Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Wenxuan Xu Date

Supercoiling and Protein-mediated loops in Bacterial Transcription

By

Wenxuan Xu

Doctor of Philosophy

Physics

Laura Finzi

Advisor

David Dunlap

Committee Member

Francesca Storici

Committee Member

Irina Artsimovitch

Committee Member

Justin Burton

Committee Member

Kurt Warncke

Committee Member

Minsu Kim

Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D, MPH

Dean of the James T. Laney School of Graduate Studies

Date

Supercoiling and Protein-mediated loops in Bacterial Transcription

By

Wenxuan Xu

B.S., University of Science and Technology of China, 2012

Advisor: Laura Finzi, PhD; David Dunlap, PhD

An abstract of

A dissertation submitted to the faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Physics

2021

Abstract

Gene regulation comprises a complex network of events. Transcription is the first step of gene expression and is regulated by proteins that bind to DNA. Such proteins may facilitate or interfere with any of the three steps of transcription: (i) initiation, (ii) elongation, (iii) termination. DNA-bound proteins may also be roadblocks along the DNA template affecting RNA polymerase (RNAP) processivity. Protein mediated looping is a ubiquitous mechanism for DNA compaction and long-range interaction between distant DNA sites, either of which can regulate transcription. Similarly, DNA supercoiling, the level of which is modulated by DNA-binding proteins, DNA-processing enzymes, including RNAP, is an inherent regulatory mechanism. To understand the interplay between transcription, protein-mediated DNA looping, and supercoiling is important to understand gene regulation. The experiments described here, which leveraged the power of magnetic tweezers and tethered particle motion microscopy and used the lac repressor protein (LacI) as a model DNA-binding and looping protein, provide insight into this problem. In particular, they show that (i) negative supercoiling can make protein-mediated looping deterministic and ergodic within the bacterial doubling time, (ii) although the LacI repressor mediating a loop is a very strong roadblock for an incoming RNAP external to the loop, (iii) RNAP may become trapped inside a loop, in which case (iv) positive supercoiling generated by RNAP progress facilitates LacI repressor dissociation from its binding site, allowing RNAP to exit from a loop.

Supercoiling and Protein-mediated loops in Bacterial Transcription

By

Wenxuan Xu

B.S., University of Science and Technology of China, 2012

Advisors: Laura Finzi, PhD; David Dunlap, PhD

A dissertation submitted to the faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Physics

2021

Acknowledgement

I owe my deepest gratitude to my advisors, Dr. Laura Finzi, Dr. David Dunlap and for their generous guidance, supports and encouragement during the five and half years of my graduate study. This dissertation would be impossible without their kindly mentorship over me. It is also because of them; single molecule biophysics will guide me through my future life.

I also want to express my heartily thanks to my committee members: Dr. Justin Burton, Dr. Kurt Warncke, Dr. Minsu Kim, Dr. Irina Artsimovitch and Dr. Francesca Storici for their time, thoughtful criticism, encourage and attendance to my annual meetings.

I am grateful to my colleagues, Yan Yan, for her generous guidance and help through the first 4 years of my research; Jin Qian, for cooperation of transcription review, Derrica McCalla, for her technical supports. I want to thank all the members in Finzi/Dunlap Lab for assisting me on research.

I would like to also thank my parents, my mother Yu Wang and My father Binfeng Xu for their constant trust on me to encourage me to overcome difficulties during my study.

I am also grateful to my dear friends Yonglun Jiang, Guanxiong Chen, Yixuan Han and Chentao Li for their company.

I want to thank Laney Graduate School to offer me financial support to finish my Ph.D. study.

During Ph.D. program, I received numerous assistants from members in the Physics Department at Emory. The dissertation is dedicated to all of them.

Table of Contents

List of figures

Chapter 1 Introduction

Chapter 2 Material and Methods

Chapter 3 Energetics of twisted DNA topologies

Chapter 4 Negative DNA supercoiling makes protein-mediated looping deterministic and ergodic within the bacterial doubling time

Chapter 5 Positive supercoiling favors transcription elongation through lac repressor-mediated DNA loops

Chapter 1 Introduction

§1.1 Motivation and hypothesis

Deoxyribonucleic acid (DNA) encodes up to tens of thousands of genes and is highly compacted inside the cell. In some cases, the linear length of the DNA compacted inside the nucleus is as much as a million-fold longer than the length of a cell. Compaction of DNA requires supercoiling and protein-mediated DNA topologies, such as loops. How can, therefore, RNAP transcribe along supercoiled, protein-decorated DNA (Figure 1.1). Given the well-known, twin-domain model of transcription [1], which states that RNAP generates positive and negative supercoiling ahead and behind itself, respectively, it is natural to hypothesize that supercoiling generated by RNAP progress may play a key role in freeing the path from DNA-bound proteins for RNAP. Protein-mediated DNA loops bring distant DNA sites in contact and compact DNA; they are also common regulatory mechanisms of transcription. For example, eukaryotic transcription is almost always regulated by the kbp-range loop between an enhancer/silencer and a promoter while prokaryotes use few hundred bp-long loops to activate or repress essential metabolic genes [2-4] and temperate bacteriophages use looping, for example, to simultaneously maintain lysogeny and guaranteeing a reliable switch to lysis [5, 6]. Such loops represent genetic switches as they turn transcription of genes on or off depending on if they are in the closed or open configuration. However, just as any protein bound to DNA may be a roadblock to transcription, so may a protein securing a loop, especially if the DNA sequences bridged by the protein are part of different genes. Thus, I set out to investigate how RNAP navigates past protein-mediated loops.

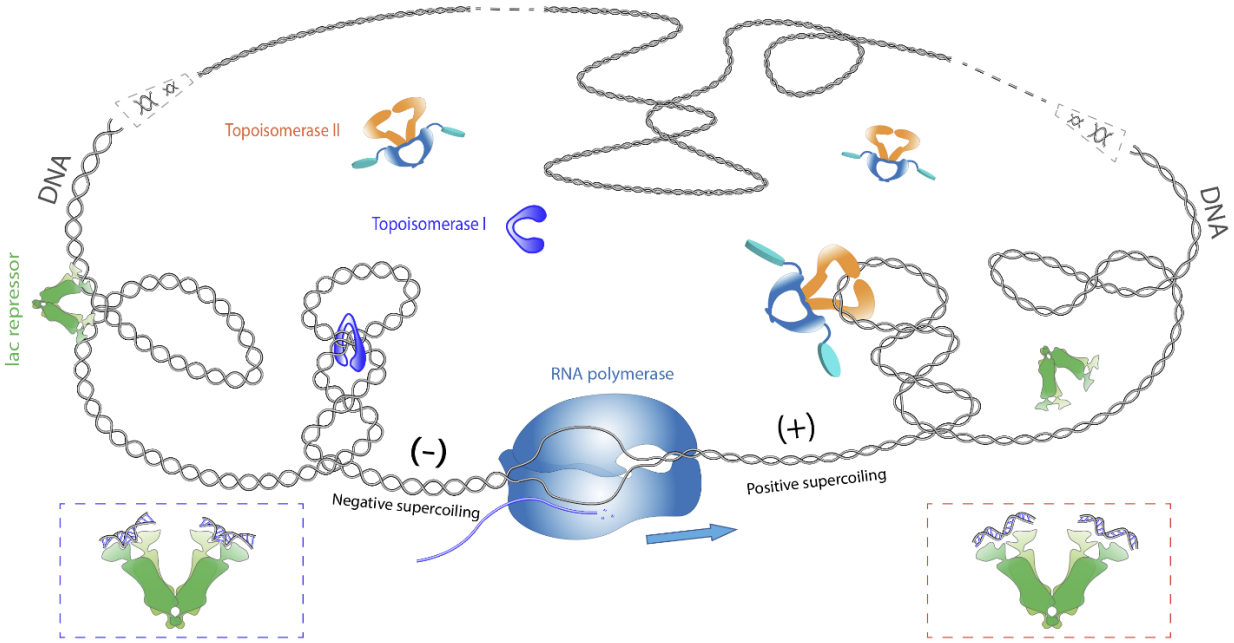


Figure 1. 1. A schematic illustration of the interplay between RNA polymerase elongation, topological DNA structures, and regulation by representative proteins and enzymes. During transcription of the torsionally constrained DNA segment, negative supercoiling accumulates upstream while positive supercoiling accumulates downstream (twin domain model [1]). Supercoiling accumulation is managed by topoisomerases, a special class of enzyme which can relax torsional stress. lac repressor is used here as an example of a transcriptionl roadblock and a DNA loop-mediating protein. Negative supercoiling enhances lac repressor binding [7] whereas positive supercoiling produces a shallower minor groove [8] that may attenuate binding, as shown in the blue and red dashed boxes, respectively.

§1.2 DNA structure

§1.2.1 DNA primary structure

DNA is a macromolecule made up two anti-parallel polynucleotide chains that coil around each other in a double helix. Each chain is composed of a sequence of deoxyribonucleotides which can contain one of four different bases: Adenine(A), Guanine(G), Thymine(T) or Cytosine(C). The double stranded helical

structure of DNA is stabilized by stacking interactions and hydrogen bonds (H-bonds) formed between pairs of bases across the two chains, according to base pairing rules (A pairs with T through two H-bonds, while C pairs with G through three H-bonds) (Figure 1.2A). The best known DNA form is *B*-DNA (Figure 1.2B, left) where the double helix is right-handed with ~10.5 base pairs per turn [9], the average distance between two adjacent base pairs is around 0.34 nm. The DNA double helix is characterized by a minor groove (narrower) and a major groove (wider). Different proteins show preferential interaction with either the major or minor groove [10]. *Z*-DNA (Figure 1.2B, right) is a left-handed, zigzag-like form of the DNA double helix [11]. Although this form is not as well-known as the *B*-form, it may occur more commonly than expected. Certain DNA sequences like $d(GC)_n$ have the tendency to adopt the *Z*-form [12] which is also induced by negative DNA supercoiling under tension, or high salt concentration [13-15].

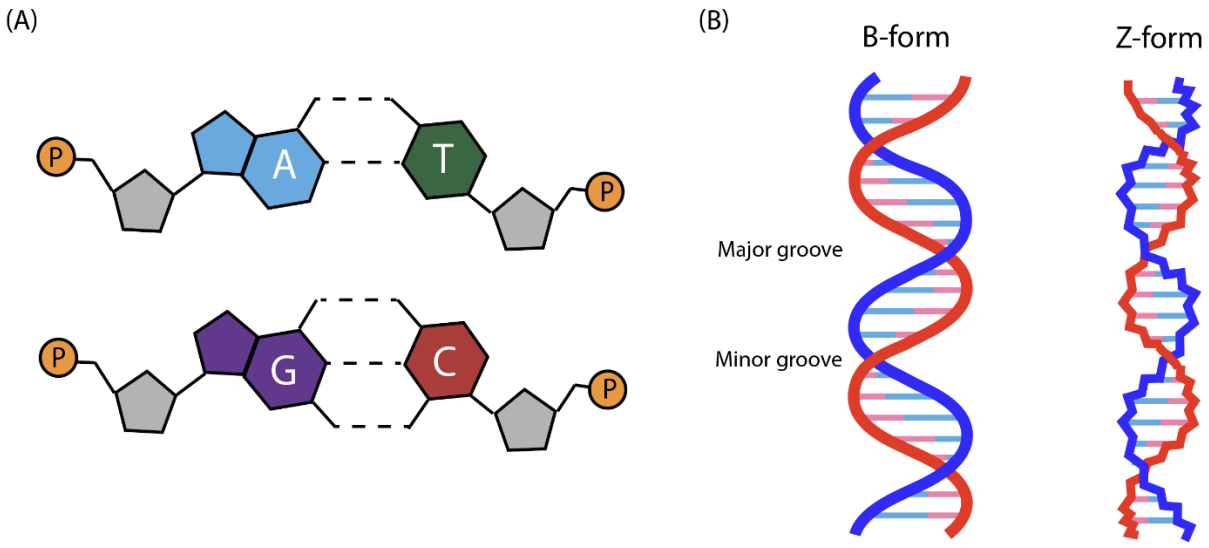


Figure 1. 2. Sketch of basic structures of DNA. (A) AT and GC Base pairing. (B) *B*- and *Z*-form DNA.

§1.2.2 DNA structure under tension and torsion

In vivo, DNA experiences both tensile and torsional stress, which mainly result from the topological constraints of DNA and the activities of DNA-processing enzymes. For example, *E. coli* RNA polymerase (RNAP) twists DNA positively ahead and negatively behind itself as it unwinds the double helix to gain

access to the template strand during transcription [1], while topoisomerases help relieve the accumulated torsional stress by nicking (breaking) one, or both, DNA strands and rejoining them after relaxation (Figure 1.1). It has also been tested *in vitro* that the *E. coli* RNAP motor is capable of generating forces up to 25 pN during elongation [16], DNA polymerase, the enzyme that carries out DNA replication [17], can generate force against large load (~50 pN) [18], etc. Therefore, tension and torsion are ubiquitous elements *in vivo*, and their effects on DNA structure have been widely studied by various *in vitro* single molecular technologies and fitted with theoretical models. Thus, the form of DNA is very likely to change dynamically as a function of local levels of tension and twist.

DNA supercoiling refers to the over- or under-winding of a double-stranded DNA molecule which is often quantified as the ratio of the number of turns added, or subtracted, to a DNA molecule, n, over the number of helical turns in its torsionally relaxed form, n_0. This ratio defines the change in linking number, ΔLk, with respect to the torsionally relaxed state. Expressed in percentage, ΔLk is called σ, or superhelical density. The linking number itself is defined as, $Lk = Tw + Wr$ where twist (Tw) is the number of times the single strands of the DNA helix intertwine, and writhe (Wr) is the number of times the main axis of double helix crosses itself. By convention right-handed intertwining is considered positive. The linking number of relaxed *B*-DNA with no torsion is given by $Lk_0 = n_0$ (bp) / 10.5 (bp/turn) in which n_0 stands for the total number of base pairs in the DNA molecule. In this case, $Lk_0 = Tw_0$ and $Wr = 0$. Thus, when DNA is under torsional stress, the change in linking number is $\Delta Lk = Lk - Lk_0 = \Delta Tw + Wr$. According to White's theorem [19], the linking number is a constant in a topologically constrained system. Genomic DNA has large negatively supercoiled regions as shown by psoralen staining of eukaryotic cells [20] or bacteria [21], and plasmids extracted from bacteria display supercoiling as low as -7.75% and rise to -6.5% as cultures reach saturation [22].

Under low tension ($f \lesssim 10$ pN), in the absence of, or below 40 pN·nm positive torque, DNA remains in *B*-form. Added torsion initially generates a twisted molecule, but when the torsional stress on the DNA exceeds a critical level, at the buckling transition [23], a "curl" forms and the DNA writhes and crosses over

itself to form a loop (Figure 1.3). The aggregate writhe and twist in DNA determine the supercoiling, which can be positive or negative. Further torsion applied to the DNA molecule adds gyres to the stem of this loop to form a plectoneme, which may encompass a long segment of DNA [24] (Figure 1.3). Plectonemes are fundamental elements of compacted DNA and facilitate interactions between distant DNA sites [25, 26] and their associated proteins [27-29]. The number of plectonemes that form depends on the applied tension and ionic strength [30]. They have been observed to diffuse along DNA, as well as to abruptly disappear at one position before reappearing at a another [31], although intrinsically curved DNA sequences, DNA deformability, and DNA-protein complexes can pin plectonemes at a specific site [32-37].

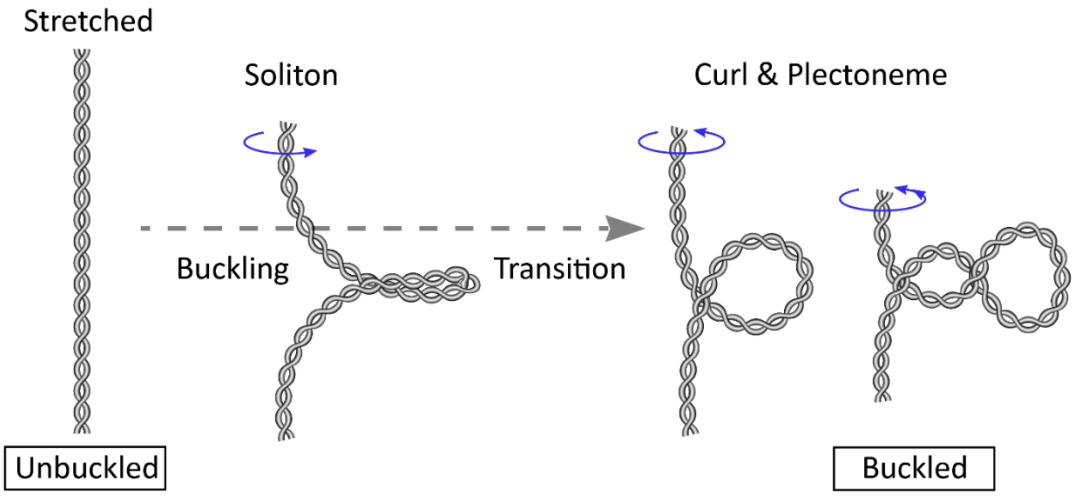


Figure 1. 3. Sketches of the plectoneme formation. Winding DNA may produce a "soliton" in which the DNA curls about the long axis, without significantly changing the tether length. Detectable tether length changes result when the curl aligns perpendicular to the direction of extension and subsequently twists to form a plectoneme.

Plectoneme formation is not the only structural response of the DNA helix to tensile and torsional stress. Previous experiments have shown that even a small negative torque (when $f < 10$ pN, $\tau < 40$ pN·nm) may induce stretches of a DNA molecule to transition from *B*- to a left-handed *L*-form (*L*-DNA) [38]. *L*-DNA

appears to be a mixture of melted and base-paired DNA with left-handed intertwining of the two strands. AT-rich regions, being less thermodynamically stable than sequences rich in GC, may melt if unwound under tension [39]. Thus, in negatively supercoiled plasmids, single-stranded regions may occur [40] as well as *Z*-DNA, one of the non-canonical, left-handed DNA forms adopted by GC-rich sequences in high salt and believed to be important during transcription [41, 42]. Only a slight negative torque of ~-3 pN·nm can trigger the *B*-*Z* transition in GC/GT-rich regions under 1.4 pN of tension [15] indicating that this transition may be more common than melting in physiological conditions, since the average torque for DNA melting is ~10 pN·nm. Although it is unlikely that a plasmid will fully convert from *B*- to *L*-DNA, certain segments may adopt this form.

§1.2.3 Nucleoid-associated protein HU

E. coli HU is one of the most abundant nucleoid-associated proteins. HU binds to all nucleic acids and DNA-RNA hybrids in a sequence nonspecific manner [43]. *In vitro*, HU acts similarly to histone proteins and can introduce negative supercoiling into torsional relaxed DNA with the help of topoisomerase I [44]. *In vivo*, HU was found to help maintenance of DNA super-helical density and modulate topoisomerase I activity [45]. HU dimers were able to bend DNA [46] and compact DNA [47], which favors looping activities *in vivo* [48, 49] (Figure 1.4).

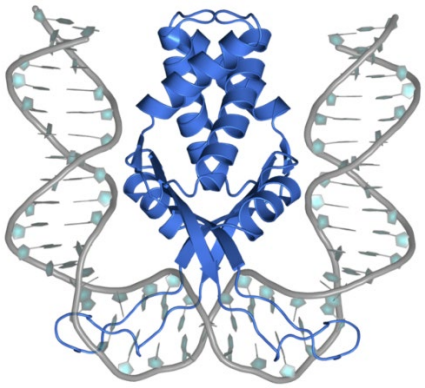


Figure 1. 4. Protein HU (blue in middle) binds with DNA (gray). (From Protein Data Bank in Europe).

§1.3 Transcription in Bacteria

§1.3.1 Basic concepts of bacterial transcription

Transcription is the first step of gene expression. Bacterial transcription refers to the process by which genetic information encoded into special DNA sequences called genes is transcribed into messenger RNA (m-RNA) by the motor enzyme RNA polymerase (RNAP). Every organism has slightly different RNA polymerases. In my studies, I used *E. coli* RNA polymerase whose core consists of 5 subunits: α, α, β′, β, ω. The three main steps of transcription are initiation, elongation (Figure 1.5) and termination. Initiation refers to the phase following promoter unwinding (open complex formation) by RNAP holoenzyme (core plus σ^{70}), when RNAP can start "reading" one of the two DNA strands. My project focused on elongation, the steady elongation of transcript RNA after the first initial 21 nucleotides. During elongation, RNAP unwinds the DNA ahead and progresses along it to transcribe one strand of DNA (template strand) into mRNA by hybridizing incoming ribonucleoside triphosphates at the active site. The newly synthesized nascent mRNA exits from the RNA exit channel. Finally, RNAP stops at a specific DNA sequence called terminator where the mRNA transcript is released, and transcription ends. There are two mechanisms of termination, rho dependent or rho independent.

