mathbf{e}_2$ plane. For the geometry shown in fig. B.1 we expect a mirror symmetry for deformation profiles along \mathbf{e}_1 such that

$$\begin{aligned} \mathbf{r}^+ &= x \mathbf{e}_1 + \frac{d}{2} \mathbf{e}_2 + u \mathbf{e}_1 + w \mathbf{e}_2, \\ \mathbf{r}^- &= x \mathbf{e}_1 - \frac{d}{2} \mathbf{e}_2 + u \mathbf{e}_1 - w \mathbf{e}_2, \end{aligned} \quad (\text{B.1})$$

where $u = u(x)$ and $w = w(x)$ are displacements along the \mathbf{e}_1 and \mathbf{e}_2 directions, respectively.

B.1.2 Step 2: Rotation of the two strands

At each point x on the \pm strands we attach an orthogonal rotation frame which is simply $\mathbf{R}_0^\pm = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3] = \mathbf{1}_{3 \times 3}$ (the identity matrix) in the reference configuration. The vectors \mathbf{e}_1 and \mathbf{e}_2 map onto $\mathbf{d}_{1,2}^+$ and $\mathbf{d}_{1,2}^-$ in the deformed configuration for the positive

and negative strand, respectively. The \mathbf{d}_i , $i = 1, 2, 3$ are again unit vectors.

$$\begin{aligned} \mathbf{d}_1^\pm &= \cos \theta \mathbf{e}_1 \pm \sin \theta \mathbf{e}_2 \approx \mathbf{e}_1 \pm \theta \mathbf{e}_2, \\ \mathbf{d}_2^\pm &= \mp \sin \theta \mathbf{e}_1 + \cos \theta \mathbf{e}_2 \approx \pm \theta \mathbf{e}_1 + \mathbf{e}_2, \\ \mathbf{R}^\pm &= \begin{bmatrix} \cos \theta & \mp \sin \theta & 0 \\ \pm \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \approx \begin{bmatrix} 1 & \mp \theta & 0 \\ \pm \theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (\text{B.2})$$

We assume small θ to keep the calculations tractable.

B.1.3 Step 3: Extension and rotation of the web

We decompose the kinematics of the web into a macroscopic deformation and a microscopic deformation [Moakher & Maddocks \(2005\)](#). The former describes the rigid displacement and rotation, while the latter is related to the force and moment transferred by the web. The macro- displacement vector \mathbf{r} is defined as $\mathbf{r} = \frac{\mathbf{r}^+ + \mathbf{r}^-}{2} = x \mathbf{e}_1 + u \mathbf{e}_1$ [Moakher & Maddocks \(2005\)](#). The macro- rotation tensor is \mathbf{R} defined as $\mathbf{R} = (\mathbf{R}^+ \mathbf{R}^{-T})^{1/2} \mathbf{R}^-$ [Moakher & Maddocks \(2005\)](#), which in our case is

$$\mathbf{R} = (\mathbf{R}^+ \mathbf{R}^{-T})^{1/2} \mathbf{R}^- = \mathbf{I}_{3 \times 3}. \quad (\text{B.3})$$

We define another tensor \mathbf{P} relating \mathbf{R}^+ and \mathbf{R}^- to \mathbf{R} . An elastic constitutive relation discussed in further sections connects the micro- rotation tensor $\mathbf{P} = (\mathbf{R}^+ \mathbf{R}^{-T})^{1/2}$ to the moment transferred by the web.

$$\mathbf{P} = (\mathbf{R}^+ \mathbf{R}^{-T})^{1/2} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \approx \begin{bmatrix} 1 & -\theta & 0 \\ \theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (\text{B.4})$$

We need to calculate the Gibbs rotation vector $\boldsymbol{\eta} = \tan \frac{\lambda}{2} \hat{\mathbf{k}}$, where λ is obtained from $1 + 2 \cos \lambda = \text{tr}(\mathbf{P})$ and $\hat{\mathbf{k}}$ is the eigenvector of \mathbf{P} *i.e.* $\mathbf{P} \hat{\mathbf{k}} = \hat{\mathbf{k}}$. We need $\boldsymbol{\eta}$ in the subsequent section to compute the moment transferred by the web [Moakher & Maddocks \(2005\)](#). By direct observations, $\lambda = \theta$ and $\hat{\mathbf{k}} = \mathbf{e}_3$, so that $\boldsymbol{\eta} = \tan \frac{\theta}{2} \mathbf{e}_3$. The Gibbs rotation vector in the reference configuration $\boldsymbol{\eta}_0 = 0$.

The micro- displacement of the web is defined by $\mathbf{w} = \frac{\mathbf{r}^+ - \mathbf{r}^-}{2}$, which is $\mathbf{w}_0 = \frac{d}{2} \mathbf{e}_2$ in the reference configuration and $\mathbf{w} = (\frac{d}{2} + w) \mathbf{e}_2$ in the current configuration. We need \mathbf{w} and \mathbf{w}_0 to compute the force transferred by the web.

B.1.4 Step 4: Governing differential equations

We calculate various strains and curvatures associated with the deformation and relate them to the contact force and moment, respectively, which go into the governing equations. For detailed discussion on the relations used in this section we refer the reader to Moakher and Maddocks [Moakher & Maddocks \(2005\)](#). The governing equations of the birod consist of three kinetic components: the contact forces in the two strands \mathbf{n}^\pm , the contact moments \mathbf{m}^\pm , and the force \mathbf{f} and moment \mathbf{c} transferred by the $-$ strand onto the $+$ strand. We compute each of these components as follows:

1. \mathbf{n}^\pm : We need strains in the current configuration \mathbf{v}^\pm and in the reference configuration \mathbf{v}_0^\pm , in the strands to compute \mathbf{n}^\pm . These strains are:

$$\begin{aligned}\mathbf{v}_0^\pm &= \frac{\partial \mathbf{r}_0^\pm}{\partial x} = \mathbf{e}_1, \\ \mathbf{v}^\pm &= \frac{\partial \mathbf{r}^\pm}{\partial x} = (1 + u_x) \mathbf{e}_1 \pm w_x \mathbf{e}_2.\end{aligned}\tag{B.5}$$

The contact forces $\mathbf{n}^\pm = \mathbf{R}^\pm \mathbf{C} \mathbf{R}^{\pm T} \mathbf{v}^\pm$ where \mathbf{C} is a second order tensor such that $\mathbf{C}_{11} = EA$, $\mathbf{C}_{22} = GA$ and $\mathbf{C}_{12} = \mathbf{C}_{21} = 0$. Here E is the stretch modulus, G shear modulus and A is the cross-sectional area of the strands. Upon performing the calculation and taking account of the fact that u, w and θ are small and upon ignoring higher order terms we get,

$$\mathbf{n}^\pm = EA u_x \mathbf{e}_1 \pm GA(w_x - \theta) \mathbf{e}_2.\tag{B.6}$$