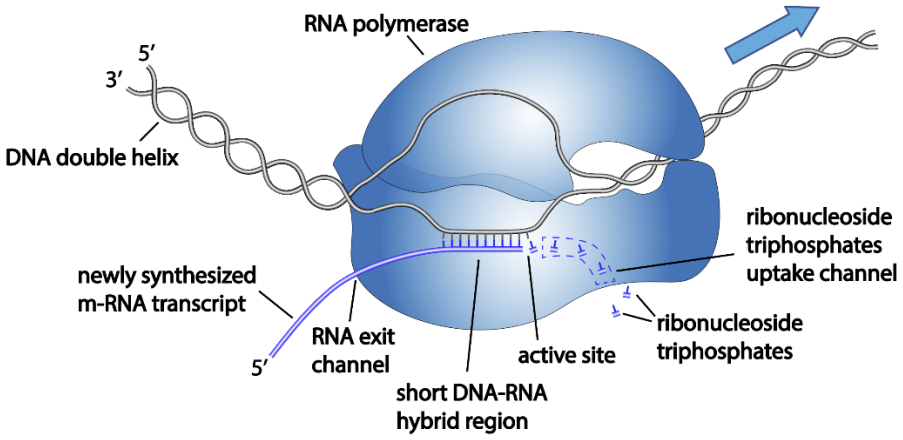


Figure 1. 5. Cartoon of elongation by RNA polymerase. The blue arrow indicates the elongation direction.

§1.3.2 Transcription regulation: DNA looping and roadblocks

Transcription is regulated by various proteins and corresponding DNA binding sequences. The most basic mechanism of transcription regulation relies on proteins (transcription factors) that interact with RNAP in order to either favor or disfavor its activity. Accordingly, transcription of a gene may be activated by an "activator" protein which binds at a site adjacent to the promoter for that gene to increase the level of expression. On the other hand, a "repressor" protein may bind to a specific site that partially overlaps with the promoter in order to sterically prevent RNAP binding to decrease the expression of a gene. External stimuli determine the level of expression of activators vs. repressors. Both repression and activation are strengthened by long-range interactions where the activator/repressor binds simultaneously not only near the promoter but also a second site further away, causing the looping out of the intervening DNA. This bridging of two distant sites causes an increase in protein local concentration and, therefore, a cooperativity effect that thermodynamically stabilizes the protein at the site near the promoter. Protein-mediated DNA looping can occur over different ranges (from a few base pairs to hundreds of kilobase pairs). For example, the AraC, Lac, Deo and Gal repressors, all have at least two specific binding sites, one of which is close to the promoter, while the other is one to few hundreds of base pairs away and mediate small loops which repress more effectively the genes these repressors regulate [2-4]. Now, the second binding site (promoter distant) of proteins like these may be located in a different gene, meaning that the RNAP transcribing that second gene will run into an obstacle constituted by a protein mediating a loop. I will call the physical obstacle represented by a protein on the path of an elongating RNAP on the DNA template a "roadblock". *In vitro*, roadblocks can be DNA-binding proteins like repressors, nucleoid architectural proteins [50], protein that direct replication [51], DNA-based molecular motors [52], repair proteins, proteins part of editing complexes, such as RNA-directed Cas9 [53]. These factors interact with transcription factories inside the dynamic, structured nucleoid [54-58] and can cause RNAP to pause and backtrack [59]. In this dissertation the Lac repressor serves as a either a looping or roadblock, or both.

§1.3.3 *lac* repressor (LacI) and LacI mediated loop

The lac repressor (LacI) is a DNA binding protein that binds to specific DNA sequences (operators). It is produced to repress the expression of the lac operon (a sequence of three genes responsible for the metabolism of lactose in *E. coli*) by binding at the O1 operator which partially overlaps the operon promoter. Structurally, lac repressor is a homo-tetramer with two pairs of dimers interacting via their C-terminal domains (CTDs) to form the "hinge" of a "V". The ends of the V each contain a DNA binding NTD (Figure 1.6A). Such V-shaped structure enables LacI to bind to two operators simultaneously and loop DNA. There are four types of operators with different binding affinities for LacI which in order of affinity are: "Oid", "O1", "O2" and "O3". Oid is an engineered operator while the other three are naturally occurring. *In vivo*, O1 slightly overlaps with the lac operon promoter while O2 is located 401 bp downstream of O1 and O3 is located 92 bp upstream of O1 [60]. These three operators cooperatively strengthen the effective binding of LacI to O1 and, consequently, the repression of the lac operon [61] (Figure 1.6B).

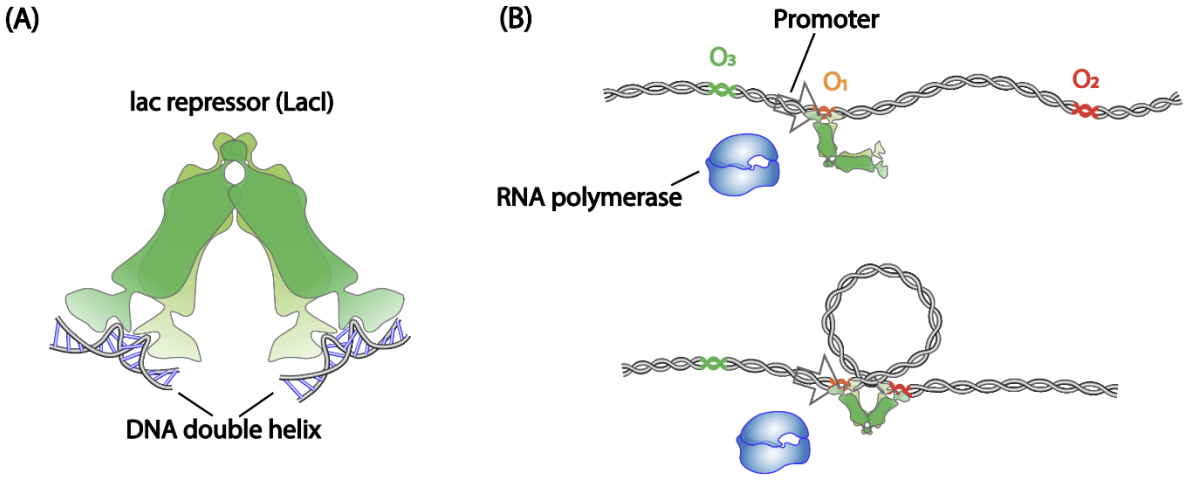


Figure 1. 6. Cartoons of LacI and its mechanism of gene expression regulation. (A) LacI tetramer bound to two DNA operators. (B) LacI represses transcription *in vivo*. The promoter (black outlined arrow) slightly overlaps with O1 operator. LacI can block RNAP initiation by simply binding to O1 (top panel) or mediated a loop (bottom panel).

Chapter 2 Material and Methods

§2.1 DNA preparation

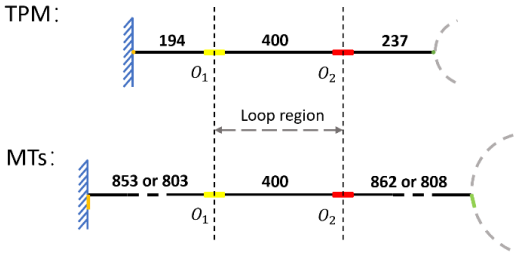
The DNA fragments used in all my studies were produced by Polymerase Chain Reaction (PCR). The reaction mixture contains the desired plasmid, primers (eurofins Genomics) designed using "A plasmid Editor" (ApE), deoxynucleotide triphosphates (New England BioLabs, NEB), Taq/Q5 DNA polymerase (NEB), and nuclease-free water. The annealing temperature was calculated using NEB Tm calculator.

Multiple biotins or digoxigenin labeled DNA tails were required for the experiments where tethers are torsional constrained, especially for magnetic tweezers (MTs) experiments. Such ~150 bp biotin- or digoxigenin-labeled DNA anchor fragments were generated by ApaI or XmaI restriction in the middle of 302 bp PCR amplicons produced using dATP, dCTP, dGTP, dTTP (Fermentas-Thermo Fisher Scientific Inc., Pittsburgh, PA, USA), and digoxigenin-11-dUTP (Roche Life Science, Indianapolis, IN, USA), or biotin-11-dUTP (Invitrogen, Life Technologies, Grand Island, NY, USA), in a molar ratio of 1:1:1:0.7:0.3. The product of ligation of the main fragment and two anchor fragments was purified and the total variation in DNA tether length, due to random placement of 30% biotin- or digoxigenin along the anchor fragments did not exceed 50 bp (16 nm) [29].

§2.1.1 DNA for the looping project in chapter 4

Details about template and primers that were used to generate the MT DNA constructs are summarized in Figure2.1.

Construct	Template	Sense primer	Anti-sense primer	Restriction enzymes
O1-400-O2 (TPM)	pO1O2[1]	D-S/pO1O2_401/1588	B-A/pO1O2_401/2418	none
O1-400-O2 (MT)	pO1O2, pZV_21_400[2]	S/pO1O2_401/929_XmaI, S/JBOIDO1_400/1799_ApaI	A/pO1O2_401/3043_ApaI, A/JBOIDO1_400/3810_XmaI	XmaI, ApaI
3352 bp segment (MT)	pYY_I1_400_BstEII [3]	S/JBOIDO1_400/1783-Apa1	A/JBOIDO1_400/5113-Xma1	XmaI, ApaI
Bio-, dig-tails	pBluKSP+	A/pIT_Loop3/1665	Pol1979R	XmaI, ApaI



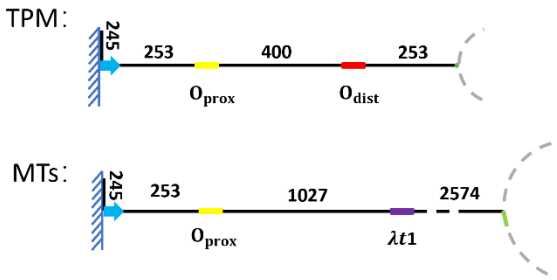
Primer name	Sequence
D-S/pO1O2_401/1588	[dig]-tgctcgcttcgctacttg
B-A/pO1O2_401/2418	[bio]-tgactgggttgaaggctc
S/pO1O2_401/929_XmaI	tgCCCGGGacccggaaagacatgc
A/pO1O2_401/3043_ApaI	ctGggCCCggtgaatccgttagcga
S/JBOIDO1_400/1997_ApaI	acccgtGGGCCCagcatcctctctcgtttcatc
A/JBOIDO1_400/3810_XmaI	gcagcgCCCGGGtgagcgaggaagcggaagag
A/pIT_Loop3/1665	ggcgattaagttgggtaacg
Pol1979R	tgtggaattgtgagcggata
A/JBOIDO1_400/5113-Xma1	acaacacccgggATCATGTAACTCGCCTTG
S/JBOIDO1_400/1783-Apa1	atgttcgggcccAGTAACCCGTATCGTGAGC

Figure 2. 1. Plasmids, primers, and templates for the looping project in chapter 4. (Top table) The DNA constructs were constructed using plasmid templates with the indicated pairs of primers. (Bottom table) Sequences for each primer. Bio- or B- is a biotin tag at the 5' end of the primer, and D- or dig- is a digoxigenin tag at 5' end of the primer. Cartoons at left bottom shows the templates of DNA tethers. Blue dashed wall on the left represents the coverslip, the gray dashed half circles on the right represent beads. Unit of DNA length is base pair.

§2.1.2 DNA for the transcription project in chapter 5

Details about template and primers that were used to generate the MT DNA constructs are summarized in figure 2.2.

Construct	O1-400-O2 (TPM)	O2-400-O1 (TPM)	O1 (MT)	Bio-, dig-tails
Template	pWX_12_400	pZV_21_400	pRS_1N_400	pBluKSP
Forward primer	S/JBOIDO1_400/2086	S/JBOIDO1_400/2086	S/JBOIDO1_400/2086	A/pIT_Loop3/1665
Reverse primer	B-A/YY400/3342	B-A/YY400/3342	A/pO1O2_401/3043_ApaI	Pol1979R
Restriction enzymes			ApaI	ApaI



Prime name	Sequences
S/JBOIDO1_400/2086	agcttgtctgtaagcggatg
B-A/YY400/3342	[bio]-accgcatcagcaagtgtat
A/pO1O2_401/3043_ApaI	ctgggcccggtgaatccgttagcga
A/pIT_Loop3/1665	ggcgattaagttgggtaacg
Pol1979R	tgtggaattgtgagcggata

Figure 2. 2. Plasmids, primers, and templates for transcription project in chapter 5. (Top table) The DNA constructs were constructed using plasmid templates with the indicated pairs of primers. (Bottom table) Sequences for each primer. Bio- or B- is a biotin tag at the 5' end of the primer. Cartoons at left bottom shows the templates of DNA tethers. Blue dashed wall on the left represents the coverslip, the gray dashed half circles on the right represent beads. λt1 is the name of terminator. Unit of DNA length is base pair.

§2.2 Chamber preparation

The bottom coverslip of the microchamber (Fisherbrand, Thermo Fisher Scientific, Waltham, MA, USA) supported a parafilm gasket produced with a laser cutter (Universal Laser Systems, VLS 860, Middletown, CT) with a central observation area connected through narrow inlet and outlet channels to inlet and outlet reservoirs, extending beyond the edges of the top coverslip (Figure 2.3). The coverslips with parafilm assembly were heated to seal its components together and form the microchamber. The narrow inlet and

outlet reduced the evaporation of the buffer, while the triangular shape of the chamber, confined the reaction in a relatively small volume and provided a gradient of tether densities to optimize throughput.

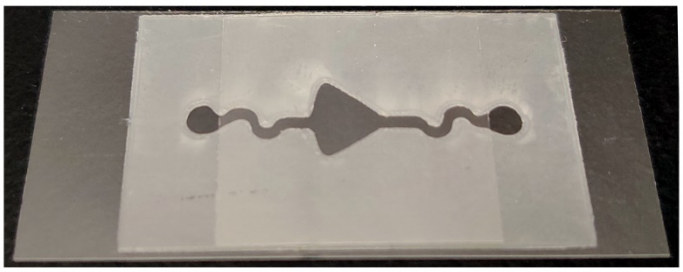


Figure 2. 3. Image of a typical micro chamber. The bottom coverslip supports a parafilm gasket with inlet and outlet reservoirs at the edges of the top coverslip connected through narrow inlet and outlet channels to the central observation area under the coverslip. The parafilm and coverslips are heated and sealed together, and the total volume of the chamber is approximate 6 μL.

The buffer used in different projects were slightly different. Basic Incubation buffer (BIB: 20 mM Tris-Cl (pH 7.4), 50 mM KCl, 1 mM DTT) and transcription buffer (TXB: 20 mM Tris-glutamate (pH 7.4), 10 mM magnesium glutamate, 50 mM potassium-glutamate, 0.2 mg/mL α-casein (Sigma-Aldrich, St. Louis, MO), 1 mM DTT) were used for the transcription experiments while phosphate buffered saline (PBS) and λ buffer (10 mM Tris-HCl (pH 7.4), 200 mM KCl, 5% DMSO, 0.1 mM EDTA, 0.1 mg/mL α-casein, 0.2 mM DTT) were used in DNA looping related non-transcriptional experiments. The entire sample preparation was performed at room temperature (~25 °C) and materials were kept on ice. The beads used in TPM measurements were 0.32 μm diameter, streptavidin-coated polystyrene beads (Spherotech, Lake forest, IL, USA), while the beads used in MT measurements were 1.0 μm diameter, streptavidin-coated super-paramagnetic beads (Dynabead MyOne Streptavidin T1, Invitrogen, Grand Island, NY).

For looping projects, the chambers were first incubated with reference beads resuspended in PBS for 5 minutes to let some beads adhere to the surface and serve as references during data acquisition and analysis. The chambers were then incubated with 4 μg/mL anti-digoxigenin (Roche Life Science, Indianapolis, IN,

USA) in PBS at room temperature for 30 min. They were then passivated with PBS supplemented with 6 mg/mL α-casein at room temperature for 1 hour. DNA tethers in λ buffer were then introduced into the chamber for 15 min incubation to be attached through a single digoxigenin (TPM), or a multiply digoxigenin-labeled tail (MTs) to the anti-digoxigenin coated coverslip. The opposite end of the DNA was attached to a streptavidin-coated bead via a single biotin (TPM), or a multiply biotin-labeled tail (MTs) by flowing in the microchamber 0.03 mg/mL beads solution and letting them incubate for 10 min (TPM, Figure 2.4A) or 0.02 mg/mL bead solution incubating for 5 min (MTs, Figure 2.9A). Both types of beads were washed in PBS twice and resuspended in λ buffer before introducing them into the chamber.

For transcription measurements, the chambers were first incubated with reference beads resuspended in BIB for 5 minutes to let some beads adhere to the surface and serve as references during data acquisition and analysis. The chambers were then incubated with 10 g/mL purified Anti-HA 11 Epitope tag antibody (BioLegend, San Diego, CA, USA) in BIB at room temperature for 1 hour. They were then passivated with BIB supplemented with 6 mg/mL α-casein at 4 °C overnight. 60 nM doubly-HA tagged *E. coli* RNA polymerase were then drawn into the chamber and incubated 30 min at room temperature to let the HA-labeled RNA polymerase bind to the anti-HA coated surface. 10 nM DNA template, 50 μM GpA (initiating dinucleotide, TriLink Bio Technologies, San Diego, CA, USA), and 10 μM ATP/UTP/GTP in TXB were introduced into the chamber to produce the stalled elongation complexes (SECs). For TPM experiments (Figure 2.4B), the end of the DNA far from the promoter was labeled with beads by flowing in the microchamber 0.03 mg/mL beads resuspended in TXB and letting them incubate for 10 min then wash out. For MTs experiments (Figure 2.9B), instead, a 0.02 mg/mL bead soμution was flowed in and incubated for 5 min then wash out. Both types of beads were washed in PBS twice and resuspended in TXB before introducing them into the chamber.

§2.3 Proteins

LacI was provided by Kathleen Matthews (Rice University). *E. coli* HU protein was overexpressed and purified in Fengfei Leng's lab (Floride International University). RNA polymerases were provided by Karen Adelman (Harvard University) and Irina Artsimovitch (Ohio State University).

§2.4 Tether particle motion (TPM) microscopy

§2.4.1 Experimental setups for TPM measurements

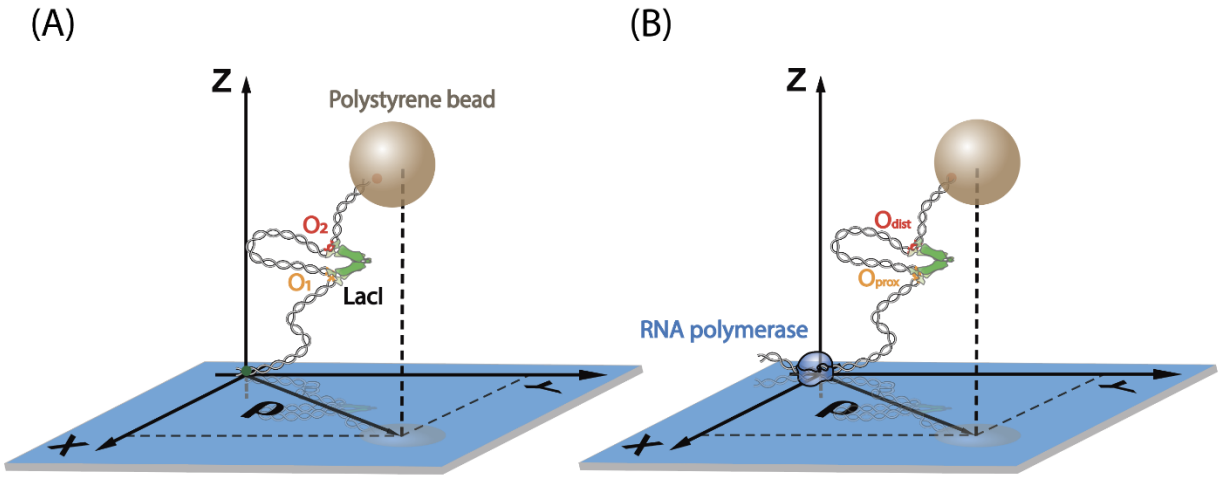


Figure 2. 4. Cartoons of the TPM experimental setup. (A) General TPM setup for looping projects. One end of DNA is immobilized on glass surface while the other end is linked to polystyrene bead. (B) General TPM setup for transcription projects. A stalled RNAP-DNA complex is immobilized on glass surface while the other end is linked to polystyrene bead. The projection of bead on XY plane gives ρ value relative to the anchor point which will then be used to analyze DNA excursion.

Tethered particle motion (TPM) microscopy is a simple but powerful tool for quantitative analysis of various tethered single polymers (DNA in this case) and their interaction with other entities (proteins in this case). The basic idea is to record the Brownian motion of microspheres tethered to a glass surface by single polymer molecules, analyze the effective length change of the tethered molecule in order to deduce its activity. Usually, the projection of tethered microspheres on the XY plane is recorded to infer the position of the anchor point and then the 2-D projection, ρ, of the 3-D position of the tethered bead. The ρ value provides a direct read-out of the effective tether length to which it can be converted using an appropriate calibration curve. Figure 2.4 shows the experimental setups of TPM used in this dissertation.

§2.4.2 TPM microscope

The microscope (Leica DM LB-100, Leica Microsystems, Wetzlar, Germany) I used is equipped with an oil-immersion objective (N-Plan 100 × 1.4) and a CV-A60 CCD camera (JAI, Copenhagen, Denmark). The video sequence was digitized with an IMAQ PCI-1409 frame grabber (National Instruments, Austin, TX) and analyzed using a custom Lab View (National Instruments, Austin, TX) routine. The exposure time of the CCD camera was minimized to avoid blurring that otherwise reduced the apparent amplitude of the motion of the bead. The exposure had to be much shorter than the time, t_b, required by the bead to traverse its range of motion. A rough estimate is given by $t_b = \sigma^2/D$, where σ is the range of motion and $D \approx 1000$ nm^2/ms is the diffusion coefficient of the bead obtained using the Stokes-Einstein formula. Typical values of t_b are about 50 ms, I used 1 ms exposures, which are safely lower than tb. Standard PAL video cameras use an interlaced format: the even lines of each image are exposed 20 ms after the odd lines. To avoid blurring from this 20 ms offset, I analyzed the even and odd fields as independent sequences. This procedure effectively doubled the pixel spacing in the vertical direction (from 64 to 128 nm in the field of view) but did not significantly impair my determination of bead positions.

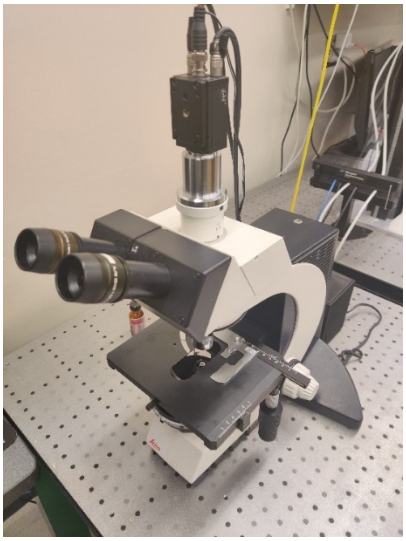


Figure 2. 5. Photograph of the TPM microscope in the Finzi-Dunlap lab used for the measurements described in this dissertation.

§2.4.3 Data collection and analysis of TPM

The absolute XY positions of each bead was recorded at 50 Hz with a custom Labview (National Instruments, Austin, TX, USA) program. Vibrational or mechanical drift in the position of each tethered bead was removed by subtracting the average position of multiple, stuck reference bead(s) within the same field of view [29, 62, 63]. The excursion of each tether was then calculated as $< \rho >_{4/8s} = <\sqrt{(x - < x >_{4/8s})^2 + (y - < y >_{4/8s})^2} >_{4/8s}$, in which $< x >_{4/8s}$, and $< y >_{4/8s}$ are eight-second (looping projects) / four-second (transcription projects) moving averages representing the coordinates of the anchor point of a bead. Changes in the excursion of the bead reflect conformational (length) changes of the DNA tether [64-66]. The beads that exhibited (*x, y*) position distributions with a ratio of the major to minor axes greater than 1.07, were discarded, since they were likely to be tethered by multiple DNA molecules [63]. The excursion data from the time records of the beads, in the same experimental conditions, which passed this "symmetry test" were pooled for the following analysis.

§2.4.3.1 Looping probability Calculation

For studies of DNA looping, the excursion data was used to generate probability distribution histograms. These summarized the average excursion distribution and, in the presence of LacI, they indicated three excursion peaks (two looped states and one unlooped state, figure 2.6) [29]. The histogram of each selected temporal trace was fitted with three Gaussians, then the looping probability was calculated by dividing the area under the Gaussians corresponding to the two looped states (peaks with shorter excursion) by the total area under all three Gaussians. The mean value of looping probability under each protein condition were weighted by the length of each trace.

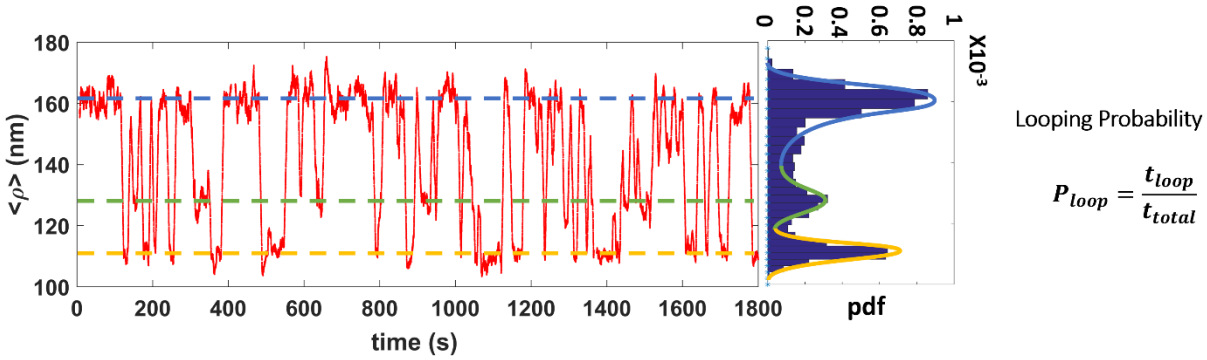
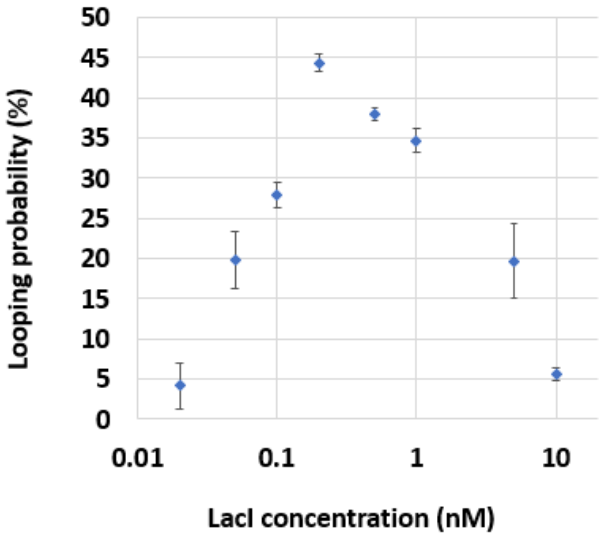


Figure 2. 6. Looping probability calculation. A representative TPM trace (red) shows loop formation and breakdown, the unlooped state (blue dashed line) and the two looped states (green and yellow dashed lines) are indicated. The probability distribution function is shown with blue bars fitted with the sum of three Gaussians.

§2.4.3.2 Looping probability distribution in transcription buffer (TXB)

The looping probability distribution in TXB buffer was measured as a function of LacI concentration and is shown in figure 2.7. The highest looping probability of ~44% was obtained with a LacI concentration of 0.2nM. The looping probability distribution in λ buffer is shown in figure 4.1.

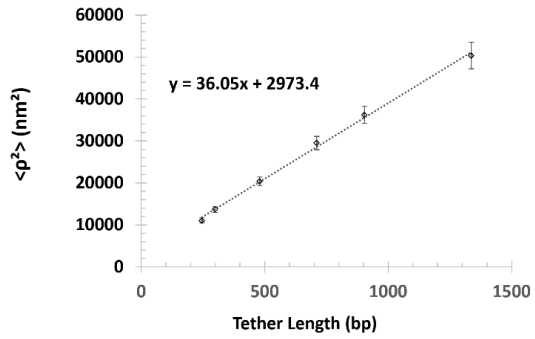


[LacI] (nM)	Looping Probability	Stand_Error
0.02	0.041461	0.028293
0.05	0.199028	0.035377
0.1	0.278917	0.015676
0.2	0.443856	0.010945
0.5	0.379909	0.007482
1	0.3471	0.015456
5	0.197158	0.046178
10	0.056030	0.007239

Figure 2. 7. Looping probability vs LacI concentration. The looping probability was measured in 904 bp-long DNA tethers with a 400 bp loop region under 0.02, 0.05, 0.1, 0.2, 0.5, 1, 5, 10nM LacI concentration in transcription buffer (TXB).

§2.4.3.3 Calibration of rho square

For transcription projects, in order to infer RNAP activity from the DNA tether length, a calibration curve relating tether length to $<\rho^2>$ values (Figure 2.8) was obtained.



Tether length (bp)	$<\rho^2>$ (nm^2)
246	11075
299	13680
480	20345
711	29508.5
904	36228.3
1337	50375

Figure 2. 8. Calibration of tether length and rho square. The recorded rho square values were converted into values of tether length by measuring rho square for tethers of DNA molecules with different known

lengths (244, 299, 480, 711, 904, 992, 1337 bp). The expression for the fitted line is: $< \rho^2 >= 36.05 *$ tether length $+ 2973.4$.

§2.5 Magnetic tweezers (MTs)

§2.5.1 Experimental set-up

Although different implementations of magnetic tweezers exist [29, 67-70], the basic design is to use a pair of permanent magnets to apply tension or torsion to super-paramagnetic beads tethered by the molecule of interested to the surface of microscope flow-chamber. Cartoons of the experimental set-up I used are shown in figure 2.9. For both projects, the distances between magnet and beads were used to vary the tension applied to DNA. For looping projects (Figure 2.9A), the DNA tether is built torsional constrained (§2.2 Chamber preparation), rotation of magnetic tweezer adds supercoiling into DNA tethers. For transcription projects (Figure 2.9B), RNAP was linked onto anti-HA coated coverslips by two HA-anti-HA linkages, while the bead labels the downstream end of DNA by a single, or multiple, biotin-streptavidin linkage(s) depending on the experimental torsional requirements.

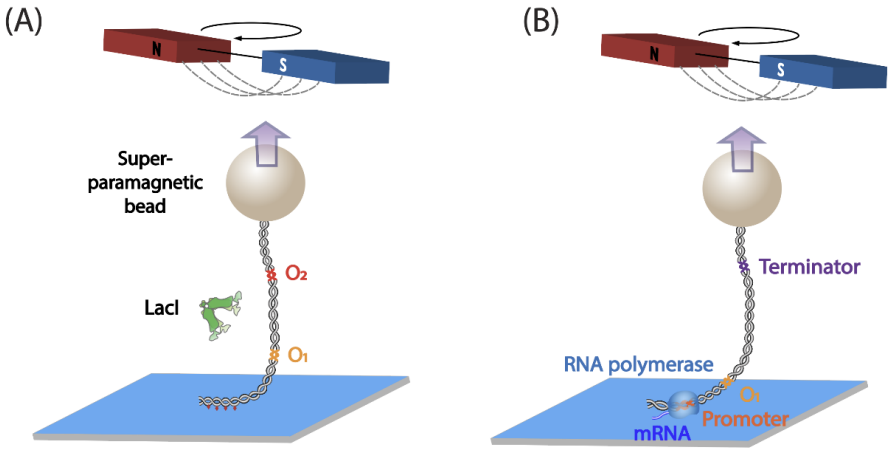


Figure 2. 9. Sketches of TPM mechanism and experimental setup. (A) General MTs setup for looping projects. (B) General MTs setup for transcription projects.

§2.5.2 MTs microscope equipment

The custom-built magnetic tweezers microscope is shown as figure 2.10. From top to bottom, it is composed of illumination unit, magnet control unit, sample stage, objective control unit and image acquisition unit. All the units are aligned on a vertical rail which stands on a vibration damped optical table. Illumination unit contains custom LED (Luxeon Star LEDs, Quadica Developments Inc. Brantford, ON, Canada), brightfield illuminator. Magnet control unit contains magnetic dipole which consisted of two 1/200x1/400X1/800 Neodymium N52 grade magnets (K&J Magnetics Inc. Pipersville, PA), spaced 1 mm apart, attached to a steel hub along with the illumination path. The steel hub was then mounted on a vertical translation stage and can be rotated using belt linkage with horizontally located motor (custom design). Sample stage is a customized aluminum plate whose XY position can be adjusted by two attached spiral micrometers. The center part of the plate is hollowed out for coverslip holding and interaction with objective. The objective control unit contains a Nikon Plan 100x/1.25 Oil immersion objective (Nikon Instruments Inc. Melville, NY), P-721 Piezo Flexure Objective Scanner (PI Physik Instrumente LP Auburn, MA) which can vertically control the position of objective. Image acquisition unit an f=160 mm tube lens (Thorlabs Inc. Newton, NJ) and a Basler acA2000-165um camera (IVS Imaging, Coppell, TX).

Real-time 3D particle tracking was implemented following a previously published scheme [71]. The XY location of each particle was tracked based on a radial symmetry detection algorithm [72]. The pixel resolution is 72.5 nm/pixel. The noise the radial symmetry algorithm can moderate image noise to confine the particles to within 5– 10% of a pixel, yielding an accuracy of around 3–7 nm. Microscope controls and 3D tracking software were written in MATLAB (Mathworks Natick, MA) and utilize Micromanager (www.micro-manager.org) to communicate with the hardware. Tracking routines and control software can be found at http://www.physics.emory.edu/faculty/finzi/research/code.shtml.

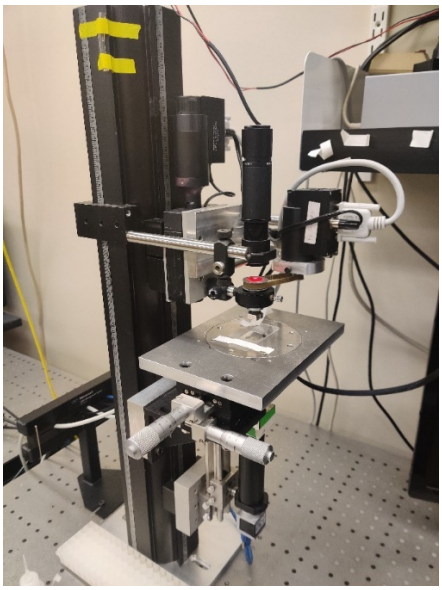


Figure 2. 10. Photograph of the magnetic tweezer microscope in the Finzi-Dunlap lab used in the experiments reported in this dissertation.

§2.5.3 Data collection and analysis of MTs

Extension-versus-time data were acquired at 164 Hz using a custom-built instrument. A 60-point moving average of the motions of beads that were stuck to the surface was used to subtract mechanical drift introduced by vibration or thermal expansion of the microscope. A 200-point moving average of the drift-corrected time-series was applied to abate the noise in each time series.

§2.5.3.1 DNA extension calibration

Z positions were determined by matching the radial profile of diffraction pattern intensity ($\hat{I}_r$) with the intensity pattern in the lookup table ($I_r[z]$) that yielded the smallest total squared difference ($arg\min[\sum_r (I_r[z] - \hat{I}_r)^2]$). Calibration experiments revealed that this scheme yielded a depth resolution of 10-20 nm. Lookup tables were built by moving the objective over a preset range of heights, the diffraction pattern of a bead at every height was recorded and piled up to form a lookup table. Figure 2.11 shows how extension was extracted from the loop up table. A reference bead, stuck to the surface of the microchamber,

is away from the focal point and showed a larger diffraction pattern than that of a bead at the focal point. In the case illustrated in figure 2.11, the pattern of the stuck bead corresponded to the pattern around 59.7 µm in the lookup table (blue dashed line) while the tethered bead which is close to the focal point showed a smaller diffraction pattern corresponding to the pattern around 56.7 µm in the lookup table (orange dashed line). Therefore, the difference in Z position between these two diffraction patterns shows the height difference between the reference bead and the tethered bead which represents the extension of the tethering DNA molecule.

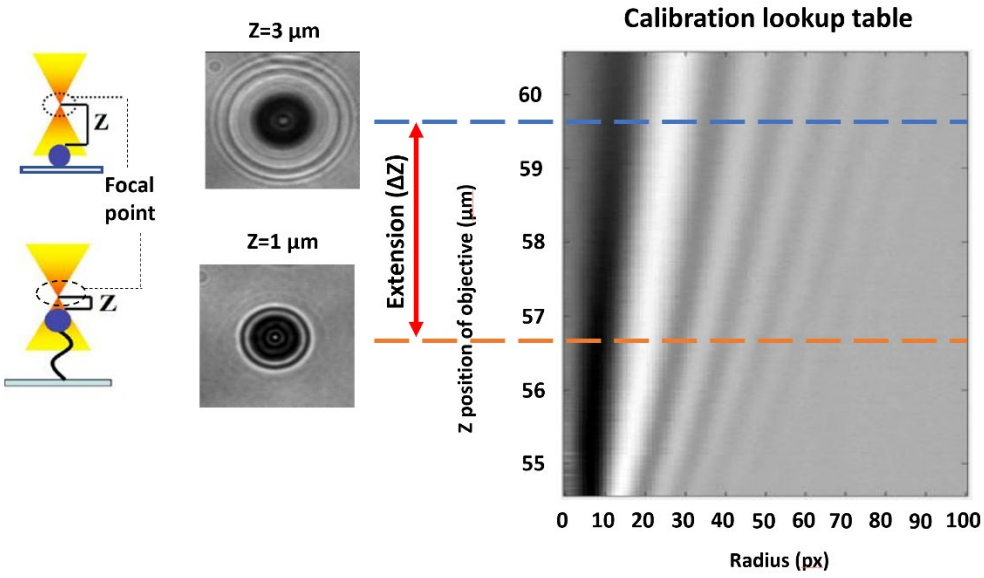


Figure 2. 11. Tether extension calibration. Cartoons on the left show the different Z values of reference bead and tethered bead relative to the focal point. Diffraction patterns in the middle show the imagine of tracked beads whose positions correspond to the cartoons on the left. Imagine on the right is the calibration lookup table which can be used to find the positions of beads according to their diffraction patterns.

§2.5.3.2 Force-magnet height curve and "Hat" curve achievement

To control the force applied to the bead/DNA tether, the force-extension curve associated with the magnet height was necessary. The average tension F can be calculated by following equation at each magnet height:

$\langle F \rangle = \frac{Lk_BT}{\langle \delta x^2 \rangle}$, Where L stands for DNA extension, k_B is Boltzmann constant, T is temperature, $<\delta x^2>$ is the mean square displacement in XY plane. Figure 2.12A shows the representative force-magnet height curve where the larger the magnet height is, the closer the magnet to the chamber. This curve enables me to achieve specific tension applying to DNA.

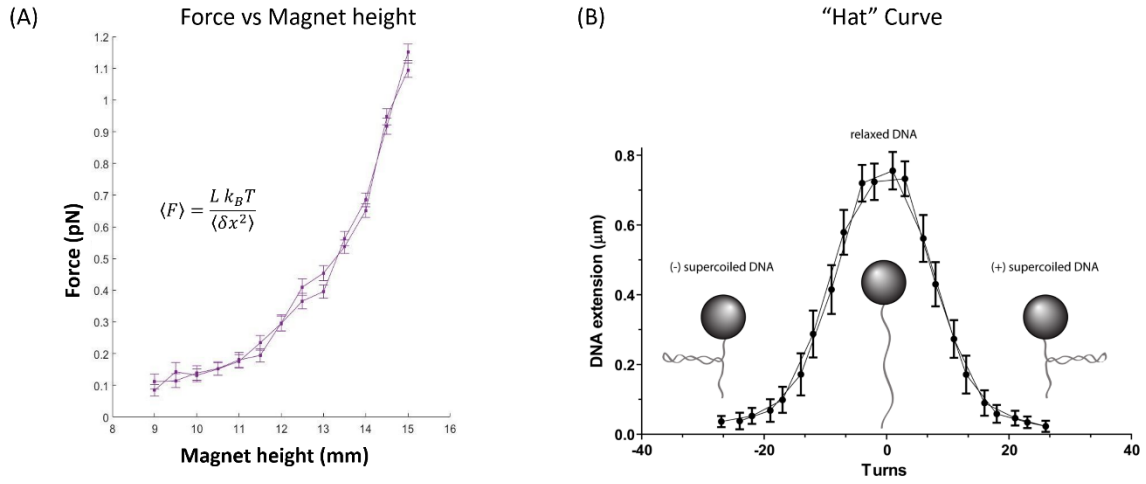


Figure 2. 12. Force vs magnet height curve and "hat" curve. (A) Force vs magnet height curve calculated using equation**:** $\boldsymbol{\langle F \rangle = \frac{Lk_BT}{\langle \delta x^2 \rangle}}$**.** (B) "Hat" curve – DNA extension vs magnet turns curve. It was generated by sequentially increase DNA supercoiling from heavily negative to heavily positive. The inserted cartoons show the plectonemes forming and relaxing during this process.