2. \mathbf{m}^\pm : For calculating the contact moments \mathbf{m}^\pm in the respective strands we need the curvature vector $\boldsymbol{\kappa}^\pm$ for the two strands, which can, in turn, be obtained by computing the axial vector of the skew-symmetric matrices $\mathbf{U}^\pm = \frac{\partial \mathbf{R}^\pm}{\partial x} \mathbf{R}^{\pm T}$.

$$\begin{aligned}\mathbf{U}^\pm &= \frac{\partial \mathbf{R}^\pm}{\partial x} \mathbf{R}^{\pm T} = \begin{bmatrix} 0 & \mp \theta_x & 0 \\ \pm \theta_x & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \boldsymbol{\kappa}^\pm &= \pm \theta_x \mathbf{e}_3.\end{aligned}\tag{B.7}$$

The contact moment \mathbf{m}^\pm is related to the curvature via a bending rigidity EI such that

$$\mathbf{m}^\pm = \pm EI \theta_x \mathbf{e}_3.\tag{B.8}$$

Here, I is the moment of inertia of the cross-section of the outer strands.

3. \mathbf{f} and \mathbf{c} : The force transferred by the web \mathbf{f} is proportional to the change in the dimensions of the web quantified by \mathbf{w} and \mathbf{w}_0 in the previous sections such that,

$$2\mathbf{f} = \mathbf{RHR}^T[\mathbf{w} - \mathbf{R}\dot{\mathbf{w}}_0] \approx Lw \mathbf{e}_2, \quad (\text{B.9})$$

where \mathbf{H} is a diagonal second order elasticity tensor such that $\mathbf{H}_{22} = L$. Similarly, the moment transferred by the web \mathbf{c} is elastically related to $\boldsymbol{\eta}$ and $\boldsymbol{\eta}_0$ calculated in the previous sections:

$$2\mathbf{c} = \frac{1}{\alpha} \mathbf{RGR}^T(\boldsymbol{\eta} - \mathbf{R}\dot{\boldsymbol{\eta}}) - \boldsymbol{\eta} \times (\mathbf{w} \times \mathbf{f}) \approx K\theta \mathbf{e}_3, \quad (\text{B.10})$$

where $\alpha = \frac{2}{1+\|\boldsymbol{\eta}\|^2}$, \mathbf{G} is a second order diagonal elasticity tensor and $K = \frac{G_{33}}{2}$.

The governing equations from box 4 in [Moakher & Maddocks \(2005\)](#) are given by,

$$\begin{aligned} \mathbf{n}_x &= 0, \\ \mathbf{m}_x + \mathbf{r}_x \times \mathbf{n} &= 0. \end{aligned} \quad (\text{B.11a})$$

$$\begin{aligned} \mathbf{n}_x^c - 2\mathbf{f} &= 0, \\ \mathbf{m}_x^c + \mathbf{r}_x \times \mathbf{n}^c - \mathbf{c} &= 0. \end{aligned} \quad (\text{B.11b})$$

In the above equations, $\mathbf{n} = \mathbf{n}^+ + \mathbf{n}^- = 2EAu_x \mathbf{e}_1$, $\mathbf{n}^c = \mathbf{n}^+ - \mathbf{n}^- = 2GA(w_x - \theta) \mathbf{e}_2$, $\mathbf{m} = \mathbf{m}^+ + \mathbf{m}^- + \mathbf{w} \times \mathbf{n}^c = 0$ and $\mathbf{m}^c = \mathbf{m}^+ - \mathbf{m}^- + \mathbf{w} \times \mathbf{n} = 2EI\theta_x \mathbf{e}_3 + (d/2 + w) \mathbf{e}_2 \times 2EAu_x \mathbf{e}_1 \approx 2[EI\theta_x - \frac{d}{2}EAu_x] \mathbf{e}_3$. Upon substituting these values into the governing equations we get,

$$\begin{aligned} EAu_{xx} &= 0, \\ 2GA(w_{xx} - \theta_x) - Lw &= 0, \\ 2(EI\theta_{xx} - d/2EAu_{xx}) + (1 + u_x) \mathbf{e}_1 \times 2GA(w_x - \theta) \mathbf{e}_2 - K\theta \mathbf{e}_3 &= 0. \end{aligned} \quad (\text{B.12})$$

We use $\theta_x = w_{xx} - \frac{L}{2GA}w$ and $u_{xx} = 0$ and get,

$$EIw_{xxxx} - \left(\frac{EIL}{2GA} + \frac{K}{2}\right)w_{xx} + \left(\frac{L}{2} + \frac{KL}{4GA}\right)w = 0. \quad (\text{B.13})$$

If we further assume that the outer strands are unshearable ($GA \rightarrow \infty$ and $\theta = w_x$), the above equation reduces to a simpler equation.

$$EIw_{xxxx} - \frac{K}{2}w_{xx} + \frac{L}{2}w = 0. \quad (\text{B.14})$$

B.1.5 Step 5,6 and 7: Interaction Energy

We substitute $w = e^{ms}$, and get eigenvalues $m = \pm\lambda, \pm\mu$. For illustration purposes, we assume λ and μ are real numbers (i.e., $K^2 - 32L > 0$) and the ladder extends from $-\infty$ in the negative \mathbf{e}_1 direction to $+\infty$ in the positive \mathbf{e}_1 direction with $w = w_x = 0$ at $x = \pm\infty$. Hence, for a force pair at $x = 0$

$$\begin{aligned} w(x) &= Ae^{\lambda x} + Be^{\mu x} \quad \text{when } x < 0, \\ w(x) &= Ae^{-\lambda x} + Be^{-\mu x} \quad \text{when } x > 0, \end{aligned} \tag{B.15}$$

for some constants A and B which could be determined using boundary conditions in step 5. For two force pairs separated by a distance a , the displacement profile $w_2(x) = w(x) + w(x - a)$. The elastic energy in the deformed configuration is computed in step 6 and is given by,

$$E[w] = EIw_{xx}^2 + \frac{1}{2}Kw_x^2 + \frac{1}{2}Lw^2. \tag{B.16}$$

Finally, we compute the interaction energy defined by $\Delta G = E[w_2] - 2E[w]$ in step 7 and find that it decreases exponentially with the distance a .

$$\begin{aligned} \Delta G &= \frac{L}{2} \left(\frac{e^{-\lambda a} (A^2\lambda^2\mu - A^2\mu^3 + A^2\lambda^3\mu a - A^2\lambda\mu^3 a - 4AB\lambda\mu^2)}{\lambda\mu(\lambda^2 - \mu^2)} + \right. \\ &\quad \left. \frac{e^{-\mu a} (4AB\lambda^2\mu + B^2\lambda^3 - B^2\lambda\mu^2 + B^2\lambda^3\mu a - B^2\lambda\mu^3 a)}{\lambda\mu(\lambda^2 - \mu^2)} \right) + \\ &\quad \frac{K}{2} \left(\frac{e^{-\lambda a} (A^2\lambda^3 - A^2\lambda\mu^2 - A^2\lambda^4 a + A^2\lambda^2\mu^2 a + 4AB\lambda^2\mu)}{(\lambda^2 - \mu^2)} + \right. \\ &\quad \left. \frac{e^{-\mu a} (-4AB\lambda\mu^2 + B^2\lambda^2\mu - B^2\mu^3 - B^2\lambda^2\mu^2 a + B^2\mu^4 a)}{(\lambda^2 - \mu^2)} \right) + \\ &\quad EI \left(\frac{e^{-\lambda a} (A^2\lambda^5 - A^2\lambda^3\mu^2 + A^2\lambda^6 a - A^2\lambda^4\mu^2 a - 4AB\lambda^2\mu^3)}{(\lambda^2 - \mu^2)} + \right. \\ &\quad \left. \frac{e^{-\mu a} (4AB\lambda^3\mu^2 + B^2\lambda^2\mu^3 - B^2\mu^5 + B^2\lambda^2\mu^4 a - B^2\mu^6 a)}{(\lambda^2 - \mu^2)} \right). \end{aligned} \tag{B.17}$$

We follow these steps for a helical birod model of DNA in the main text.

B.2 Kinematics of the $-$ strand

In the main text we gave detailed derivations for the strains, curvatures, etc., for the $+$ strand in our birod. We now shift our attention to the complimentary $-$ strand. The

reference configuration of this strand is denoted by position vector \mathbf{r}_0^- .

$$\mathbf{r}_0^- = b(\cos(\omega x + \alpha) \mathbf{e}_1 + \sin(\omega x + \alpha) \mathbf{e}_2) + x \mathbf{e}_3. \quad (\text{B.18})$$

Along the same lines as the + strand, we conceive the deformed configuration to be a helix wrapped around a curved axis defined by curvatures k_1, k_2 and k_3 along the directors $\mathbf{d}_1, \mathbf{d}_2$ and \mathbf{d}_3 , respectively.

$$\mathbf{r}^-(x) = (b + r^-)(\cos(\omega x + \alpha + \beta^-) \mathbf{d}_1 + \sin(\omega x + \alpha + \beta^-) \mathbf{d}_2) + \int_0^x (1 + b\xi) \mathbf{d}_3 dx. \quad (\text{B.19})$$

We use the same apparatus *mutatis mutandis* described for the + strand to calculate various quantities of interest. The results are:

$$\mathbf{R}^- = [\mathbf{n}^- \quad \mathbf{b}^- \quad \mathbf{t}^-] = \mathbf{Z}\mathbf{R}_0^-(\mathbf{1} + \mathbf{\Theta}^-). \quad (\text{B.20})$$

where $\mathbf{\Theta}^-$ is a skew symmetric tensor.

$$\mathbf{\Theta}^- = \begin{bmatrix} 0 & -\theta_3^- & \theta_2^- \\ \theta_3^- & 0 & -\theta_1^- \\ -\theta_2^- & \theta_1^- & 0 \end{bmatrix}, \quad (\text{B.21})$$

$$\text{where } \theta_1^- = (r^- \omega + b(\beta_x^- + k_3)), \quad \theta_2^- = -r_x^- \cos k + \beta^- \sin k, \\ \theta_3^- = \frac{g^-}{\omega \sin k} - \frac{(r_x^- \cos k - \beta^- \sin k) \cos k}{\omega \sin k}.$$

We compute curvature κ^- as follows,

$$\Omega^- = (\mathbf{t}_x^- \cdot \mathbf{t}_x^-)^{1/2} = \omega \sin k - (r_{xx}^- + \xi) \cos k + (\beta_x^- + k_3) \sin k, \\ \kappa^- = \Omega^- - \omega \sin k = -(r_{xx}^- + \xi) \cos k + (\beta_x^- + k_3) \sin k. \quad (\text{B.22})$$

We obtain the moment \mathbf{m}^- as follows,

$$\mathbf{m}^- = EI\kappa^- (\cos k \cos \frac{\alpha}{2} \mathbf{f}_1 + \cos k \sin \frac{\alpha}{2} \mathbf{f}_2 + \sin k \mathbf{f}_3), \quad (\text{B.23})$$

where $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ are given as follows.

$$\mathbf{f}_1 = \left(\sin(\omega x + \frac{\alpha}{2}) \mathbf{d}_1 - \cos(\omega x + \frac{\alpha}{2}) \mathbf{d}_2 \right), \quad \mathbf{f}_2 = \left(\cos(\omega x + \frac{\alpha}{2}) \mathbf{d}_1 + \sin(\omega x + \frac{\alpha}{2}) \mathbf{d}_2 \right), \quad \mathbf{f}_3 = \mathbf{d}_3. \quad (\text{B.24})$$

B.3 Evaluation of material properties of the web

In this section, we consider a deformation of the double-helical structure induced by a stretching force F and torque T on one end. We assume that the helix retains its helical configuration, but with changed geometrical parameters. Thus, r , β and e are independent of x . Our goal is to compute the strains and curvatures, then evaluate the energy, and then identify the stretch modulus, twist modulus and twist-stretch coupling modulus of the double-helical structure from this energy expression. The computation of strains, curvatures, etc., of the helix proceeds as in the main text.

$$\begin{aligned}\mathbf{r}^+ &= (a+r)(\cos \omega x(1+\beta) \mathbf{e}_1 + \sin \omega x(1+\beta) \mathbf{e}_2) + x(1+e), \\ \mathbf{r}_x^+ &= -(a+r)(\cos \omega x(1+\beta) \mathbf{e}_1 + \sin \omega x(1+\beta) \mathbf{e}_2) + x(1+e),\end{aligned}\tag{B.25}$$

We assume $r, \beta, e \sim O(\varepsilon)$, hence

$$\begin{aligned}\mathbf{r}^+ &= (a+r)(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) + a\omega\beta x(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + x(1+e) \mathbf{e}_3, \\ \mathbf{r}_x^+ &= (a+r)\omega(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + a\omega\beta(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) - a\omega^2, \\ &\quad \beta x(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) + (1+(ex)_x) \mathbf{e}_3, \\ &= -a\omega^2\beta x(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) + \omega(a+r+a\beta)(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + (1+(ex)_x) \mathbf{e}_3.\end{aligned}\tag{B.26}$$