Under low tension, a torsionally constrained DNA tether can be twisted by rotating the pair of magnets. Increasing torsion induces plectonemes of increasing size. As plectonemic gyres are added in the DNA, its extension decreases monotonically and the slope of the sides of the hat curve yields the size of each plectonemic gyre. Figure 2.12B shows a representative extension vs magnet turns curve ("hat" curve). Hat curve shows the relationship between supercoiling level ($\sigma = \Delta Lk / Lk_0$ = turns/ Lk_0) and DNA extension. When multiple coilable tethers were recorded simultaneously in the same frame of view, the relaxed states

of each tether did not necessarily occur at the same rotation point of the magnet. In this case, hat curves are useful to indicate the relaxed state of each tether.

§2.6 Transcription traces and pausing time measurement

§2.6.1 Transcription on torsional relaxed DNA and no LacI-mediated looping

For transcription measurements on torsionally relaxed DNA, the single biotin label at the end far from the T7A1 promoter acted as a swivel to relax any torsion the tether accumulated during transcription (Figures 2.4B and 2.9B). Before adding NTPs, the extension of the tether was recorded for approximately 1 minute. Immediately after addition of 1 mM NTPs, turbulence lasting almost one minute produced spurious length measurements. When the turbulence subsided, many tethers returned to the previously measured extension value and shortly thereafter transcription elongation resumed, and the DNA extension decreased. When LacI bound at operator site, traces showed a clear pause with no clear extension changes for some time. These pause times were then calculated by simply picking two points located at the beginning and end of the pause, respectively. Pause duration was also determined by the scatter points that show no significant DNA extension changes comparing with the elongation process. As RNAP read through the operator, the DNA extension continues to decrease until RNAP reaches, or read through the terminator to the end of the DNA template at the flow chamber surface (Figure 2.13).

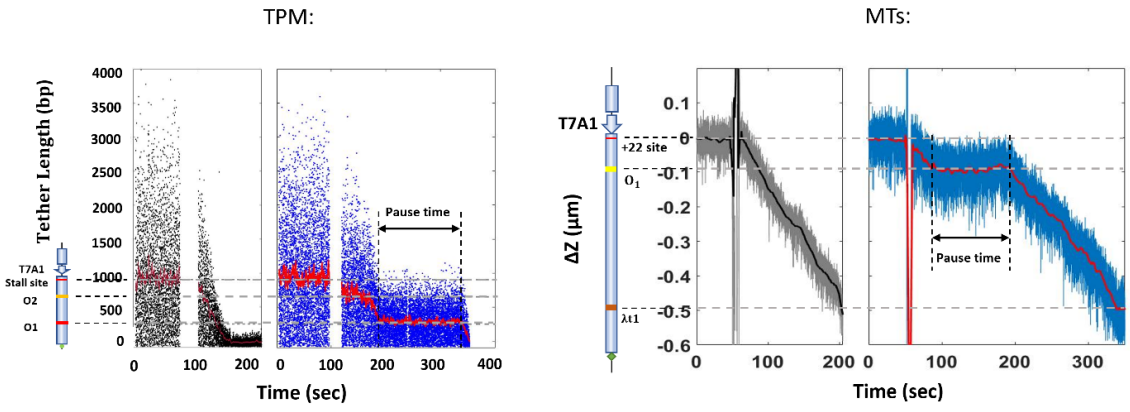


Figure 2. 13. Representative transcription traces on torsionally relaxed DNA with no loop formation. Gray traces are control experiments without addition of LacI while blue traces are with LacI which shows

clear pauses at positions corresponding to the operators. The bar cartoons are the DNA templates used in each project and show the positions of T7A1 promoter, +22 start site, operators, λt1 terminator (MTs).

§2.6.2 Transcription on torsionally-constrained DNA and no LacI-mediated looping

A DNA tail labeled with multiple biotins was ligated at the end far from the T7A1 promoter to torsionally constrain the tether during transcription (Figure 2.4B and Figure 2.9B). The experimental was designed so that as RNAP encounters O1, ΔLk = +5. Since the distance between the +22 start site and the O1 operator is in the DNA construct is 253bp, the corresponding linking number of that stretch of DNA is 253(bp) / 10.5(bp/turn) =24; Therefore, starting transcription on a DNA template with initial -19 turns will ensure ΔLk = +5 as RNAP encounters O1. The "hat" curve tells the height change from torsionally relaxed state to ΔLk = +5 which can be used to recognize RNAP pausing event at O1 operator as shown in Figure 2.14. The distance between 0 turn and +5 turn in "hat" curve tells the distance between the relaxed state and the +5 turn DNA state, facilitating recognition of a pause caused by the LacI obstacle. Note that -19 turns made the tether plectonemic and short at the beginning of the experiment. As transcription starts, positive supercoiling generated by RNAP annihilates the negative plectonemes present in the tether and causes it to extend, until the torsionally relaxed state is reached (highest extension). Further transcription adds positive plectonemes into the tether which shorten the tether again until RNAP is stopped by excessive accumulation of positive supercoiling.

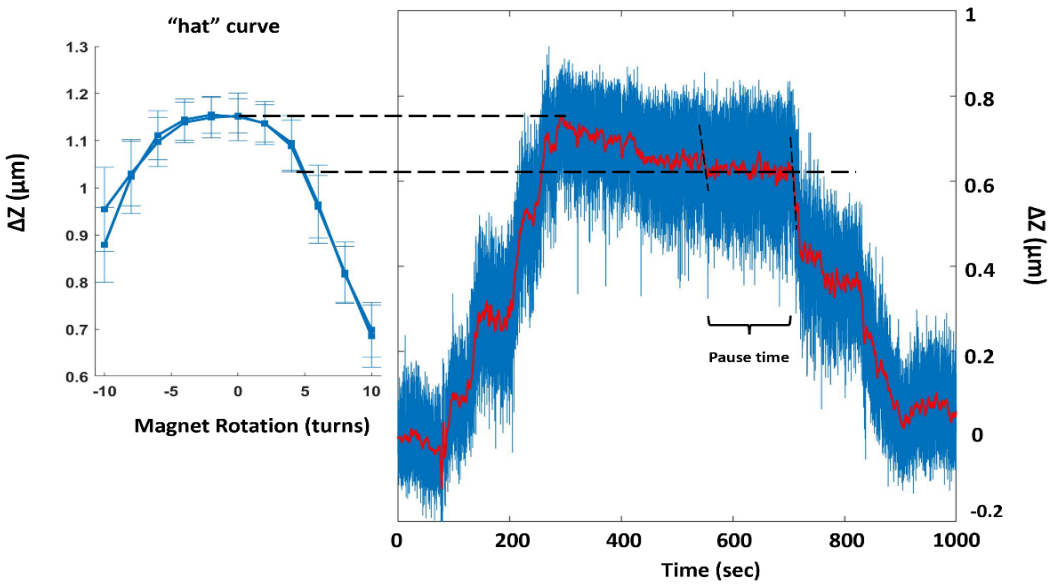


Figure 2. 14. Measurement of pause time on supercoiled DNA. (Left) "hat" curve representing the relationship between tether extension and magnet rotation. (Right) A representative trace that starts transcription after -19 turns had been introduced in the DNA template of a stalled TEC. The two horizontal, black dashed lines show that the elongation pause occurred after RNAP had introduced +5 turns in the template. Given the distance of 253 bp between the stall and the O1 site, this is expected to be exactly in front of O1: [(253bp/10.4bp/turn) – 19 turns] = ~5 turns.

§2.6.3 Transcription on torsional relaxed DNA with LacI-mediated looping

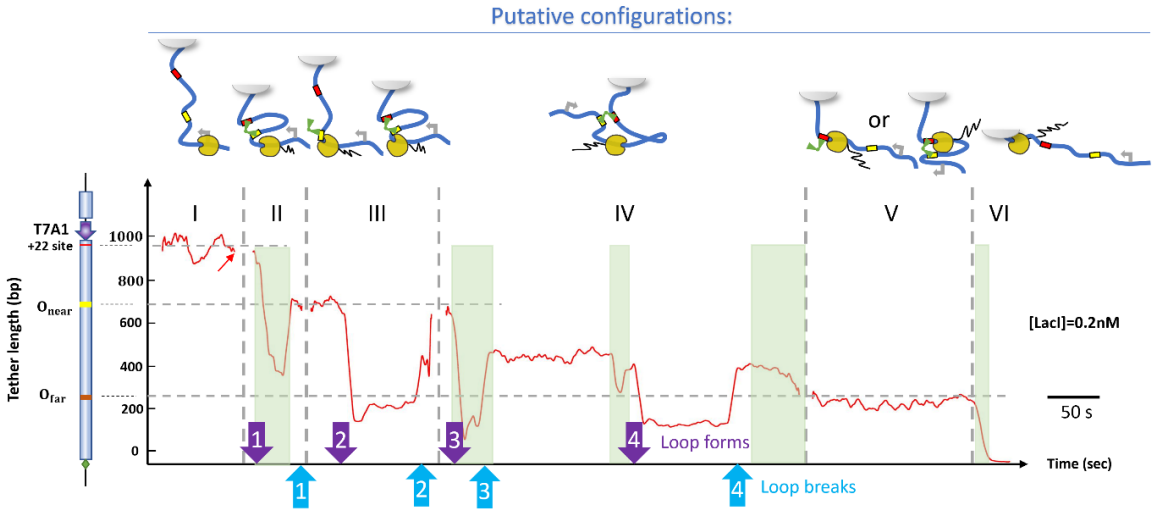


Figure 2. 15. An illustrative example which assembles segments from several traces that transcribing with loop formation. Top series of cartoon show putative configurations that correspond to different intervals in the example. The green areas indicate transcriptional activities, the purple arrows (PA) mark the time that loop forms while the blue arrows (BA) mark the time that loop breaks.

A LacI-mediated loop can produce discrete jumps or plateaus as transcription decreases the tether length. To facilitate interpretation, an illustrative example was assembled from segments drawn from different traces (Figure 2.15). Note that transcription draws the bead toward the anchor point and shortens the tether. In interval (I), the DNA tether length remains constant before introducing NTPs and LacI (red arrow). Transcription begins shortly after introducing NTPs in interval II and a loop form (PA 1), but the loop ruptures (BA 1) as RNAP approaches O_{near} and RNAP pauses at this location suggesting that LacI is bound to O_{near}. In interval (III), a loop forms (PA 2) while RNAP is paused in front of O_{near} and ruptures (BA 2) shortly thereafter, suddenly restoring the tether length to O_{near}. At the beginning of interval (IV), RNAP begins to transcribe the inter-operator region downstream of O_{near} as a loop form (PA 3). Loop breaks down (BA 3) after about 40 seconds and the new tether length corresponds to RNAP approximately in the middle of the loop segment, 200 bp from O_{near}. Then, RNAP stalls for 150 second until the loop reforms (PA 4).

Judging from the tether length at the time loop (BA 3) ruptures, RNAP has not progressed further while circumscribed by loop. Transcription resumes shortly after loop ruptures (BA 4) and then RNAP pauses at O_{far}, interval (V). This indicates that LacI is bound to O_{far}. In this region, loop formation may occur but does not significantly change the tether length and cannot be detected. In interval (VI) RNAP has passed O_{far} and transcribes to the end of the DNA template. The x lengths of (III) and (V) intervals represent the pause time at O_{near} and O_{far}, respectively. The x lengths of interval (IV) are the time RNAP spent inside the loop region.

Chapter 3 Energetics of twisted DNA topologies

§3.1 Summary

Conformational changes in DNA that result from torsional stress, modulate subsequent transactions and are essential for the regulation of DNA biochemistry. The energy changes associated with these torsion-induced conformational changes direct this regulation and are important for understanding cellular energy transactions and balance [73]. However, given the different types and length scales of conformational responses to torsion in DNA, it has been challenging to theoretically describe the energetics of torsion-driven, topological or conformational changes in DNA with a single model.

In this chapter [74] I review the main theoretical models used to calculate free energy changes associated with common, torsion-induced conformational changes in DNA and provide the resulting equations hoping to facilitate quantitative analysis of both *in vitro* and *in vivo* studies. This chapter begins with a summary of work regarding the energy change of the negative supercoiling-induced *B*- to *L*-DNA transition, followed by a discussion of the energetics associated with the transition to *Z*-form DNA. Lastly, it describes the energy changes associated with the formation of DNA curls and plectonemes, which can regulate DNA-protein interactions and promote crosstalk between distant DNA elements, respectively. The salient formulas and parameters for each scenario are summarized in table format to facilitate comparison and provide a concise, user-friendly resource.

§3.2 Free energy of torsion-induced conformational changes in DNA

DNA adopts various conformations and topologies depending on the tension and torque to which it is subjected, and the dominant conformations over a range of tension and torque have been theoretically predicted [38]. I confine my attention to the energy of DNA conformations most relevant *in vivo*, including *B*-form, melted, intertwined, *Z*-form, and plectonemic DNA (Figure. 3.1).

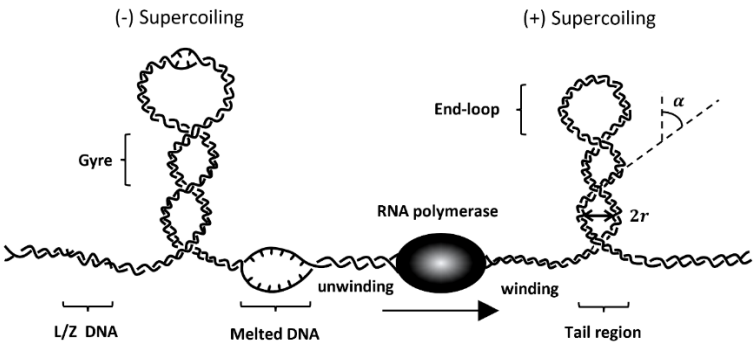


Figure 3. 1. Sketches of the twin-domain model. During transcription of a template that is torsionally constrained at each end, negative supercoiling accumulates upstream while positive supercoiling accumulates downstream. This RNA polymerase-generated supercoiling can produce plectonemes, though negative supercoiling may induce DNA melting or transition to the *L/Z*-form before plectonemes are formed at the buckling transition. Transcription is not the only supercoiling-generating process in the cell, but it is used here as an example.

§3.2.1 Energies of DNA melting and of the *B-Z* transition

As *B*-DNA is progressively unwound, the magnitude of (-) torque increases. The *B-L* transition is triggered beyond a threshold of negative torque and continues until the entire molecule adopts the *L*-form. In this transition, if tension is high enough (but still physiological, $f < 20$ pN) to prevent buckling of a negatively twisted DNA molecule, base pairs will start melting when the torque reaches about -10 pN·nm. With further negative twist, the torque remains constant as *B*-DNA transitions to *L*-DNA. Since *L*-DNA has a much higher twist persistence length ($C_L \in$ (10-20) nm) [75] than intertwined, melted DNA ($C_{melt} \approx$ 1nm, Eq. 1), *L*-DNA is believed to be composed of intertwined, melted DNA and other non-canonical structures like *Z*-DNA [76]. Previous models have shown that less energy is required to denature an A:T base pair than a C:G base pair, $\Delta G \approx 1.9$ and 2.9 k_BT, respectively [77], so DNA melting occurs preferentially in AT-rich regions. Conversely, the B-Z transition is most common in DNA regions containing dinucleotide repeats of d(pCpG) or d(pGpT)[78]. Both melting and the *B-Z* transition are highly sensitive to increases of temperature and pH, which destabilize hydrogen bonding in base pairs, and salt concentration, which reduces repulsions between

the sugar phosphate backbones to stabilize base pairing. This discussion is limited to the most common experimental conditions, namely room temperature (25 °C), pH 7.5-7.7, and salt concentrations from 10 to 400 mM.

The energy necessary to melt *B*-DNA can be divided into two contributions (E_{bpM} in Eq. 1, Table 1): (1) The energy required to melt an initial base pair which depends on the base pair type and averages $\mathcal{E}_{ini,melt} \approx$ 9-11 k_BT [79] , and (2) the energy associated with melting successive base pairs, which depends on each base pair and its neighbors [77, 80, 81]. After melting, the twist energy of the melted region changes as torsional stress within what was previously dsDNA is sustained by intertwined, but unpaired, single strands and can be expressed as E_{twM} in Eq. 1 in Table 1. The twist persistence length, C_{melt}, varies from approximately 1 to 3 nm as [Na^+] drops from 400 mM to 10 mM [80, 81], as expected since the ionic screening of electrostatic repulsions falls with the salt concentration. C_{melt} for (AAT) repeats can even exceed 15 nm in Tris buffer with no salt, which indicates that like *B*-form DNA, the twist rigidity of intertwined, melted DNA increases as monovalent cation concentrations decrease. For non-hybridized, intertwined strands, the average twist angle of each disrupted base pair, $\bar{\theta}$, is determined by minimizing the overall energy [82]. An intertwined, melted region contributes $\Delta Lk_{melted} = (n\bar{\theta})/2\pi$ to the linking number, where n is the number of denatured base pairs after loss of n/10.5 turns due to melting. If the overall linking number of the molecule is fixed, the linking number change of DNA segments surrounding the melted region will be $-\Delta Lk_{melted}$, [82] which can be plugged into Eqs. *3*, *4*, or *5* in Table 1 (depending on the model used; see section ii) to calculate the twist energy of the flanking *B*-DNA.

Similar to melting, the free energy for the *B*-*Z* transition can be calculated considering two contributions (E_{bpZ} in Eq. 2, Table 1). One is the energy of disrupted base stacking including the "domain wall energy", $\mathcal{E}_{wall}$, of ~8.4 k_BT at each junction between *B*- and *Z*-form segments [76, 83, 84]. The other is the energy associated with disrupting each additional base pair, following the first one, which depends on the nature of the base pair, for example, ~1.1 k_BT/bp, for d(pCpG) repeats, ~2.4 k_BT/bp, for d(pCpA) repeats [83, 84]. After a *B*-*Z* transition has initiated, considering only d(pCpG) repeats with m base pairs, in a conventional zipper model

[84], the change in twist in going from *B*-form (10.5 bp/turn) to *Z*-form (-12 bp/turn) is $\Delta Lk_{B\text{-}Z} = -am - 2b$, where a is the change in twist for a single *GC* base pair flipped from *B* to *Z* and is given by: $a = 2\cdot(^1/_{10.5} + ^1/_{12})$, while $b = 0.4$, is the twist at the boundary between the *B*- and *Z*-forms. Such a change in twist consequently alters the linking number of flanking DNA and its twist energy as well. The twist energy of *Z*-DNA is expressed as E_{twZ} in Eq. 2, Table 1, where the twist persistence length, C_z, is determined to be approximately 7 nm in Tris buffer without salt[85].

DNA melting and the *B*-*Z* transition alter the DNA conformation over a long-range and may affect the DNA-binding energy of proteins. For example, RNA polymerase binds more readily if the DNA is partially melted, and there is a lowered energy barrier for strand separation. Similarly, *Z*-form DNA is the binding target of certain proteins, such as human ADAR1, a prototypic *Z*-DNA binding protein (ZBP), which otherwise must induce *Z*-form DNA upon binding [86]. Knowledge of the energetics of different DNA conformations facilitates a quantitative understanding of how DNA topological changes regulate protein binding and enzyme activity.

§3.2.2 Energies of extended, curled, and plectonemic DNA

This section focuses on the buckling transition and the associated energy change. Studies of the buckling transition have often focused on positively supercoiled DNA, where DNA remains in the *B*-form under tension up to 10 pN and torsion up to 40 pN·nm [38].

Initially, twisting a DNA molecule under tension produces an extended conformation (Figure 1.3). One of the simplest models for this is the twistable worm-like chain (TWLC) (Table 1, Eqs. *3* and *6*), which treats DNA as an inextensible rod with independent flexural and torsional elasticities [87]. However, given the intrinsic helicity of DNA, a twist-bend coupling parameter, G, improves the accuracy of theoretical descriptions [88, 89].