The inextensibility condition gives,

$$\begin{aligned}|\mathbf{r}_x^+| &= |\mathbf{r}_{0x}^+|, \\ (ex)_x + \omega^2 a(r+\beta) &= 0,\end{aligned}\tag{B.27}$$

\mathbf{t}_0^+ , \mathbf{n}_0^+ and \mathbf{b}_0^+ are the tangent, normal and binormal to the + strand in the reference configuration. We calculate tangent \mathbf{t}^+ to the deformed configuration.

$$\begin{aligned}\mathbf{t}^+ &= -\sin k\beta x(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) + (\sin k + \omega r \cos k + \beta \sin k) \\ &\quad (-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + (\cos k - \omega \sin k(r+a\beta)) \mathbf{e}_3, \\ &= \mathbf{t}_0^+ + \omega\beta x \sin k \mathbf{n}_0^+ + (\omega r + \beta \tan k) \mathbf{b}_0^+,\end{aligned}\tag{B.28}$$

Next, we calculate the curvature κ^+ .

$$\begin{aligned}\mathbf{t}_x^+ &= -(\omega \sin k 2\omega\beta \sin k + \omega^2 r \cos k)(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) \\ &\quad - \omega^2 \sin k\beta x(-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2). \\ K^2 &= \omega \sin k + 2\omega\beta \sin k + \omega^2 r \cos k. \\ \kappa^+ &= K - \omega \sin k = 2\omega\beta \sin k + \omega^2 r \cos k.\end{aligned}\tag{B.29}$$

We go on to calculate the normal in the deformed configuration \mathbf{n}^+ .

$$\begin{aligned}\mathbf{n}^+ &= -(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) - \omega \beta x (-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2), \\ &= \mathbf{n}_0^+ - \omega \beta x \sin k \mathbf{t}_0^+ + \omega \beta x \cos k \mathbf{b}_0^+.\end{aligned}\quad (\text{B.30})$$

We are now in a position to calculate the deformed Frenet-Serret frame \mathbf{R}^+ .

$$\mathbf{R}^+ = [\mathbf{n}^+ \quad \mathbf{b}^+ \quad \mathbf{t}^+] = \mathbf{R}_0^+ (\mathbf{1} + \mathbf{\Theta}^+). \quad (\text{B.31})$$

where $\mathbf{\Theta}^+$ is a skew symmetric tensor.

$$\mathbf{\Theta}^+ = \begin{bmatrix} 0 & -\theta_3^+ & \theta_2^+ \\ \theta_3^+ & 0 & -\theta_1^+ \\ -\theta_2^+ & \theta_1^+ & 0 \end{bmatrix}, \quad (\text{B.32})$$

$$\text{where } \theta_1^+ = \omega r + \beta \tan k, \quad \theta_2^+ = \omega \beta x \sin k, \quad \theta_3^+ = \omega \beta x \cos k.$$

For the negative strand we follow the same procedure.

$$\begin{aligned}\mathbf{R}^- &= [\mathbf{n}^- \quad \mathbf{b}^- \quad \mathbf{t}^-] = \mathbf{R}_0^- (\mathbf{1} + \mathbf{\Theta}^-), \\ \mathbf{\Theta}^- &= \mathbf{\Theta}^+, \\ \kappa^- &= \kappa^+.\end{aligned}\quad (\text{B.33})$$

After performing all the calculations

$$\begin{aligned}E &= \int_0^L (EI(2\omega\beta \sin k + \omega^2 r \cos k)^2 + \frac{1}{2}H_1\omega^2(r + a\beta)^2 + \frac{1}{2}L_1r^2) - M\theta - F\Delta x, \\ \Delta x &= eL, \quad \theta = \beta L.\end{aligned}\quad (\text{B.34})$$

We substitute $r = -\frac{e}{\omega^2 a} - a\beta$ from eqn. (B.27) and compute the elastic constants as follows.

$$\begin{aligned}\frac{\partial E}{\partial \beta} &= 0, \quad \frac{\partial E}{\partial e} = 0, \\ S &= \frac{\partial^2 E}{\partial e^2}, \quad g = \frac{\partial^2 E}{\partial e \partial \beta}, \quad C = \frac{\partial^2 E}{\partial \beta^2}.\end{aligned}\quad (\text{B.35})$$

Then, by trial and error we pick values of $L_1, L_2, L_3, H_1, H_2, H_3, K_c, K_e, EI$ to match the S, g, C known from experiments. Our choice of the material parameters L_1, H_2, K_c , etc., is not unique.

B.4 Choice of eigenvalues obtained in section 5

In section 5, we solve the governing differential equation eqn. 3.33 by substituting $y(x) = y_0 e^{-\lambda x}$ where $y = (r, f, \xi, k_3, \beta^\pm, n_i^c, n_i)$ $i = 1, 2, 3$. We look for the values of λ corresponding to a non-trivial solution of the governing equations. For this we need to solve the eigenvalue problem $\mathcal{A}(\lambda)\mathbf{v}_0 = 0$, where \mathcal{A} is a function of λ and elastic constants (eqn. 4.1) and $\mathbf{v}_0 = [r_0, f_0, \xi_0, k_{30}, \beta_0^+, \beta_0^-, n_{i0}^c, n_{i0}]^T$ $i = 1, 2, 3$. We set $\det \mathcal{A}(\lambda) = 0$ and get following solutions for λ .

$$\begin{aligned}
 x_1 &= -1.5 \times 10^4(1 + i), & x_2 &= -1.5 \times 10^4(-1 + i), & x_3 &= -4 \times 10^3, \\
 x_4 &= 1.2 \times 10^3(-1 - 3.2i), & x_5 &= 1.2 \times 10^3(-1 + 3.2i), & x_6 &= -0.68, \\
 x_7 &= -0.42, & x_8 &= -0.36, & x_9 &= -5.2 \times 10^{-10}, \\
 x_{10} &= -1.9i, & x_{11} &= 1.9i, & x_{12} &= -3.8i, & x_{13} &= 3.8i, & x_{14} &= -6.2i, & x_{15} &= 6.2i, \\
 x_{16} &= 5.2 \times 10^{-10}, & x_{17} &= 0.36, & x_{18} &= 0.42, & x_{19} &= 0.68, & x_{20} &= 2.3 \times 10^3(1.4 - i), \\
 x_{21} &= 2.3 \times 10^3(1.4 + i), & x_{22} &= 1.5 \times 10^4(1 - i), & x_{23} &= 1.5 \times 10^4(1 + i).
 \end{aligned}
 \tag{B.36}$$

Among these 23 eigenvalues we neglect the eigenvalues $x_{1,2,3,4,5,20,21,22,23}$ whose magnitude is $> 10^3$ because the corresponding decay length is tiny which leads to large numerical errors given that we need to compute third derivatives. Then, there are small eigenvalues $x_{9,16}$ whose magnitude is close to zero ($< 10^{-3}$) and purely imaginary eigenvalues $x_{10,11,12,13,14,15}$ which when substituted in $e^{-\lambda x}$ result in a constant or a sinusoidal function, respectively, that do not decay to 0 as $x \rightarrow \pm\infty$. Hence, we must neglect these too. This leaves us with $x_{6,7,8,17,18,19}$, which are used in section 5.

B.5 Results for $\alpha = \pi$ radians

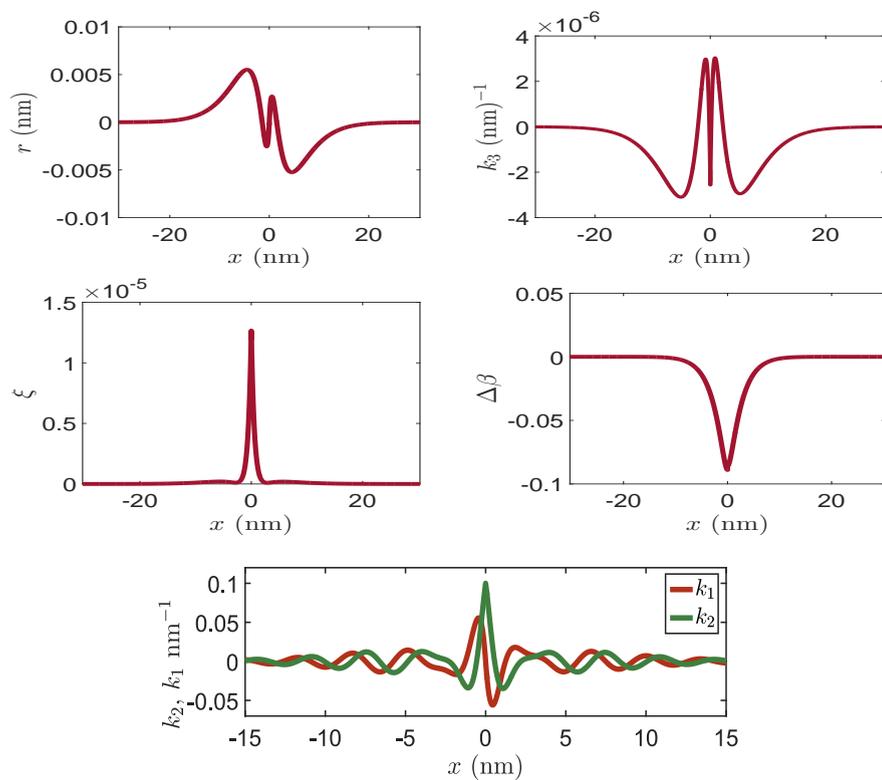


FIGURE B.2: Variation of strain variables for $\alpha = \pi$ radians. Notice that the curves are symmetric about the site of protein binding. As mentioned in section 6, the curves are not symmetric if we choose $\alpha = 2.1$ radians.

Appendix C

Dependence of interaction energy on boundary conditions and elastic constants.

The elastic constants and the boundary conditions used in chapter 3 and chapter 4 have been chosen in a way that we get sinusoidal dependence of the interaction energy (shown in fig. 3.6 (a) and fig. 4.5(b)) while recovering the stretch, twist and twist-stretch coupling moduli of dsDNA (shown in Appendix B). The aim in this appendix is to examine the robustness of the sinusoidal behavior to the choice of these constants and boundary conditions, and to shed slight on the approach to fit the data. Succintly, the amplitude of the interaction energy depends only the amount of the distortion caused by the protein at the binding site encoded in the magnitude of the boundary conditions (see fig. 3.8) while the decay length of the exponential depends only on the elastic constants (see fig. 4.7). Let us examine them in more detail:

- **Elastic constants:** In chapter 3, we choose the elastic constants in a way such that the eigenvalue λ obtained in eqn. 3.14 leads to the correct behavior observed in simulations and experiments. Here we examine the effect of the choice of the elastic constants on the real and imaginary parts of the eigenvalue λ . Let

$$\lambda = -\frac{1}{l_d} + i\nu \tag{C.1}$$

We parametrize the elasticity of the web by a single parameter A such that $L_1 = L_2 = L_3 = A pN/nm^2$ and $H_1 = H_2 = H_3 = A pN.$, and examine the effect of A on decay length l_d and period ν . The results are given in fig.C.1. We find that the decay length decreases as the elasticity of the web increases which is

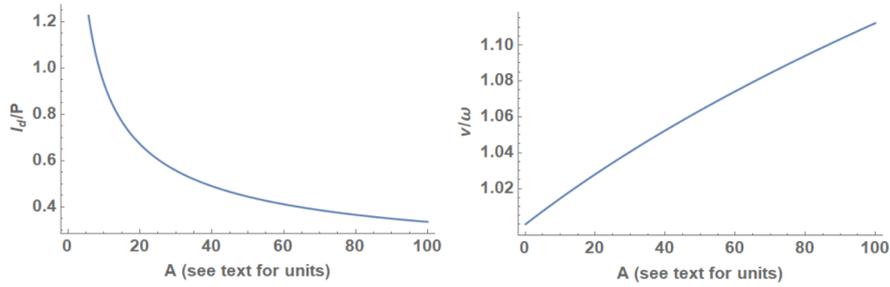


FIGURE C.1: Effect of elastic constants on the decay length l_d (nm) and period ν . ω and P are the period and pitch, respectively for dsDNA. The elasticity of the web is parametrized by a single parameter A such that $L_1 = L_2 = L_3 = A \text{ pN/nm}^2$ and $H_1 = H_2 = H_3 = A \text{ pN}$. We choose $A = 10$ since it gives decay length in the right ballpark [Kim et al. \(2013\)](#).

what is expected since the base-pairs with higher GC content are known to have lower decay lengths (for more details see fig. 4.7). Also, change in the period ν is minimal even after a two order of magnitude increase in the parameter A . We choose $A \approx 10$ since it gives correct decay length.

- **Boundary conditions:** As pointed out in the chapters 3 and 4, the mechanical distortions caused by the protein at the binding site are not well understood in literature. Within the scope of the birod model discussed here, the protein applies a set of boundary conditions. The number of the boundary conditions required to solve the problem depend on the number of unknown constants in the solution to the governing differential equations (See eqn. 3.16 and eqn. 4.37).