In the absence of tension, $f=0$, and externally applied torsion, $\tau=0$, when taking into consideration the coupling that must exist between helical DNA bending and the consequent twist, the intrinsic bending

stiffness, A, and twisting stiffness, C, should be calculated as the renormalized, G-dependent values κ_b, and κ_t, respectively [90]. Then, κ_b, κ_t can be used to calculate the tension-dependent ($f > 0$) effective twisting stiffness, C_{eff} (Table 1, Eq. 3). This was found to be in excellent agreement with the one derived from a coarse-grained model of DNA, which used improved conformational parameters and salt dependence, and in reasonable agreement with experiments, especially at low tension ($f \leq 1$ pN) [91]. For small values of G, Nomidis *et al.* [91] proposed a calculation of the free energy of extended DNA under both tension and externally applied torsion based on perturbation theory (Table 1, Eqs. *4* and *7*). Their approach identified a very large characteristic force, $f_0 \approx 600$ pN, above which C_{eff} finally approaches the intrinsic value C. For almost all single-molecule measurements on DNA with mechanical properties under physiologically relevant conditions, the forces at play are a lot smaller than f_0, making it justifiable and accurate to use C_{eff}, instead of C.

If the extensibility of DNA is taken into consideration, instead of using the TWLC model, the chiral, extensible worm-like chain (CEWLC) model can be used, which takes into consideration the coupling between twist and stretch, expressed as the coupling parameter g [92]. Using this model, the free energy of a DNA molecule under both tension and torsional stress was formulated by Marko [92] (Table 1, Eq. *8*) and used to fit experimental data and determine the value of g [93, 94]. Under low tension, the extension of double-stranded DNA increases upon winding and a negative g characterizes this behavior. DNA has a narrower minor groove, but a positive base pair inclination, which leads to a slenderer and more extended double helix than that of RNA [95]. Under the same conditions, double-stranded RNA is characterized by a positive g [94, 96], with a more compact conformation, a narrower major groove, and reduced helical extension compared to DNA.

Either the TWLC or CEWLC, with G or g, respectively, $\neq 0$, are still simplified models of DNA mechanics under torsion, since the bending and twisting stiffness are DNA sequence dependent [97, 98]. Both the TWLC and CEWLC models have been used to successfully predict the torsional and bending stiffness of DNA [91, 92]. However, combination of these two models into one that couples G and g would provide a unified model

that allows simultaneous consideration of both the asymmetric nature and extensibility of double-stranded DNA. More rigorous constitutive theoretical frameworks have been developed for the short-range mechanical behavior of DNA as a function of sequence [32], defects [34-36], and distortions [37]. These models have proved useful to predict, for example, sequence-dependent local DNA topologies, such as the positioning of nucleosomes [99, 100], or the position of DNA supercoils [32], based on energy cost of deforming specific DNA sequences.

When the torque on a DNA molecule exceeds a critical value, τ', accumulated strain will trigger buckling [23] and plectoneme nucleation which abruptly shrinks the overall extension. This buckling transition is first characterized by a "soliton" transition state [101, 102] with energy given by Eq. *9* in Table 1. After a soliton is formed, a "curled" intermediate state [30, 102] is the following step towards the formation of plectonemes (Figure 1.3). Near the buckling transition, the extension of DNA molecules toggles between two distinct values, the frequencies of which display the probabilities of these states [103]. The critical value of linking number at the buckling transition, which corresponds to the torsion that causes the curled and extended DNA states to be equally probable, increases with the applied tension and molecule length [104]. The analytical formula is available [30], and the critical torque, τ', can be easily calculated by using Eq. *7* (Table 1).

Theoretical expressions for the free energy of stretched-twisted DNA (Table 1. Eq. 6), curls (Table 1. Eq. *10*), and plectonemic DNA (Table 1. Eq. 11) describe the partitioning of energy associated with a single DNA molecule [30, 105]. These expressions can be used to calculate the number of curls, plectonemes, the size of plectonemic gyres, the probability of each conformation, and the average extension of DNA as a function of tension, DNA length, and salt concentration.

For salt concentrations lower than 50 mM, the persistence length of DNA increases due to reduced electrostatic screening. This contributes to the electrostatic repulsion energy term, $U(r, \alpha)$ in Eq. *11* of Table 1, and the number of curls and plectonemes along DNA decreases. Even under the most common experimental conditions in which $f > 0.25$ pN and the monovalent salt concentration is in the range of 100

– 200 mM, a buckling transition rarely involves more than one plectoneme [30] due to the high energy required to curl the DNA and repulsion between the negatively charged helices. As might be expected since tension opposes bending fluctuations and reduces entropy, the DNA length per plectonemic gyre decreases as tension f increases. However, under very low forces ($f < 0.25$ pN), which negligibly penalize the formation of curls and plectonemic end-loops, multiple plectonemes may occur [30].

In vivo, distortions or sequence-dependent architecture of the DNA as well as protein binding may favor the formation of multiple plectonemes or curls along DNA. Instead of plectonemes randomly nucleating along DNA, the DNA sequence can favor plectonemes and curls in particular positions due to the intrinsic curvature or deformability of specific DNA sequences, although the AT versus GC content *per se* is not necessarily a determinant for plectoneme formation [32]. In addition, local defects induced by DNA mismatch, or damage, can also spur the formation and localization of plectonemes, which suggests that plectoneme formation might be a mechanism for locating the site of lesions requiring repair *in vivo* [36, 102]. Epigenetically, proteins that bind to and stabilize curls might lower the energy required to induce curls at multiple protein binding sites [30, 99, 100]. In turn, mild torsion generated by motor enzymes may facilitate transcriptional initiation while highly plectonemic regions might stall transcriptional elongation.

The nucleation energy largely determines whether specific DNA sequences with mismatches, kinks, or other deformities pin plectonemes, and various models with which to calculate it have been proposed [32, 34, 35, 37, 106]. Most begin with the bending or twisting energy of DNA and introduce additional mechanical features. Given that DNA buckling and plectonemes directly affect DNA long-range interactions, it is important to understand how the sequence and tension within a DNA segment, as well as the surrounding salt concentration, affect the free energies of these topologies. This improves my understanding of DNA dynamics and regulation *in vivo*.

§3.3 Table of energy equations

Common symbols:

L_s: Stretched DNA contour length; C_{eff}: Effective twist persistence length (twist modulus/k_BT); ΔLk: linking number change in stretched *B*-form DNA; f: tension; τ: torque; ω_0: intrinsic twist of DNA, $\omega_0 = \frac{2\pi}{3.57}$ nm^{-1}; G: twist-bend coupling parameter (G : 30 -- 40 nm [90, 91]); g: twist-stretch coupling parameter (unitless) (g : -22 ~ -17 [93, 94, 96]); A: bend persistence length; C: twist persistence length.

Energy type	Equations	Annotations		Ref. #	Eqn. #
Energy of intertwined, strand separated DNA	$E_{Melted} = E_{twM} + E_{bpM}$ E_{twM}: Energy of intertwined, melted region: $E_{twM} = \frac{1}{2} n\, \kappa_{melt} \bar{\theta}^2 = \frac{k_B T C_{melt}}{2L_{melt}} (2\pi \Delta Lk_{melted})^2$ E_{bpM}: Energy cost for denature base pairs in a melted region: $E_{bpM} = \varepsilon_{ini,melt} + \sum_{i=1}^{n} \varepsilon_{melt,i}$	n: number of unpaired base pairs. $\bar{\theta}$: average twist angle of each disrupted base pair. Length of melted region: $L_{melt} \approx n * 0.54nm/bp$ [75] , $2\pi \Delta Lk_{melted} = n\bar{\theta}$. κ_{melt}: twist modulus of two intertwined strands, ~$2.3 k_BT$. Twist persistence length: $C_{melt} = \frac{L_{melt} * \kappa_{melt}}{n k_B T} \approx 1.2nm$. $\varepsilon_{ini,melt}$: initial energy for melting: $9 - 11\ k_BT$. $\sum_{i=1}^{n} \varepsilon_{melt,i}$: sum of energy cost for each individual base pair melting, where $\varepsilon_{melt,i}$ depend on each base pair and its neighbors, particular values can be found [77].	Change of linking number that melted region contributes to flanking DNA: $\Delta Lk_{B-melt} = \frac{n}{10.5} - \frac{n\bar{\theta}}{2\pi}$	[75, 77, 79-82, 85, 107]	(1)
Energy of *Z*-form DNA (fixed two boundaries)	$E_Z = E_{twZ} + E_{bpZ}$ E_{twZ}: Energy of twisting *Z*-DNA: $E_{twZ} = \frac{k_B T C_Z}{2L_Z} (2\pi \Delta Lk_Z)^2$ E_{bpM}: Energy cost for B-Z transition: $E_{bpM} = 2\varepsilon_{wall} + \sum_{i=1}^{m} \varepsilon_{B-Z,i}$	L_Z: Length of *Z*-form DNA. ΔLk_Z: twist change in *Z*-form DNA. C_Z: Twist persistence length, $\sim 7nm$ for $\mathrm{d(pCpG)_n}$ repeats. ε_{wall}: domain wall energy, ~$8.4\ k_BT$. m: number of base pairs that undergo B-Z transition, $L_Z = m * 0.37nm/bp$ [108]. $\sum_{i=1}^{m} \varepsilon_{B-Z,i}$: sum of energy cost for disrupting each additional base pair, following the nucleus of *Z*-DNA, where $\varepsilon_{B-Z,i} \approx 1.1\ k_BT$ for d(pCpG) repeats, $\varepsilon_{B-Z,i} \approx 2.4\ k_BT$ for d(pCpA) repeats [83, 84].	Change of linking number that B-Z transition contributes to flanking DNA: $\Delta Lk_{B-Z} = -am - 2b$, $a = 2\left(\frac{1}{10.5} + \frac{1}{12}\right)$, $b = 0.4$.	[83-85, 108, 109]	(2)
Twist energy of stretched *B*-form DNA	$E_{twB} = \frac{k_B T C_{eff}}{2L_s} (2\pi \Delta Lk_B)^2$	$\frac{1}{C_{eff}} = \frac{1}{C^*} + \frac{1}{4A^*} \sqrt{\frac{k_B T}{f A^*}}$	TWLC model (G=0): $A^* = A\ (\approx 50\text{nm})$ $C^* = C\ (\approx 100\text{nm})$	[105]	(3)
			TWLC (G≠0 & very large G) Non-perturbation theory: $A^* = \kappa_b = A \frac{1 - \frac{\varepsilon^2}{A^2} - \frac{G^2}{AC}(1 + \frac{\varepsilon}{A})}{1 - \frac{G^2}{2AC}}$ $C^* = \kappa_t = C \frac{1 - \frac{\varepsilon}{A} - \frac{G^2}{AC}}{1 - \frac{\varepsilon}{A}}$ ε: the bending anisotropy.	[90]	(4)
			TWLC (G≠0&G is small) Perturbation theory ($f \in (0.1pN, 10pN)$: $\frac{1}{A^*} = \frac{1}{A}(1 + \frac{G^2}{2AC})$ $\frac{1}{C^*} = \frac{1}{C}(1 + \frac{G^2}{AC})$	[91]	(5)

Stretched DNA, TWLC model (G=0)	$E(f,\Delta Lk) = [-f + \sqrt{\frac{fk_BT}{A}} + \frac{C_{eff}k_BT}{2}\left(\frac{2\pi\Delta Lk}{L_s}\right)^2]L_s$	Definitions of $f, A, C_{eff}, \Delta Lk, L_s$ see above.		[105]	(6)
Stretched-twisted DNA, (perturbation theory) (G≠0)	$E(f,\tau) \approx \left(-f + \sqrt{\frac{fk_BT}{A^*}} + \Gamma\tau - \frac{\tau^2}{2k_BTC_{eff}}\right)L_s$	Torque:$\tau \approx \frac{2\pi k_BTC_{eff}}{L_s}\Delta Lk$	Proportionality constant Γ. $\Gamma = \frac{G^2}{8A^2C^2\omega_0}$	[91]	(7)
Stretched-twisted DNA, (CTWLC model)	$\frac{E}{k_BTL_s} = \frac{1}{A}\sqrt{\frac{Af}{k_BT} - \frac{1}{4}(C_{eff}\omega_0\sigma + \frac{gf}{K_0})^2} + \frac{C_{eff}}{2}\omega_0^2\sigma^2 - \frac{f}{k_BT} - \frac{k_BT}{2K_0}(\frac{f}{k_BT} - g\omega_0\sigma)^2$	σ: supercoiling level $\sigma = \frac{\Delta Lk}{Lk_0}$	K_0: the stretch modulus. $K_0 \approx 1200pN$ $f, A, C_{eff}, \Delta Lk, L_s, g, \omega_0$ see above	[92, 94]	(8)
Energy barrier of DNA from unbuckled to soliton state	$E_s = \frac{8k_BTA}{l}\tanh\left(\frac{L}{2l}\right) - 2\pi Wr_s(\tau + \frac{\pi k_BTC_{eff}Wr_s}{L})^2$	L: DNA length in nm.	l: soliton length scale: $[\left(\frac{A}{f}\right)^{-1} - \left(\frac{2C_{eff}}{f}\right)^{-2}]^{-\frac{1}{2}}$	Wr_s: writhe in soliton: $\frac{2}{\pi}tan^{-1}[\frac{2A}{\tau l}\tanh(\frac{L}{2l})]$	[101, 102]	(9)
Curled DNA	$E_c = (8 - \frac{3.14f(0.8 + 2.2\kappa_D^{-1})^2}{Ak_BT})\sqrt{k_BTAf}$	κ_D^{-1}: Debye length, f, A see above.			[30]	(10)
Plectonemic DNA	$E_p = \frac{2\pi^2Ck_BT}{L_p + q\Gamma}\left(\Delta Lk_p - Wr_p\right)^2 + L_p\left[\frac{Ak_BT(sin\alpha)^4}{2r^2} + U(r,\alpha)\right] + q\varepsilon_p\sqrt{Afk_BT}$	q: number of plectonemes along DNA, L_p: plectoneme length, ΔLk_p: linking number change of plectoneme, Wr_p: total writhe of the plectonemic regions: $Wr_p = \frac{L_psin2\alpha}{4\pi r} + q\omega_p$ ($\omega_p \approx 1$), $U(r,\alpha)$: electrostatic repulsion and entropic confinement free energy, Γ: length of the end-loop and tail region of a plectoneme (Figure. 3.1), more details please refer to [30]. --Top part of left expression represents the twist energy of plectoneme. --The middle term in the sum represents bending energy as well as the electrostatic repulsion and entropic confinement free energy (whole plectonemes except end-loop and tail region). --Bottom part contains the energy contribution of end-loop and tail region			[30]	(11)

Table 1. Energy expressions for DNA under different conditions of tension and torsion. Some variable names have been changed and values/units have been converted from those used in the original reports for uniformity.

§3.4 Remarks

As described in Figure 3.2, the interplay between DNA processing, DNA conformational and topological changes, and regulation via protein binding is tuned by a complex network of energy contributions. *In vivo*, DNA conformations are dynamic due to protein binding and processing by such enzymes as polymerases and helicases which can generate tension and torque. In turn, different DNA conformations affect protein/enzyme-DNA interactions. Highly supercoiled regions stall transcription while transcription-generated positive supercoiling destabilizes nucleosomes and other protein roadblocks [110] to facilitate elongation. Previous experiments have shown that elongation by RNAP will stall if the torque in the DNA template exceeds approximately 11 pN·nm [111], and topoisomerases are thought to relax excessive torsion to sustain transcription [112-114]. Transcription-coupled DNA supercoiling has also been shown to enable multiple RNAPs moving in the same direction to elongate faster than a single RNAP, since the supercoiling emanating from one reduces the supercoiling produced by others ahead or behind. Likewise, repression of an upstream promoter antagonizes downstream transcription and can cause downstream RNAPs to prematurely dissociate [115, 116]. Interestingly, the torque at which transcription stalls in negatively supercoiled DNA is similar to the torque that facilitates the separation of duplex DNA into single strands [109]. Backtracking on an unwound upstream DNA may be more favorable than on canonical *B*-form DNA, thereby stalling RNA polymerases. Furthermore, melted regions resulting from unwound upstream DNA may favor the formation of *R*-loops that might stall transcriptional elongation. For example, in the human *c-myc* promoter, transcription-induced torsional stress can melt the distant upstream element (FUSE) [117], and elongation from a promoter may recruit additional RNAPs to the same or further upstream promoters [116, 118]. Negative supercoiling also favors the binding of other transcriptionally relevant proteins such as TATA binding proteins [119].

Knowledge of the free energy of different DNA conformations and topologies provides a quantitative basis for understanding the relationship between genomic structure and function. It also provides insight into the likelihood of architectural/topological and conformational rearrangements that, for example, modulate

DNA transactions by inducing transitions between *B*- and *Z*-form DNA [41, 42], by juxtaposing distant sites for regulatory proteins, or by melting DNA to favor the recruitment of enzymes to particular sites.

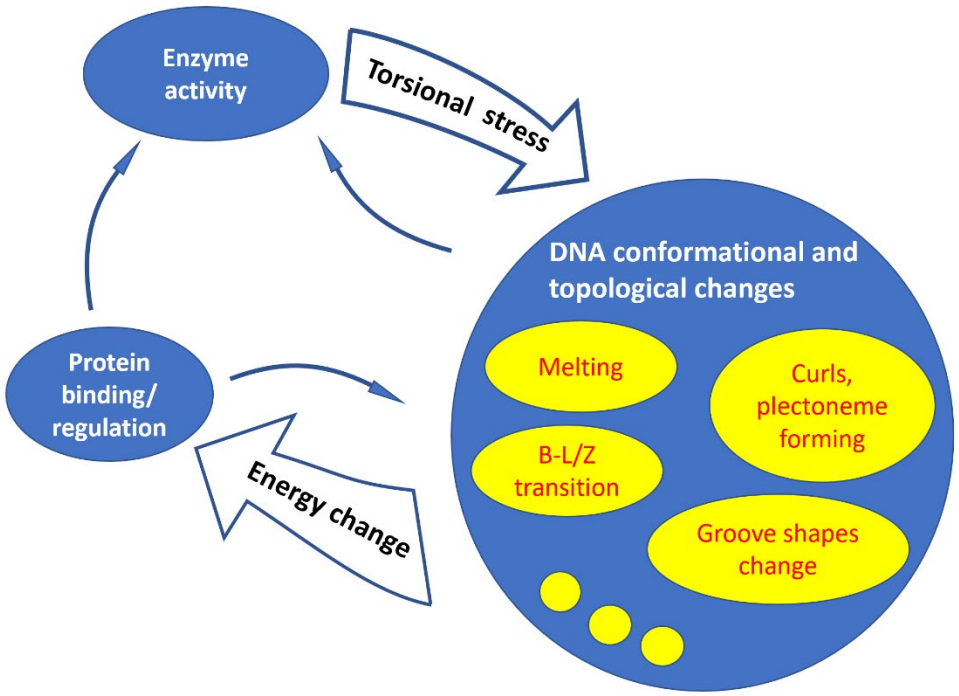


Figure 3. 2. Schematic representation of the interplay between torsion-generating DNA transactions, like transcription, conformational and topological changes in DNA, and protein binding. DNA enzyme-induced torsional stress causes local and long-range DNA conformational changes that include DNA melting, *B*-to-*L/Z* transition, major and minor groove changes, curls, and plectoneme formation. This affects protein binding. For example, plectonemes facilitate the juxtaposition of binding sites for looping proteins. Also, DNA bending, kinking, melting or *B-L/Z* transition can affect the energy barrier required for specific DNA binding proteins, or motors, to process DNA.

§3.5 Discussion and Conclusions

In this chapter, I reviewed the expressions with which to estimate energies of different DNA conformations under specific conditions of tension, torque, salt concentration, pH, and temperature. The formation of

melted DNA, *L/Z*-DNA or plectonemic domains follow a quite similar pattern: first nucleation, which requires relatively high energy to overcome a barrier, followed by progressive expansion of: denatured base pairs within a bubble, *L/Z*-form DNA between *B*-*L/Z* junctions, or more gyres in a plectoneme that require relatively less energy. Comparison of the energies associated with nucleation and expansion of these different DNA topologies under different tension, torsion, salt conditions will give insight into their likelihood *in vivo*.

Note that these expressions do not consider DNA sequences or other DNA defects. However, the mechanical properties of individual base pair steps have been theoretically assessed [120-122], and they have been used in the calculation of the energy for specific sequences to show that intrinsic curvature lowers the energy required to bend and effectively pin plectonemes [32]. Similarly, DNA defects, such as mismatches and kinks, decrease bending energy which can be predicted using available models [79-82]. By incorporating sequence-specific parameters into the general energy expressions in Table 1, I might parse the interactions that dictate protein binding and the syntax of gene expression. Knowing the energy of mismatched, kinked, or melted DNA sequences could be an incisive tool to define a sequence or a mismatch to pin a plectoneme. Calibrating these features according to tension in the DNA and salt concentrations will establish the energy required for long-range interactions.