In chapter 3, we imposed boundary conditions on the change in radius (r) and change in phase-angle (β). We could have imposed boundary conditions on any two of the strain parameters (r, β, k_3, ξ). The magnitude of the strain parameters at the binding site (r_0 and β_0) determines the amplitude of the interaction energy.

In chapter 4, we imposed boundary conditions on the change in radius r and the curvatures k_1 and k_2 . We note here that if the magnitude of the curvatures at the binding site is zero $k_{10} = k_{20} = 0$, the interaction energy lacks the sinusoidal character (see fig. 4.5(a)). Here also, we deliberately choose the boundary conditions which give us the sinusoidal character to match the experimental data.

Appendix D

Supplement to Chapter 5

We used the Metropolis algorithm [Pathria \(1984\)](#) to perform the MC simulations in the text. Here is a stepwise procedure for the algorithm:

1. The discretized birod is kinematically described by a vector $\mathbf{y} = [\boldsymbol{\theta}, \mathbf{v}]$, where $\boldsymbol{\theta}$ describes the macroscopic deformation of the centerline and \mathbf{v} is the inter-strands distance. Randomly initialize the vector \mathbf{y} . Let call this configuration C_1 .
2. Randomly perturb the configuration C_1 to a nearby configuration C_2 .
3. Compute the difference in the energy of the two states $\Delta E = E(C_2) - E(C_1)$.
4. If $\Delta E < 0$, then accept the new configuration.
5. If $\Delta E > 0$, then choose a random number $r \in (0, 1)$. Accept the new configuration if $r < e^{-\Delta E/k_B T}$. Reject otherwise.

This code for this algorithm is given below (Due to formatting issues the indentation might be wrong.):

```
import numpy as np
import random

class birod():
    # BIROD SYSTEM INITIALIZATION
    def __init__(self, EI, L1, L2, H, F, a, lam, Lcont, kB-T):
        self.EI = EI
        self.L1 = L1
        self.L2 = L2
        self.H = H
        self.F = F
        self.a = a
```

```

        self.lam = lam
        self.Lcont = Lcont
        self.kB_T = kB_T

    def compute_mean(self, n):

        #DISCRETIZATION OF BIROD
        l = self.Lcont/n

        #ENERGY FUNCTIONAL (Wrong Indentation)
        ener = lambda yvec: \
            (self.EI/l)*np.sum(np.square(yvec[1:n]-yvec[:n-1])) \
            + (4.0*self.L1*self.a**2.0*1)* np.sum(np.square(yvec[:n])) \
            + (self.EI/l**3.0)\
            * np.sum(np.square(yvec[n+2:]-2*yvec[n+1:-1]+yvec[n:-2])) \
            + (self.L2*1)*np.sum(np.square(np.exp(-self.lam*yvec[n:])-1)) \
            + (self.H/l)*np.sum(np.square(yvec[n+1:]-yvec[n:-1])) \

        # DOF LIST AVAILABLE FOR PERTURBATION
        # 0:n ARE THETA'S AND n:2n ARE V'S.
        ll = [i for i in range(1,n)] + [i + n for i in range(1,n)]

        # INITIALIZATION
        yvec = np.zeros((2*n,))

        # NUMBER OF MC ITERATIONS
        niter = 1000000

        # REGISTERS FOR VARIOUS QUANTITIES
        Ereg = np.zeros((niter,))
        qtyreg = np.zeros((niter,))
        Ereg[0] = ener(yvec)
        accep_states = 0

        # Seeding the random process, helps debug the code
        np.random.seed(20)
        random.seed(20)

        for i in range(1,niter):

            e1 = Ereg[i-1]

        #CHOOSE THE DOF TO PERTURB
            j = random.choice(ll)

        # STATE PETURBATION.
        # CHOOSE CAREFULLY. YOU SHOULD REALISTICALLY ACCEPT 20-40%
        # OF YOUR STATES. YOU COULD IN PRINCIPAL USE ANY PERTURBATION

```

```

# AND THE CODE WILL CONVERGE eventually. TO PERTURB YOU COULD
# USE ANY DISTRIBUTION SYMMETRIC ABOUT ZERO. NORMAL N(0,1)
# WORKS WHILE UNIFORM DISTRIBUTION U(0,1) DOES NOT SINCE
# IT NOT SYMMETRIC. AS LONG YOU ARE NOT BIASING, YOU
# SHOULD BE OKAY HERE.
    dely = np.random.normal(0,1.5)
    yvec[j] += dely
    e2 = ener(yvec)

# COMPUTING \Delta E
    dele = e2 - e1
    if dele > 0:
        r1 = np.random.uniform(0,1)
        if r1 < np.exp(-dele/self.kB.T):
# ACCEPTED STATES FOR \Delta E > 0
            Ereg[i] = e2
            accep_states += 1
        else:
# REJECTED STATES FOR \Delta E > 0
            yvec[j] -= dely
            Ereg[i] = Ereg[i-1]
    else:
# ACCEPTED STATES FOR \Delta E < 0
        Ereg[i] = e2
        accep_states += 1

# REGISTER FOR QTY OF INTEREST.
# IN OUR CASE THE INTERSTRAND DISTANCE
    qtyreg[i] = np.mean(yvec[n:])

# COMPUTE THE AVERAGE OVER LAST N/2 STATES
# THEREBY MAKING SURE THAT THE PROCESS HAS STABILIZED
return (2.0/(niter-1))*np.sum(qtyreg[int(niter/2):]), accep_states