I focused on melted DNA, *Z*-DNA and plectonemic *B*-DNA which are ubiquitous *in vivo*. Melted DNA plays a critical role in regulating gene expression by recruiting RNA polymerases or altering elongation. Tools with which to predict melting energies may help indentify potential promoters [122, 123] and indicate DNA melting behind polymerases. Moreover, since melted DNA regions are potential targets for a variety of DNA-binding proteins, learning to control DNA topology with sequence variations would improve my ability to maniuplate gene expression.

DNA melting not only makes bases accessible but also relieves torsional stress in the flanking DNA. The transitions to left-handed forms do this as well and appear to be associated with disease. *Z*-DNA is implicated in diseases of genomic instability and immune responses including cancer and systemic lupus

erythematosus [124]. Indeed, the level of *Z*-DNA assayed with *Z*-DNA antibodies reflected different disease stages of lupus [125]. The ability to predict the energies of *Z*-DNA formation along a sequence might allow specific removal/placement of *Z*-DNA to treat such diseases. For example, *Z*-DNA near promoters favors mutagenesis [126] perhaps due to mechanical constraints that interrupt transcription. The mechanical properties of *Z*-DNA also affect nucleosome organization. *Z*-DNA is stiffer than *B*-DNA or melted DNA [75] and does not easily wrap histone octamers, so the ability to predict *Z*-DNA formation would help identify nucleosomal regions [127].

Of course, there is still much to learn about how DNA topology influences protein binding and function. For example, how (un)winding affects the site-specific protein binding to DNA is unclear, especially at the level of the interactions between protein binding domains and the DNA backbone and base pairs. The suggestion, from unconstrained MD simulations [8], that local DNA twist may change the local bendability of DNA and the shape of the major and minor grooves, in such a way as to affect protein binding, has yet to be experimentally demonstrated. Knowledge of how protein binding changes DNA structure will improve prediction of DNA topology in complex mixtures using the approaches described herein.

In summary, the effect of tension and torsion on DNA conformation has, at this point, been under investigation for about three decades and several general models exist to estimate the free energy of each of these conformations. These are valuable tools with which to begin to quantitively describe the free energy changes of DNA associated with protein-DNA interactions and predict regulatory behavior. Further computational, theoretical, and experimental work could help: improve my understanding of how torsionally induced DNA conformations affect protein binding; determine the *in vivo* likelihood of mismatched, kinked, or melted DNA topologies under different tension, torsion, salt conditions; and predict sequences likely to form left-handed conformations including *Z*-DNA which appears to be an important epigenetic signal [128].

Chapter 4 Negative DNA supercoiling makes protein-mediated looping deterministic and ergodic within the bacterial doubling time

§4.1 Summary

In this chapter [129] I describe how looping probability is affected by Lac repressor (LacI), HU concentration and negative supercoiling level to show that that negative supercoiling makes LacI-mediated looping deterministic and ergodic within the bacterial doubling time. I used TPM to measure the looping in a DNA fragment with two Lac repressor binding sites, O1 and O2, that form a 400 bp loop. Adjusting the concentration of LacI changes the average looping probability and increasing the concentration of HU at a fixed LacI concentration will also increase the average looping probability. Although the average looping probabilities displayed trends as a function of conditions, individual traces in each condition showed heterogeneous looping behavior with looping probabilities that varied from 0% to 100% unless observations lasted much longer than the doubling time of *E. coli* bacteria. Such heterogeniety in molecular switches might not be advantageous for live bacteria. However, when I used MTs to negatively supercoil DNA, I observed quite homogeneous behavior in the ensemble of molecules during observations commensurate with the doubling time of the relevant bacterium.

§4.2 Supercoiling makes protein-mediated looping deterministic

§4.2.1 Looping probabilities of different tethers varies widely

As shown in the cartoon in Figure 4.1, I used a DNA template containing two LacI binding sites. At low LacI concentrations, neither operator is likely to be occupied by a LacI tetramer, and the looping probability is low, so the DNA tether remains extended, and the attached bead exhibits large excursions. When the concentration is high, both operators become occupied by distinct tetramers which cannot bind to each other to form loops. Excursions are large for beads attached to these tethers as well. Only at intermediate concentrations, in which one tetramer may bridge two operators, does the probability of looping increase significantly. When loops form, the tether is less extended, and the excursions of the attached bead are restricted. Average looping probability, calculated as the time spent in the looped state over the total

observation time, is indicated by crosses in Figure 4.1, which summarizes ~30 min-long TPM [63, 64, 130] measurements of LacI-mediated DNA looping between the strong O1 and weaker O2 operators separated by 400 bp [29].

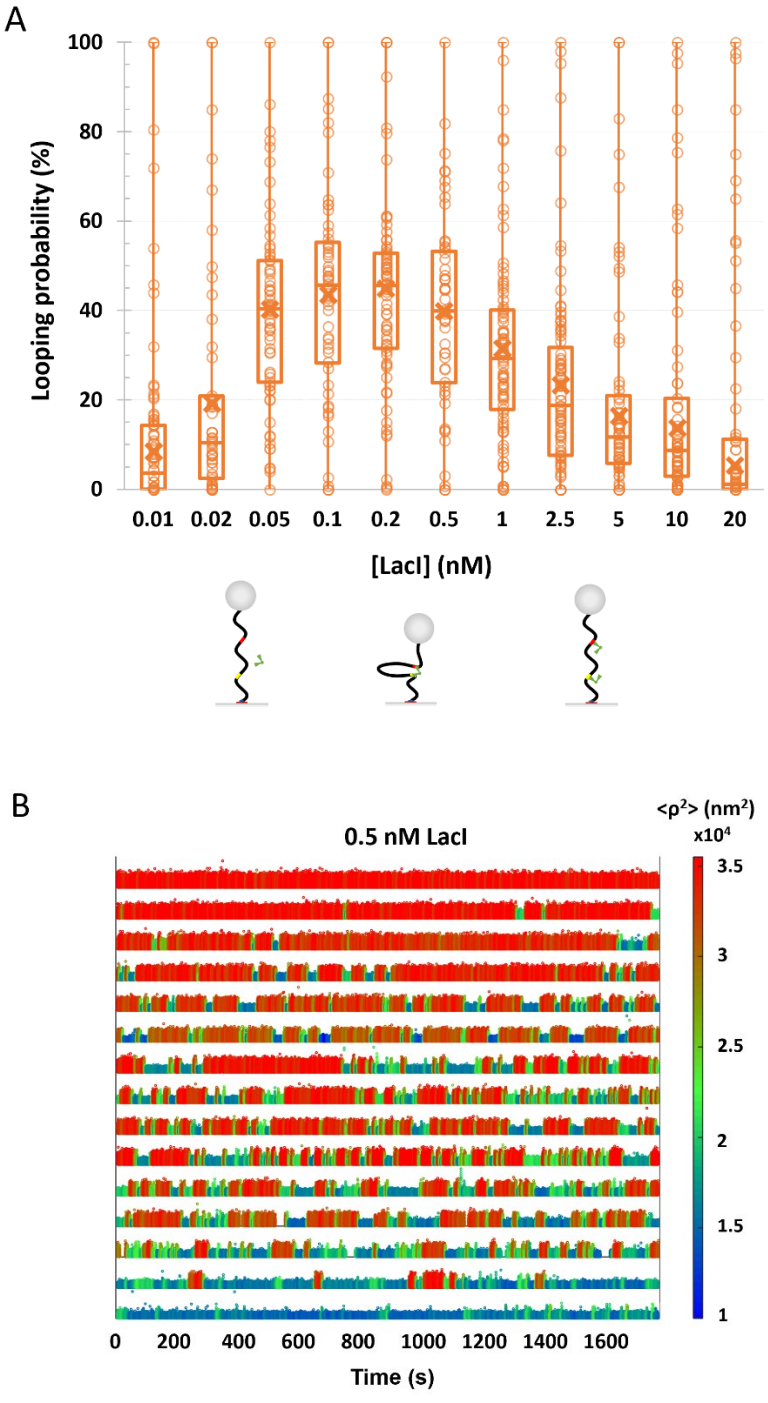


Figure 4. 1. The looping probabilities of different DNA tethers vary widely. (A) The calculated looping probabilities of individual tethers exposed to a range of LacI concentrations are summarized in a box-whisker plot. As the LacI concentration was titrated from 0 to 20 nM, the looping probability increased from 0 to approximately 45 and then fell again to 0 when repressor molecules saturated the binding sites. Note however, that at each concentration of LacI, the looping probabilities based on 30-minute observations varied from 0 to 100%. The upper and lower borders of the boxes indicate the upper and lower quartiles. The midline and cross of each box indicate the median and average of the distribution of looping probabilities. Schematic diagrams of prevalent DNA/LacI configurations are depicted below their corresponding LacI concentrations. (B) Representative temporal records of the TPM excursion parameter $<\rho^2>$ with 0.5 nM [LacI] are ranked by looping probability from 0 (top) to 1 (bottom). Unlooped states are depicted in red, and looped states are shown in blue and green. The actual $<\rho^2>$ values are encoded using the color scale at right.

Although the average behavior of the population of DNA tethers is clear and follows expectation, the looping probabilities of individual DNA tethers under any given LacI concentration are very heterogeneous. Indeed, the whiskers of each box in the upper panel of figure 4.1, range from 0 to 100% looping probability. Raw data displaying this behavior is shown in representative temporal records for DNA tethers exposed to 0.5 nM LacI (Figure 4.1B). Each time record corresponds to a different DNA tether. Unlooped states are depicted in red, and looped states are shown in blue and green corresponding to loops in anti-parallel or parallel configurations [131]. Clearly, loop formation and breakdown occur randomly. However, some tethers are never looped, some are always looped, and some toggle between looped and unlooped states with various degrees of probability.

§4.2.2 Differences between lac repressor molecules contribute to the heterogeneity of looping

To verify that the heterogeneity observed was indeed due to the activity of single lac repressors and not variations among DNA molecules introduced during PCR, a control experiment was performed in which excursions of the tethered beads were monitored for 30 min in λ buffer, and then the first solution containing LacI was washed out with buffer containing a high salt concentration (1 M KCl). Then the same tethered beads were monitored for 20 min in λ buffer. Excursions displayed no looping which verified elimination of LacI. Finally, another volume of solution containing LacI was introduced into the microchamber, and excursions of the bead were monitored for another 30 min. The percentages of time spent by individual tethers in the looped state during intervals when LacI was present were calculated and plotted. Figure 4.2 shows the lack of correlation between the looping probabilities measured for individual DNA tethers with distinct LacI proteins in the two observation periods, suggesting that the the extent of looping is dictated by variations in the activity of associated individual LacI proteins but does reflect heterogeneity amoung DNA tether. While the activity of individual LacI enzymes varied, additional experiments on supercoiled DNA molecules established conditions in which all DNA tethers exhibited a narrow range of looping probabilities within intervals shoter than the doubling time of *E. coli* in nutrient-rich media. Thus, variations among LacI proteins may only modestly affect looping probabilities on negatively supercoiled DNA.

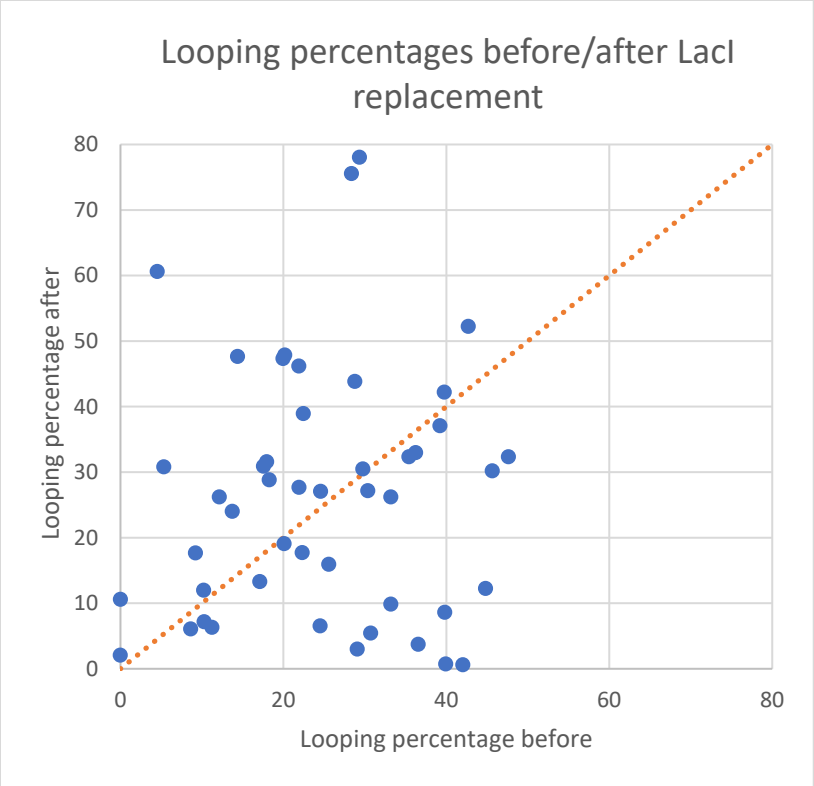


Figure 4. 2. Looping percentages are uncorrelated before and after LacI replacement. Using tethered particle motion analysis, the percentage of time spent in the looped state was measured for a field of view of individual tethers in a first measurement interval. Then LacI was washed out and re-introduced for a second measurement interval. The mean looping percentages before and after are 25+/-13 and 29+/-19, which are indistinguishable by t-test (0.61). However, the mean looping percentages for the individual tethers in the first and second intervals are poorly correlated, with a Pearson's correlation coefficient of 0.066, and do not extend along the line Y=X as would be expected if looping behavior depended on the specific DNA tether. The lack of correlation between looping percentages measured for individual tethers before and after LacI replacement indicates that the percentage of looping is not conditioned by a particular DNA molecule.

§4.2.3 HU protein does not reduce variation in looping probabilities

It is difficult to rationalize how *E. coli* bacteria that should calibrate a response to lactose could contend with such extreme variation in the stability of lac-mediated loops. Thus, there might be factors *in vivo* that make looping a deterministic process. Since protein-mediated looping is a ubiquitous regulatory mechanism across kingdoms, this question is relevant for cells of all organisms. I hypothesized that genome-compacting proteins and DNA supercoiling, which are common to all species, might decrease the variation in looping probability, because they reduce the three-demensional distance between sites joined by the looping protein. As a model genome-compacting protein, I chose the nucleoid-associated protein HU which is abundant in bacteria, binds non-specifically to DNA, contributes to the overall architecture of the genome, facilitates protein-mediated looping, and influences DNA replication and transcription [28, 29, 44, 132-134].

To determine whether HU reduces variation in looping probabilities, tethered particle motion (TPM) experiments were conducted at a LacI concentration of 2.5 nM while the HU concentration was titrated from 0 to ~1000 nM. The LacI concentration was chosen such that it would be easy to measure with negligible uncertainty, and increases or decreases in the looping probability would be obvious [29]. As shown previously by our lab and others, the magnitude of excursions of beads tethered to single DNA molecules in the salt condition used here decreases as HU concentrations increase [29, 47, 135-137], and this protein-induced DNA compaction facilitates looping [29]. Indeed, as shown in Figure 4.3A, increasing the HU concentration increased the median looping probability from around 20 to 80%. However, the looping probabilities of single DNA tethers ranged from 0 to 100% at each HU concentration as shown by the whiskers in the plot. The representative temporal traces in Figure 4.3B, show tethers that never looped, tethers that remained looped throughout the observation, and others that toggled between the looped and unlooped states. This is similar to what was observed for LacI-induced looping without additional factors (Figure 4.1), indicating that HU, despite its ability to compact DNA and favor looping overall, did not reduce the variation of looping probabilities between DNA tethers.

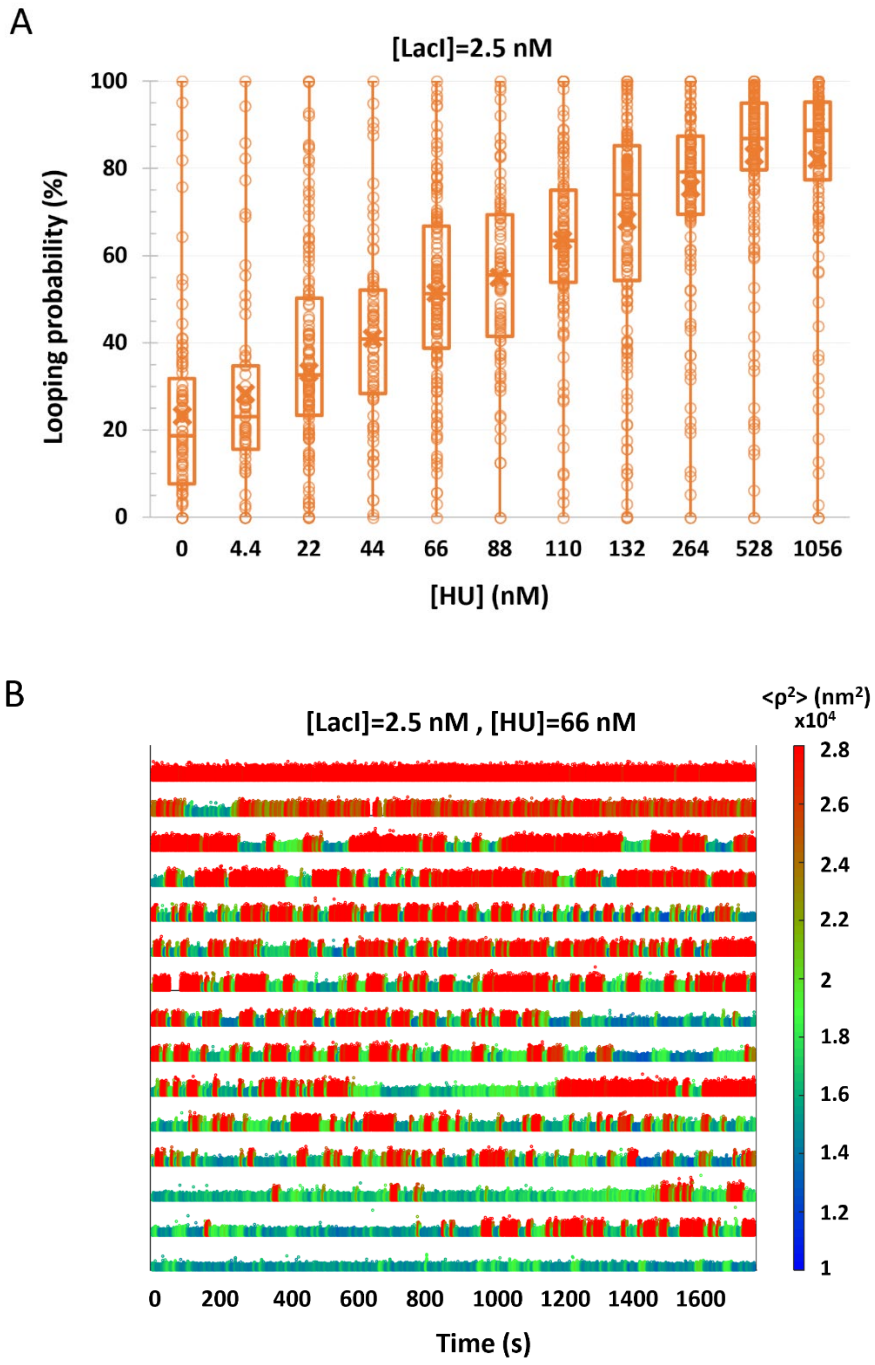


Figure 4. 3. HU did not reduce the variation of looping probabilities amongst DNA tethers. (**A**) The calculated looping probabilities of individual tethers, exposed to 2.5 nM and a range of HU protein concentrations, are summarized in a box-whisker plot. As the HU concentration was titrated from 0 to 1056 nM, the average looping probability progressively increased from 25 to just above 80. Note however, that

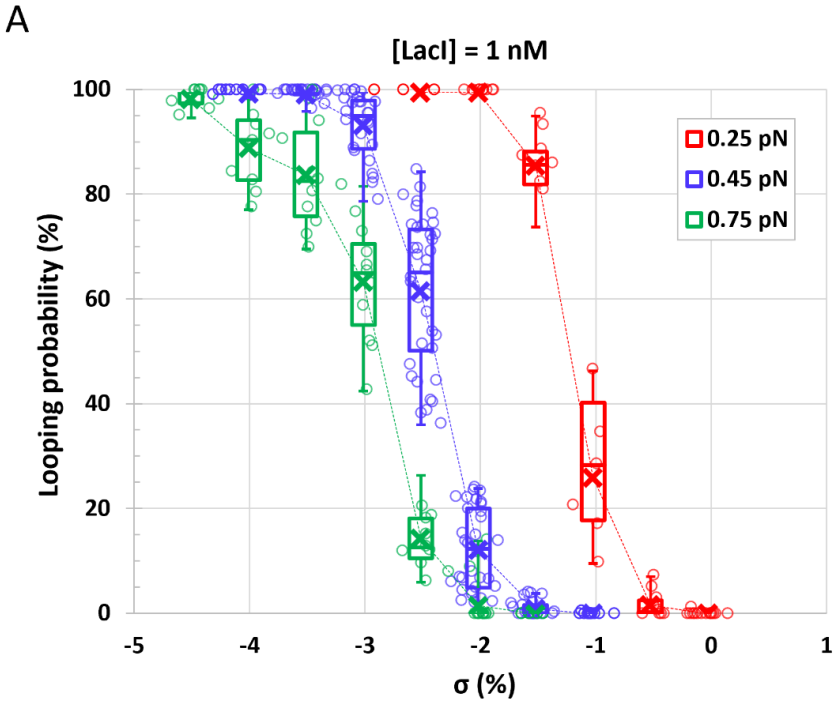
at each concentration of HU, the looping probabilities for individual DNA tethers varied from 0 to 100%. The upper and lower borders of the boxes indicate the upper and lower quartiles. The midline and cross of each box indicate the median and average of the distribution of looping probabilities. (**B**) Representative temporal records of the TPM excursion parameter $<\rho^2>$ with 2.5 nM [LacI] and 66 nM [HU] are ranked by looping probability from 0 (top) to 1 (bottom). Unlooped states are depicted in red, and looped states are shown in blue and green. The actual $<\rho^2>$ values are encoded using the color scale at right.