```

LISTING D.1: Code for MC simulations

References

- Abramowitz, M., & Stegun, I. A. (1948). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Vol. 55). US Government printing office.
- Allemand, J. F., Bensimon, D., Lavery, R., & Croquette, V. (1998). Stretched and overwound dna forms a pauling-like structure with exposed bases. *Proceedings of the National Academy of Sciences*, *95*(24), 14152–14157.
- Argudo, D., & Purohit, P. K. (2014a). Equilibrium and kinetics of dna overstretching modeled with a quartic energy landscape. *Biophysical journal*, *107*(9), 2151–2163.
- Argudo, D., & Purohit, P. K. (2014b). Torsion of dna modeled as a heterogeneous fluctuating rod. *Journal of the Mechanics and Physics of Solids*, *62*, 228–256.
- Blake, R., & Delcourt, S. G. (1998). Thermal stability of dna. *Nucleic acids research*, *26*(14), 3323–3332.
- Bryant, Z., Stone, M. D., Gore, J., Smith, S. B., Cozzarelli, N. R., & Bustamante, C. (2003). Structural transitions and elasticity from torque measurements on dna. *Nature*, *424*(6946), 338–341.
- Crothers, D. M. (1968). Calculation of melting curves for dna. *Biopolymers: Original Research on Biomolecules*, *6*(10), 1391–1404.
- Dauxois, T., Peyrard, M., & Bishop, A. R. (1993). Entropy-driven dna denaturation. *Physical Review E*, *47*(1), R44.
- de Gennes, P.-G. (2001). Maximum pull out force on dna hybrids. *arXiv preprint physics/0110011*.
- Drsata, T., Zgarbova, M., Jurecka, P., Sponer, J., & Lankas, F. (2016). On the use of molecular dynamics simulations for probing allostery through dna. *Biophysical journal*, *110*(4), 874–876.
- Drsata, T., Zgarbova, M., Spackova, N., Jurecka, P., Sponer, J., & Lankas, F. (2014). Mechanical model of dna allostery. *The journal of physical chemistry letters*, *5*(21), 3831–3835.

- Eslami-Mossallam, B., Schiessel, H., & van Noort, J. (2016). Nucleosome dynamics: Sequence matters. *Advances in colloid and interface science*, *232*, 101–113.
- Frank-Kamenetskiĭ. (n.d.). Polyelectrolyte model of dna.
- Fu, H., Chen, H., Zhang, X., Qu, Y., Marko, J. F., & Yan, J. (2011). Transition dynamics and selection of the distinct s-dna and strand unpeeling modes of double helix overstretching. *Nucleic acids research*, *39*(8), 3473–3481.
- Gibbs-Davis, J. M., Schatz, G. C., & Nguyen, S. T. (2007). Sharp melting transitions in dna hybrids without aggregate dissolution: Proof of neighboring-duplex cooperativity. *Journal of the American Chemical Society*, *129*(50), 15535–15540.
- Gore, J., Bryant, Z., Nöllmann, M., Le, M. U., Cozzarelli, N. R., & Bustamante, C. (2006). Dna overwinds when stretched. *Nature*, *442*(7104), 836–839.
- Grosberg, A. Y., Khokhlov, A. R., Stanley, H. E., Mallinckrodt, A. J., & McKay, S. (1995). Statistical physics of macromolecules. *Computers in Physics*, *9*(2), 171–172.
- Gu, C., Zhang, J., Yang, Y. I., Chen, X., Ge, H., Sun, Y., ... Gao, Y. Q. (2015). Dna structural correlation in short and long ranges. *The Journal of Physical Chemistry B*, *119*(44), 13980–13990.
- Gurtin, M. E. (1999). *Configurational forces as basic concepts of continuum physics* (Vol. 137). Springer Science & Business Media.
- Hancock, S. P., Ghane, T., Cascio, D., Rohs, R., Di Felice, R., & Johnson, R. C. (2013). Control of dna minor groove width and fis protein binding by the purine 2-amino group. *Nucleic acids research*, *41*(13), 6750–6760.
- Kamien, R. D., Lubensky, T., Nelson, P., & O’Hern, C. S. (1996). Twist-stretch elasticity of dna. *MRS Online Proceedings Library Archive*, *463*.
- Kim, S., Broströmer, E., Xing, D., Jin, J., Chong, S., Ge, H., ... others (2013). Probing allostery through dna. *Science*, *339*(6121), 816–819.
- King, G. A., Gross, P., Bockelmann, U., Modesti, M., Wuite, G. J., & Peterman, E. J. (2013). Revealing the competition between peeled ssdna, melting bubbles, and s-dna during dna overstretching using fluorescence microscopy. *Proceedings of the National Academy of Sciences*, *110*(10), 3859–3864.
- King, G. A., Peterman, E. J., & Wuite, G. J. (2016). Unravelling the structural plasticity of stretched dna under torsional constraint. *Nature communications*, *7*(1), 1–7.

- Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985). The molecular origin of dna-drug specificity in netropsin and distamycin. *Proceedings of the National Academy of Sciences*, *82*(5), 1376–1380.
- Koslover, E. F., & Spakowitz, A. J. (2009). Twist-and tension-mediated elastic coupling between dna-binding proteins. *Physical review letters*, *102*(17), 178102.
- Krugh, T. R., & Young, M. A. (1977). Daunorubicin and adriamycin facilitate actinomycin d binding to poly (da–dt)· poly (da–dt). *Nature*, *269*(5629), 627–628.
- Kwieceński, J., Chapman, S. J., & Goriely, A. (2017). Self-assembly of a filament by curvature-inducing proteins. *Physica D: Nonlinear Phenomena*, *344*, 68–80.
- Landau, L., & Lifshitz, E. (1969). Statistical physics, 1969. *Google Scholar*, 237–241.
- Lankaš, F., Gonzalez, O., Heffler, L., Stoll, G., Moakher, M., & Maddocks, J. H. (2009). On the parameterization of rigid base and basepair models of dna from molecular dynamics simulations. *Physical Chemistry Chemical Physics*, *11*(45), 10565–10588.
- Lankas, F., Sponer, J., Hobza, P., & Langowski, J. (2000). Sequence-dependent elastic properties of dna. *Journal of molecular biology*, *299*(3), 695–709.
- Leger, J., Romano, G., Sarkar, A., Robert, J., Bourdieu, L., Chatenay, D., & Marko, J. (1999). Structural transitions of a twisted and stretched dna molecule. *Physical review letters*, *83*(5), 1066.
- Lehman, G., & McTague, J. (1968). Melting of dna. *The Journal of Chemical Physics*, *49*(7), 3170–3179.
- Liang, X., & Purohit, P. K. (2018a). A method to compute elastic and entropic interactions of membrane inclusions. *Extreme mechanics letters*, *18*, 29–35.
- Liang, X., & Purohit, P. K. (2018b). A method to compute elastic and entropic interactions of membrane inclusions. *Extreme mechanics letters*, *18*, 29–35.
- Manning, R. S., Maddocks, J. H., & Kahn, J. D. (1996). A continuum rod model of sequence-dependent dna structure. *The Journal of chemical physics*, *105*(13), 5626–5646.
- Marko, J., & Siggia, E. (1995). Statistical mechanics of supercoiled dna. *Physical Review E*, *52*(3), 2912.
- Marko, J. F. (2007). Torque and dynamics of linking number relaxation in stretched supercoiled dna. *Physical Review E*, *76*(2), 021926.

- Marko, J. F., & Neukirch, S. (2013). Global force-torque phase diagram for the dna double helix: structural transitions, triple points, and collapsed plectonemes. *Physical Review E*, 88(6), 062722.
- Marko, J. F., & Siggia, E. D. (1995). Stretching dna. *Macromolecules*, 28(26), 8759–8770.
- Mitra, A., & Sept, D. (2008). Taxol allosterically alters the dynamics of the tubulin dimer and increases the flexibility of microtubules. *Biophysical journal*, 95(7), 3252–3258.
- Moakher, M., & Maddocks, J. (2005). Double-strand elastic rod theory. *Archive for rational mechanics and analysis*, 177, 53-91.
- Nelson, P. (2004). *Biological physics*. WH Freeman New York.
- Ngo, K. X., Umeki, N., Kijima, S. T., Kodera, N., Ueno, H., Furutani-Umezu, N., . . . others (2016). Allosteric regulation by cooperative conformational changes of actin filaments drives mutually exclusive binding with cofilin and myosin. *Scientific reports*, 6, 35449.
- Nishigaki, K., Husimi, Y., Masuda, M., Kaneko, K., & Tanaka, T. (1984). Strand dissociation and cooperative melting of double-stranded dnas detected by denaturant gradient gel electrophoresis. *The Journal of Biochemistry*, 95(3), 627–635.
- Olson, W. K. (1996). Simulating dna at low resolution. *Current Opinion in Structural Biology*, 6(2), 242–256.
- Ousaka, N., & Inai, Y. (2009). Transfer of noncovalent chiral information along an optically inactive helical peptide chain: allosteric control of asymmetry of the c-terminal site by external molecule that binds to the n-terminal site. *The Journal of organic chemistry*, 74(4), 1429–1439.
- Pathria, R. (1984). *Statistical mechanics p. 184*. Pergamon Press.
- Pérez, A., Luque, F. J., & Orozco, M. (2012). Frontiers in molecular dynamics simulations of dna. *Accounts of chemical research*, 45(2), 196–205.
- Petkeviciute, D. (2012). *A dna coarse-grain rigid base model and parameter estimation from molecular dynamics simulations* (Tech. Rep.). EPFL.
- Petrosyan, R. (2017). Improved approximations for some polymer extension models. *Rheologica Acta*, 56(1), 21–26.
- Peyrard, M., & Bishop, A. R. (1989). Statistical mechanics of a nonlinear model for dna denaturation. *Physical review letters*, 62(23), 2755.

- Phillips, R., & Rob, P. (2001). *Crystals, defects and microstructures: modeling across scales*. Cambridge University Press.
- Purohit, P. K. (2008). Plectoneme formation in twisted fluctuating rods. *Journal of the Mechanics and Physics of Solids*, 56(5), 1715–1729.
- Purohit, P. K., Arsenault, M. E., Goldman, Y., & Bau, H. H. (2008). The mechanics of short rod-like molecules in tension. *International Journal of Non-Linear Mechanics*, 43(10), 1056–1063.
- Purohit, P. K., Kondev, J., & Phillips, R. (2003). Mechanics of dna packaging in viruses. *Proceedings of the National Academy of Sciences*, 100(6), 3173–3178.
- Purohit, P. K., & Nelson, P. C. (2006). Effect of supercoiling on formation of protein-mediated dna loops. *Physical review E*, 74(6), 061907.
- Raj, R., & Purohit, P. K. (2011). Phase boundaries as agents of structural change in macromolecules. *Journal of the Mechanics and Physics of Solids*, 59(10), 2044–2069.
- Rich, A., & Zhang, S. (2003). Z-dna: the long road to biological function. *Nature Reviews Genetics*, 4(7), 566–572.
- Ridge, G. S., Bailly, C., Graves, D. E., & Waring, M. J. (1994). Daunomycin modifies the sequence-selective recognition of dna by actinomycin. *Nucleic acids research*, 22(24), 5241–5246.
- Rouzina, I., & Bloomfield, V. A. (1999). Heat capacity effects on the melting of dna. 2. analysis of nearest-neighbor base pair effects. *Biophysical journal*, 77(6), 3252–3255.
- Rouzina, I., & Bloomfield, V. A. (2001a). Force-induced melting of the dna double helix 1. thermodynamic analysis. *Biophysical journal*, 80(2), 882–893.
- Rouzina, I., & Bloomfield, V. A. (2001b). Force-induced melting of the dna double helix. 2. effect of solution conditions. *Biophysical journal*, 80(2), 894–900.
- Sarkar, A., Leger, J.-F., Chatenay, D., & Marko, J. F. (2001). Structural transitions in dna driven by external force and torque. *Physical Review E*, 63(5), 051903.
- Singh, J., & Purohit, P. K. (2017). Structural transitions in torsionally constrained dna and their dependence on solution electrostatics. *Acta biomaterialia*, 55, 214–225.
- Storm, C., & Nelson, P. C. (2003). Theory of high-force dna stretching and overstretching. *Physical Review E*, 67(5), 051906.
- Weber, G., Essex, J. W., & Neylon, C. (2009). Probing the microscopic flexibility of dna from melting temperatures. *Nature Physics*, 5(10), 769–773.

- Wiggins, P. A., Dame, R. T., Noom, M. C., & Wuite, G. J. (2009). Protein-mediated molecular bridging: a key mechanism in biopolymer organization. *Biophysical journal*, *97*(7), 1997–2003.
- Williams, M. C., Wenner, J. R., Rouzina, I., & Bloomfield, V. A. (2001). Effect of ph on the overstretching transition of double-stranded dna: evidence of force-induced dna melting. *Biophysical Journal*, *80*(2), 874–881.
- Wolfe, K. C., Hastings, W. A., Dutta, S., Long, A., Shapiro, B. A., Woolf, T. B., ... Chirikjian, G. S. (2012). Multiscale modeling of double-helical dna and rna: A unification through lie groups. *The Journal of Physical Chemistry B*, *116*(29), 8556–8572.
- Zhang, F., & Collins, M. A. (1995). Model simulations of dna dynamics. *Physical Review E*, *52*(4), 4217.
- Zhang, X., Chen, H., Fu, H., Doyle, P. S., & Yan, J. (2012). Two distinct overstretched dna structures revealed by single-molecule thermodynamics measurements. *Proceedings of the National Academy of Sciences*, *109*(21), 8103–8108.
- Zhang, X., Chen, H., Le, S., Rouzina, I., Doyle, P. S., & Yan, J. (2013). Revealing the competition between peeled ssdna, melting bubbles, and s-dna during dna overstretching by single-molecule calorimetry. *Proceedings of the National Academy of Sciences*, *110*(10), 3865–3870.
- Zhao, Q., & Purohit, P. K. (2016). Phase boundaries with discontinuous stretch and twist in dna. *International Journal of Solids and Structures*, *94*, 238–246.