§4.2.4 DNA supercoiling reduces variation in looping probabilities

Like nucleoid-associated proteins, DNA supercoiling contributes to genome compaction.. In live cells, DNA supercoiling is ubiquitous and dynamic [20, 21, 138], genomes are negatively supercoiled overall. Furthermore, it has been shown that DNA unwinding under low, physiological forces compacts DNA in a way that facilitates looping by proteins [29]. To measure the effect of supercoiling on the variation in looping probabilities, magnetic tweezers [139] were used to modulate supercoiling and measure LacI-induced looping, focusing on the level of heterogeneity of the looping probabilities of different tethers. In these experiments, the magnetic field generated by a pair of magnets above the microchamber attracts and orients the beads applying tension to the DNA. In addition, if the magnets are rotated, DNA will be twisted. If DNA is twisted under low tension, either winding or unwinding will induce plectonemes that decrease the extension of the DNA. LacI-mediated loops persist long enough in supercoiled DNA tethers to reduce the average tether length. These stable intervals of reduced extension mark loop formation and breakdown and endure longer as (-) supercoiling increases.

Figure 4.4 summarizes measurements of LacI-mediated looping probability for the 2115 and 2011 bp DNA tethers unwound to different degrees under three different tensions. Tension completely suppressed looping in torsionally relaxed DNA, $\sigma \approx 0$ %, but negative supercoiling compensated and progressively increased the looping probability at all three levels of tension. This is consistent with enhanced the binding of LacI

to negatively supercoiled DNA [7]. As expected based on previously published data increasing (negative) supercoiling drove the looping probability to 100% [29]. However, what is remarkable is that looping probabilities measured for different DNA tethers were tightly grouped around the median values. Whereas the ensemble of torsionally unconstrained DNA tethers in TPM experiments displayed the entire range of looping probabilities, negative supercoiling dramatically reduced variation and produced quite uniform, deterministic behavior at each of the three tensions. This uniformity is illustrated in Figure 4.4B, in which individual time traces, recorded under 0.45pN of tension and σ ≈ -2.5% are quite similar and exhibit frequent switching between extended, unlooped (red) and shortened, looped (blue/green) states. As introduced in section 1.2.2, *in vivo* supercoiling levels range between -7.75 and -6.5%. Here, much lower levels of supercoiling shift the equilibrium completely to the looped state.



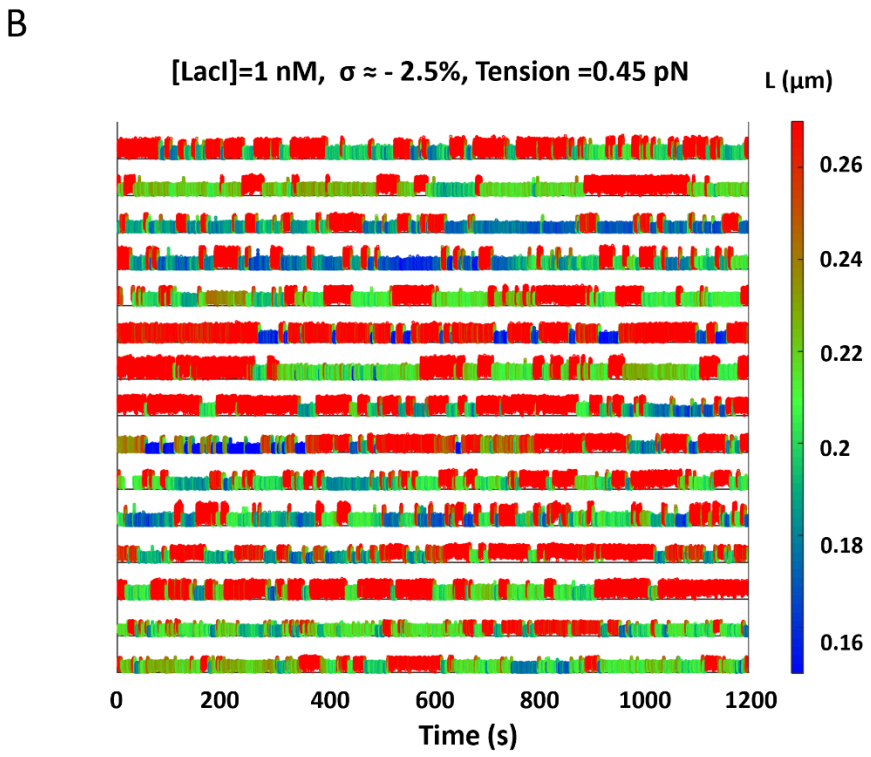


Figure 4. 4. Supercoiling dramatically reduces the variation of looping probabilities amongst DNA tethers. (A) The looping probabilities of individual tethers, exposed to 1 nM LacI and negatively supercoiled to varying degrees, are summarized in a box-whisker plot. As the supercoiling was varied from 0 to almost -5% (σ), the average looping probability progressively increased from 0 to 100. Note that at each level of supercoiling, the looping probabilities for individual DNA tethers formed compact distributions. The upper and lower borders of the boxes indicate the upper and lower quartiles. The midline and cross of each box indicate the median and average of the distribution of looping probabilities. The whiskers and quartiles are only distinct for intermediate values of negative supercoiling. For low and high levels of negative supercoiling the whiskers collapse to the median value. (B) Representative temporal records of the instantaneous lengths of DNA tethers exposed to 1 nM [LacI] while stretched by 0.45 pN of tension and supercoiled to σ = -2.5% are shown. Unlooped states are depicted in red, and looped states are shown in blue and green. The actual tether length values are encoded using the color scale at right.

§4.3 Supercoiling induced ergodicity within a biologically relevant timescale

It was puzzling that single supercoiled molecules observed over a period of 20 minutes behaved ergodically, while single, torsionally relaxed molecules observed for up to 30 min did not. I hypothesized that supercoiling changed the dynamics of looping and that torsionaly relaxed molecules might display ergodic behavior over longer intervals, Therefore, I conducted several much longer, up to five-hour, recordings of LacI-mediated looping in torsionally relaxed DNA using TPM (Figure 4.5).

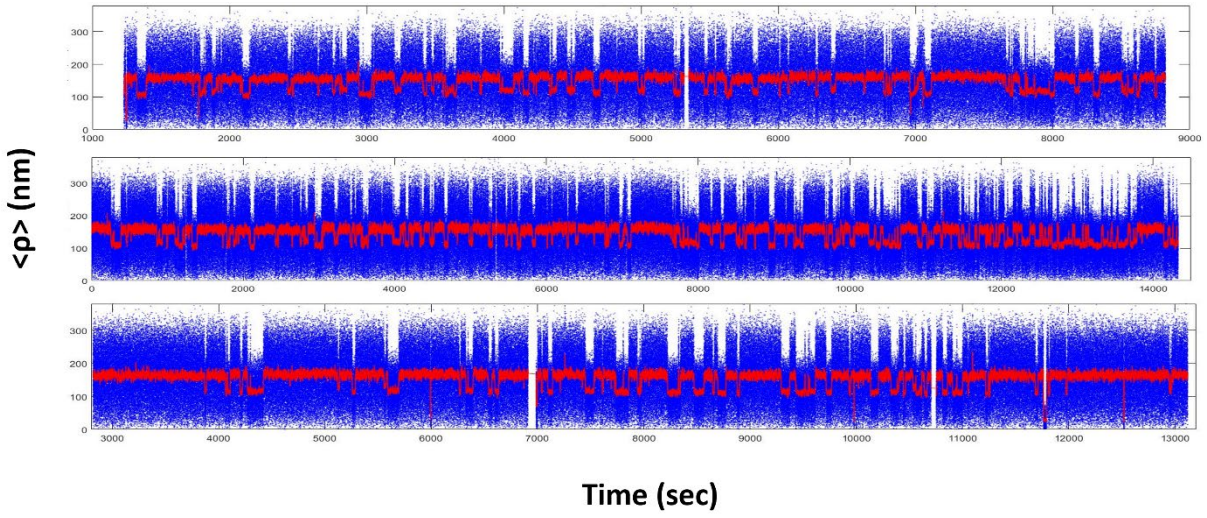


Figure 4. 5. Long TPM recordings. Three representative recordings are shown from a set of fifty lasting over one hour. The blue dots represent the amplitude of two-dimensional projections of momentary excursions of the tethered bead; the red trace is the 8-second moving average of these values.

I then measured the looping probabilities over temporal windows of different lengths, ranging from 10 min to the entire 5 hour-long recording. Figure 4.6A is an overlay of the cumulative probability distributions for looping percentages calculated for entire 5-hour records (black) or divided into shorter segments of 10 (blue), 30 (green), 40 (red), 60 (cyan), 80 (magenta), and 100 (yellow) min. At least 60 min of observation are required to accurately sample the dynamics of LacI-mediated looping in a DNA construct containing

the O1 and O2 operators separated by 400 bp, i.e., ergodicity is attained with recordings of no less than 60 min. The looping probability for negatively supercoiled DNA was analyzed in a similar manner and showed that a cumulative distribution of looping probabilities equivalent to that attained in 20 min-long measurements is achieved in just 12 minutes (Figure 4.6B, Figure 4.7).

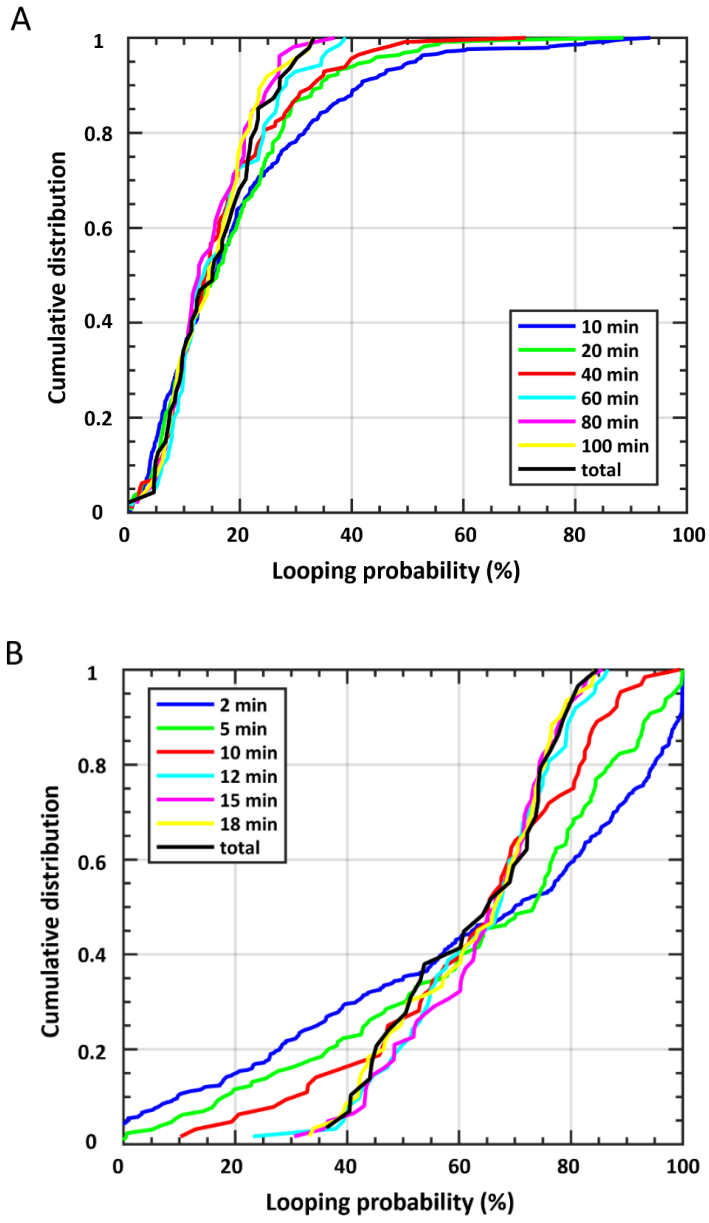


Figure 4. 6. Minimum observation times for ergodicity. Cumulative distributions of the looping probability for (A) different intervals of TPM observations (10, 30, 40, 60, 80, 100 min, or 5 hours) and (B)

different intervals of MT observations (2, 5, 10, 12, 15, 18, or 20 min) are plotted along with the cumulative distribution for the entire set of observations (black curves). All records used in the MT analysis were for 0.25 pN tension and 1.5% supercoiling. Only for observations of at least 60 min (*A*) or 12 min (*B*) do the cumulative distributions of the looping probabilities for individual molecules resemble the distributions for the maximum observation intervals. With shorter observation intervals, the looping probabilities measured for individual molecules varies widely. For example, in the TPM data in *A,* looping probabilities measured for molecules observed for 10-minute intervals (blue) displayed a significant fraction of high probabilities that did not appear for the molecules observed for 60 min or more. In *B*, looping probabilities spanning the entire range result for 2- or 5-min intervals of observation but narrow to the 40-80% range when observation times are 12 min or longer

The TPM records in this work were between 20 and 30 minute-long, slightly below, or equal to the doubling time of *E. coli* which is approximately 30 minutes in the laboratory [140]. The heterogeneity of looping probabilities observed in these records without supercoiling were extreme and seemingly at odds with a molecular system designed to respond to the presence of lactose. Such a system was expected to be ergodic, such that sufficiently long observations of single members of the ensemble would have exhibited the statistical behavior of the whole ensemble. When several much longer, five-hour, recordings of LacI-mediated looping in torsionally relaxed DNA were acquired using TPM and the looping probabilities were measured over temporal windows of different lengths it was clear that observations for periods shorter than 60 min exhibit a tail of high looping probabilities that is not present in distributions from longer observations. Only recordings greater than or equal to 60 min produce distributions like the five-hour distribution. In other words, ergodicity results when the dynamics of LacI-mediated looping in a DNA construct containing the O1 and O2 operators separated by 400 bp are observed for intervals of 60 min or more. This is due to the inherent stochastic nature of protein-mediated looping. If, however, DNA molecules are supercoiled, the reduced dimensionality modulating juxtaposition of the protein binding sites accelerates

dynamics such that the statistical behavior of the ensemble can be revealed in much shorter observations of a single molecule. Indeed, analysis of looping probability in records for unwound DNA show that a cumulative distribution of looping probabilities equivalent to that attained in 20 min-long measurements is achieved in just 12 minutes (Figure 4.6B, Figure 4.7). Thus, looping dynamics in supercoiled DNA are deterministic within the time scale of the doubling time of the bacteria, effectively the cell cycle.

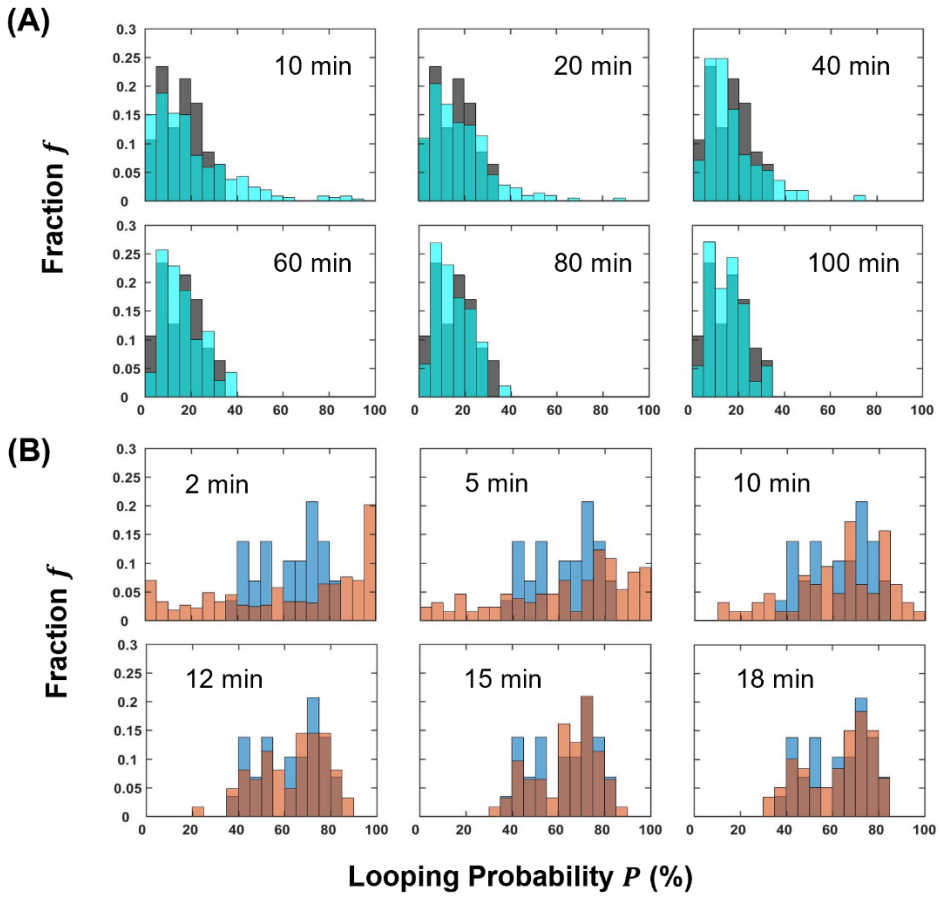


Figure 4. 7. Sufficiently long observations are ergodic. (A) Distributions of the looping probabilities derived from TPM records lasting 5 hours (gray) or split into shorter segments as indicated (cyan). **(B)** Distributions of the looping probabilities derived from MT records lasting 20 min (blue) or split into shorter segments as indicated (brown). All data was recorded at 0.25 pN of tension and -1.5% supercoiling. The

distributions of looping probabilities for the corresponding full-length record are included for comparison in each panel.

§4.4 Discussion and conclusion

§4.4.1 HU protein induced supercoiling is insufficient to promote uniform looping dynamics

Since HU binding is known to supercoil DNA [141], it might similarly alter looping in a torsionally constrained DNA tether. This was not possible in a TPM experiment (Figure 2.4A) in which single-bond attachments of the DNA to the surfaces would swivel to release any torsion. However, previously published data [29] showed that 1056 nM HU which produces a median value of 85% average looping probability in DNA tethers under no tension (Figure 4.3), produced just 3% or 5% in DNA tethers under 0.25 or 0.45 pN of tension, respectively. If HU were actually supercoiling the DNA, then it should have sustained the looping probability under tension. The fact that it did not suggests that HU binding only compacts but does not significantly supercoil DNA in the conditions used here. To verify this, extension versus twist curves were measured for DNA in the presence of HU. In this assay, any unwinding associated with protein or small molecule binding to the DNA will alter the intrinsic twist of the molecule and shift the maximum extension of the molecule with respect to that of bare DNA [142-144]. As can be seen in Figure 4.8, increasing the concentration of HU up to 1000 nM steadily contracted a DNA tether stretched by 0.45 pN in 200 mM KCl but negatively supercoiled the molecule by only -1.35 turns. This is equivalent to -0.4% supercoiling, which is insufficient to produce writhe [33].

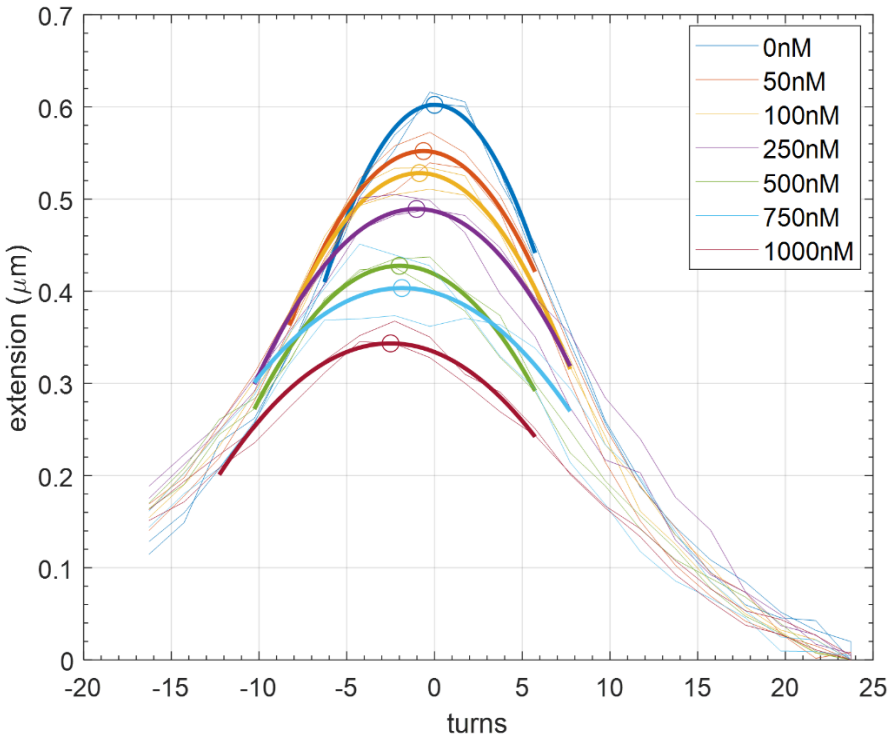


Figure 4. 8. HU binding significantly contracts but mildly unwinds DNA. A single 3352 bp DNA tether under 0.45 pN of tension in a magnetic tweezer was repeatedly wound and unwound in a buffer containing 200 mM KCl with various concentrations of HU ranging from 0 to 1000 nM. Increasing HU concentrations steadily reduced the extension of the DNA tether by nearly 50% but produced only very mild negative supercoiling. A parabolic curve (thick) was fit to each pair of winding and unwinding curves corresponding to different concentrations of HU. The vertex of the parabola represents the shift of the twist versus extension curve due to the supercoiling induced by HU protein binding to the DNA tether. (Figure courtesy of Yan Yan, PhD. And Alex Zhang)

Furthermore, HU remained bound to DNA tethers twisted to high levels of positive or negative supercoiling but could be washed out with 300 mM KCl (Figure 4.9). Thus, HU enhances looping by contracting DNA tethers but does not change the dimensionality of the looping mechanism to promote uniform looping dynamics.

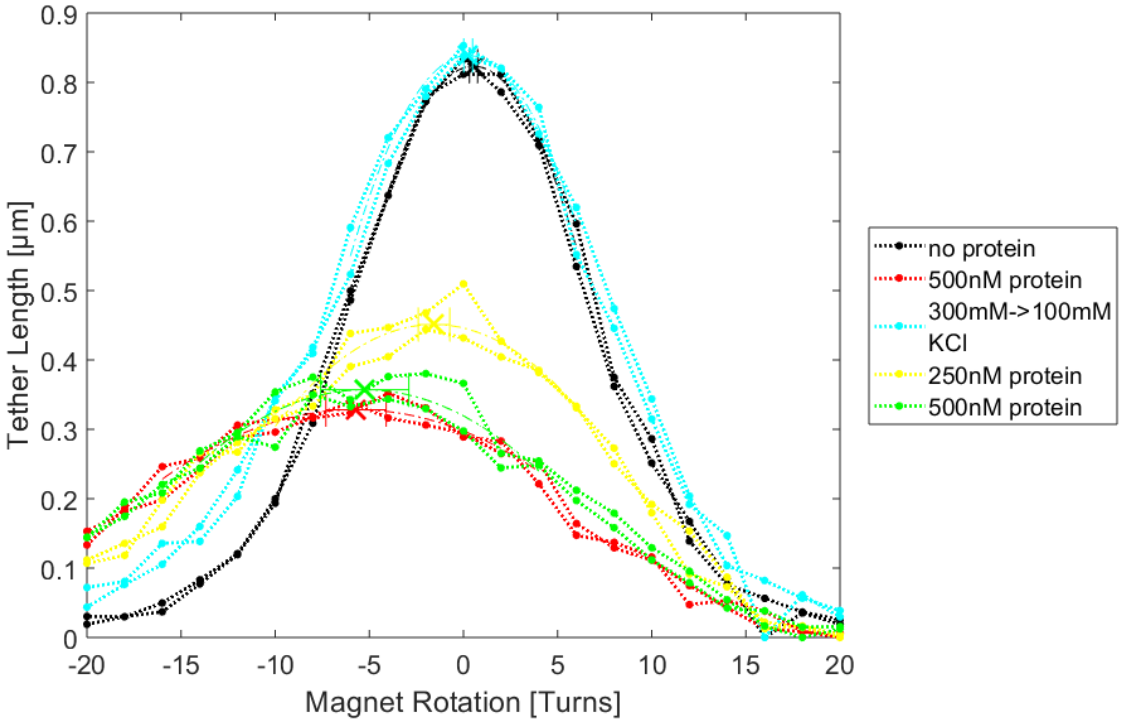


Figure 4. 9. HU does not dissociate under a wide range of supercoiling levels. Extension versus twist curves were recorded in 100 mM KCl, after introducing 500 nM HU, after washing away HU with high salt, after re-introducing 250 and then 500 nM HU. 500 nM protein shifted and reduced the maximum extension of the curve. 300 mM KCl dissociated HU and restored the curve to the original position and height. Introduction of 250 nM HU partially shifted and reduced the maximum extension of the curve. Re-introduction of 500 nM HU shifted and reduced maximum extension of the curve to the levels observed previously. (Figure courtesy of Yan Yan, PhD. And Alex Zhang)

§4.4.2 DNA supercoiling may lower the energy barrier of looping by juxtaposing operators through 1-D diffusion

In the absence of supercoiling, the probabilities of LacI-mediated DNA looping for individual DNA tethers with, as well as without, HU range from 0 to 100%. In these conditions, the LacI binding sites must

juxtapose by 3-D diffusion opposed by a high energy barrier before the protein can connect them. HU protein helps to overcome this barrier and enhance looping by compacting the DNA to reduce the separation between the operators to be bridged. In contrast, supercoiling induces plectonemes which allow slithering of DNA segments past one another [31, 32]. Thus, the operators may juxtapose through 1-D diffusion, across a much lower energy barrier. Supercoiling reduced the dimensionality of the path to juxtaposition, to produce looping dynamics in DNA tethers that could not be achieved by adding HU to facilitate juxtaposition in three dimensions.

Figure 4.10 is an illustration of hypothetical energy landscapes for looping, which involves bending and possibly twisting the DNA molecule, as shown in the cartoons on the left. Energy is color-encoded according to the scale at right. The end-to-end distance of the loop segment is represented along the vertical axis. Near-zero end-to-end distance corresponds to the looped state, while the unlooped states are more extended, up to 400 bp. The horizontal axis indicates the supercoiling in the DNA. Panel (*A*) describes the experimental conditions in a TPM measurement, with no applied tension or torsion, the DNA may coil, bend, and juxtapose the two operators via 3-D diffusion, as shown in the superimposed cartoon. Different pathways may be followed through the broad and shallow saddle point separating the looped and unlooped states. In panel (B), sub-pN tension gently extends DNA in the absence of imposed supercoiling, and the work associated with drawing the two operator sites together increases the energy barrier, effectively attenuating loop formation. These conditions correspond to a magnetic tweezer measurement in which the magnets exert tension but are not rotated to twist the DNA as shown in the superimposed cartoon. However, in plectonemic DNA as shown in the cartoon superimposed on panel (C), operator sites can juxtapose as DNA segments slither past each other in the plectoneme. This one-dimensional search for juxtaposition between looped and unlooped states changes the dynamics of looping with respect to the three-dimensional searches required in torsionally unrestrained DNA tethers.

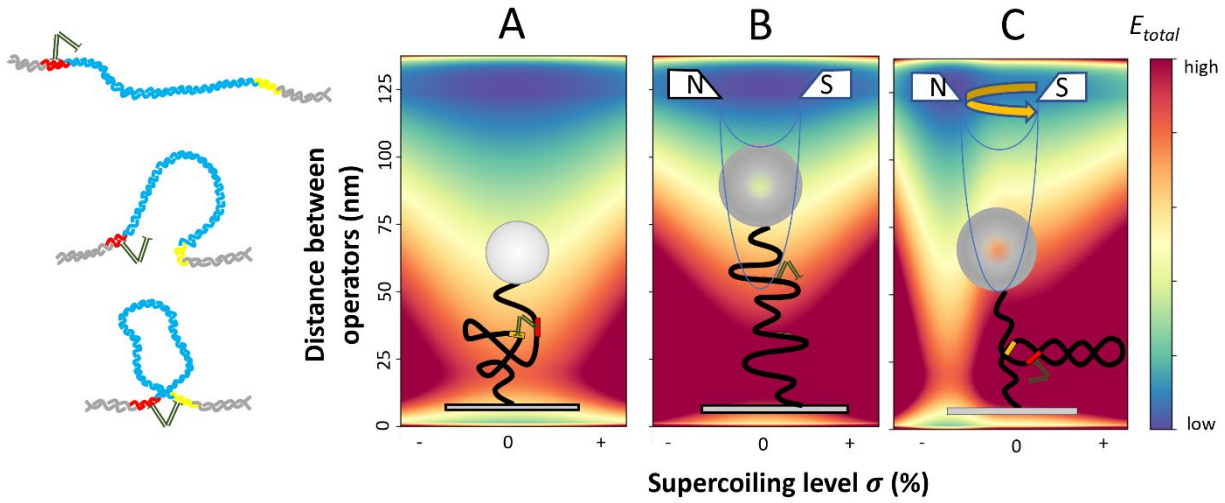


Figure 4. 10. Energy landscapes for loop closure by LacI in different conditions of tension and torsion. The illustrations at the left show possible conformations of DNA tethers corresponding to different separations between the operators. *A*-*C* are three hypothetical energy landscapes for a DNA tether under different conditions of torsion and tension. The energy values are qualitatively encoded using the color scale at right. The *y*-axis indicates the distance between the protein binding sites that constitute the junction and may vary between zero and 133 nm for a 400 bp DNA segment. DNA supercoiling varies along the *x*-axis. Superimposed on each panel are illustrations of likely DNA conformations under the different conditions of torsion and tension.

§4.4.3 Conclusion

Recently, LacI was observed to hop along the double helix [145]. This feature together with negative supercoiling is probably key for the protein to efficiently locate a binding site, contact a secondary binding site, and maintain a constant ratio of looped versus unlooped states in the various states of tension and torsion that likely develop during the cell cycle. This ratio gives a definitive regulatory outcome to looping based on inherently stochastic molecular activity. Additional parameters may contribute to achieving a specific outcome, but for straightforward protein-mediated looping, DNA supercoiling appears to be key.

Chapter 5 Positive supercoiling favors transcription elongation through lac repressor-mediated DNA loops

§5.1 Summary

Although loop formation can be tuned by negative supercoiling, how RNA polymerase, which produces supercoiling, passes through loop structure is also interesting to me. In this chapter, I describe experiments on the passage of RNAP through unlooped and looped LacI obstacles. The difference in the duration of pauses that RNAP exhibits at the two types of obstacles indicates the blocking efficiency. Transcription monitored using TPM showed that RNAP paused longer at a looped compared to an unlooped LacI obstacle. Furthermore, elongation through a looped segment was significantly delayed with respect to elongation of an identical unlooped segment. Given that a loop constrains supercoiling [146], and accumulated supercoiling can significantly affect the RNAP behavior [111], the significant delay within the loop may be due to supercoiling. Remarkably, when RNAP transiting within a loop encountered the LacI loop closure, it paused less compared than at unlooped LacI obstacles. To mimic this supercoiling effect at unlooped LacI obstacles, using magnetic tweezers I preloaded a DNA tether with negative supercoiling such that RNAP would create positive levels of supercoiling after elongation from the promoter to an obstacle. The results showed that positive supercoiling diminished pausing and helped RNAP overcome a LacI obstacle. Positive supercoiling propagating ahead of RNAP appears to facilitate elongation along protein-coated DNA.

§5.2 RNAP pauses longer at entry to than exit from LacI-loops

§5.2.1 Monitoring elongation through LacI-mediated loops with tethered particle motion

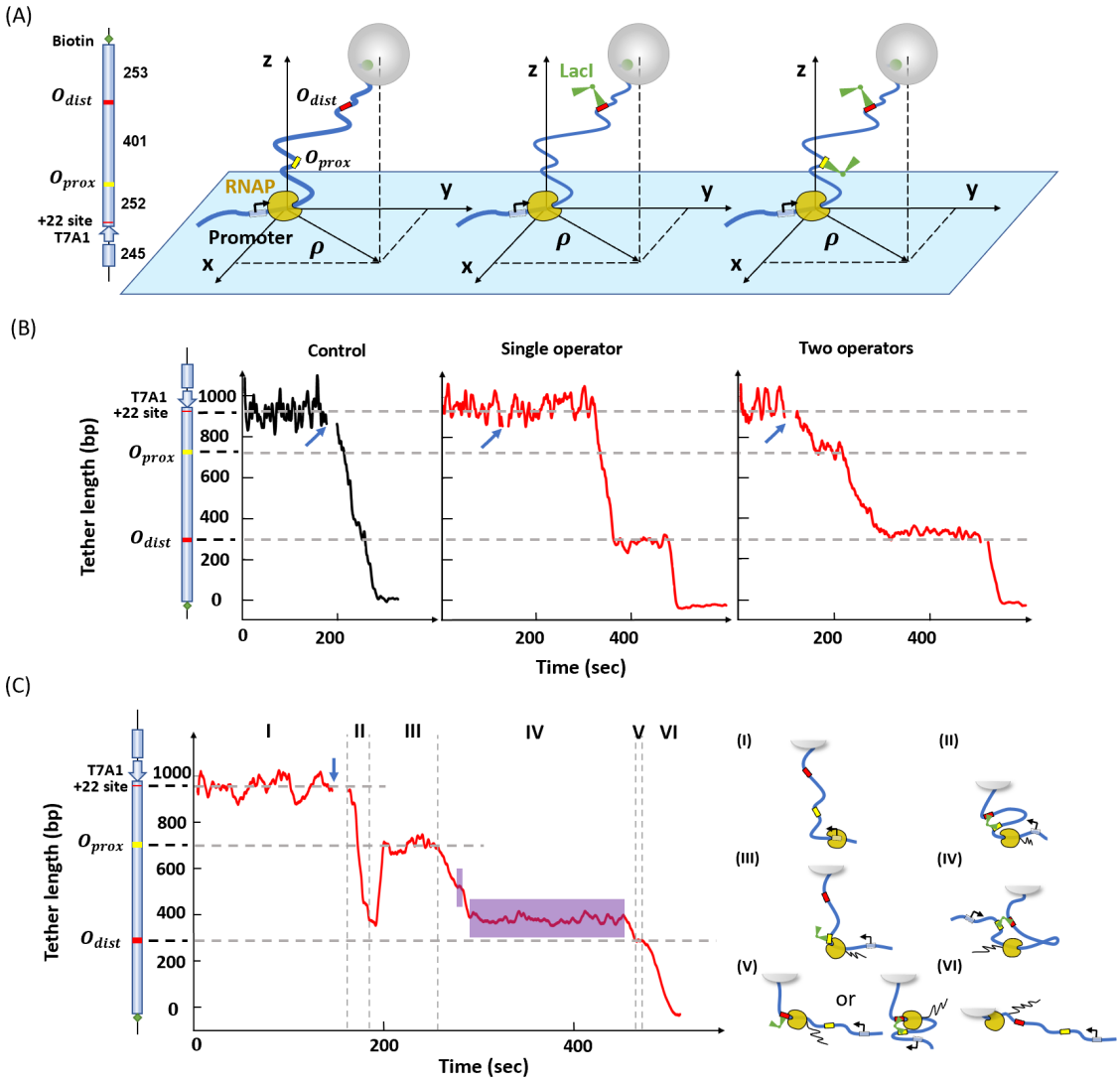


Figure 5. 1. RNAP can transcribe through a LacI-mediated loop after a pause. (A) On the left is a schematic diagram of the transcription template used to tether a microbead. Diagrams of the TPM samples at right show DNA (thick blue lines), LacI binding sites (yellow and red segments) and LacI (green, V-shape). The black arrow near the promoter indicates the direction of transcription starting at the +22 site. From left to right they depict RNAP elongation without LacI (left), with LacI bound to single operator (middle), or with distinct LacI tetramers bound to the operators (right). (B) Plots of the amplitude of

excursions of tethered beads versus time corresponding to these cartoons are shown in Panel B. At left is a control experiment without LacI. Data in the middle shows transcription in the presence of 10 nM LacI on a DNA template with only the distal LacI binding site. The data at right shows transcription in the presence of 10 nM LacI on a DNA template with two LacI binding sites. At this concentration, each binding site is likely to be occupied by a distinct LacI tetramers, so looping is rare. The blue arrows indicate the time at which NTPs were introduced. (C) The data displayed at left shows transcription on a DNA template with two LacI binding sites in the presence of 0.2 nm LacI, conditions that promote looping. The vertical dashed lines identify six intervals (I - VI) in the progress of RNAP along the DNA template and the cartoons on the right depict the likely conformations of transcription elongation complexes in each interval. The purple areas in region IV indicate random pauses between operators. Schematics of the DNA templates left at in panels B and C shows the features of the DNA template used in TPM measurements. From top to bottom: a T7A1 promoter, a stall site at +22, a promoter-proximal binding site (O_{prox}) and promoter-distal binding site (O_{dist}). The dashed horizontal lines in the data plotted to the right indicate the positions of the LacI binding sites in the construct, and the expected locations of pauses in the data.

As RNAP transcribed the DNA template containing two LacI binding sites, three scenarios were possible: RNAP might encounter an unencumbered binding site, a binding site bound by LacI in unlooped DNA (Figure 5.1A, B), or a LacI bridging two operators to secure a DNA loop (Figure 5.1C). In addition, the progress of RNAP, LacI might randomly bind to/dissociate from either the promoter-proximal or -distal binding site, O_{prox} and O_{dist}, respectively, producing/breaking intermittent loops. In data that satisfied screening criteria, Figure 5.1B for example, prior to addition of all four nucleotides the average excursion of the bead remained constant at a value consistent with the DNA tether length. Adding NTPs with or without LacI to initiate transcription, indicated by the blue arrows in Figures 5.1B and 5.1C, caused a short-lived disturbance, which was omitted from the plot. Shortly afterward, RNAP began elongation producing a progressive decrease in tether length that continued uninterrupted to the end of the template, unless LacI

was present. In fact, control experiments with 1 mM NTPs revealed no pausing by RNAP at either of the two LacI binding sites in the absence of LacI (Figure 5.1B, left). In contrast, in the presence of 10 nM LacI, RNAP clearly paused at O_{dist} on a DNA template containing only this LacI binding site (Figure 5.1B, middle) or paused in front of both O_{prox} and O_{dist} on a DNA template containing both LacI binding sites (Figure 5.1B, right).

The probability of looping in a DNA template containing two binding sites can be adjusted by varying the LacI concentration (Figure 2.7 or Figure 4.1A) [27]. For a 400-bp loop between the O2 and O1 binding sites, a maximum looping probability of ~44% could be achieved with 0.2 nM LacI, while 10 nM LacI decreased the probability to ~6% as the two binding sites became occupied by distinct LacI tetramers (Figure 2.7). Indeed, with 10 nM LacI, RNAP paused at positions corresponding to the promoter-proximal and -distal LacI binding sites and no loops were observed (Figure 5.1B, right).

However, with 0.2 nM LacI, loops occurred during transcription (Figure 5.1C) and the pattern was more complex. The record can be divided into six intervals (I-IV). In interval (I) the DNA tether length is constant before introducing 1 mM NTPs and 0.2 nM LacI. Transcription begins shortly after introducing NTPs and in interval II, a loop forms, but it ruptures at the beginning of interval III as RNAP approaches O_{prox} and pauses indicating that LacI is bound. In interval (III), RNAP paused at O_{prox} for approximately 50 s. In interval (IV), RNAP proceeded to transcribe the segment between the LacI binding sites and paused at two random locations (purple areas) before reaching the distal operator.

Traces without loop formation (Figure 5.1B) did not reveal any random pauses. Thus, I concluded that pauses within the loop were induced by the accumulation of supercoiling within a loop, as depicted in Figure 5.1C cartoon IV. A previous study [111] had shown that a torque of +11 pN·nm ahead, or -11 pN·nm behind RNAP can stall its progress. Therefore, I calculated the possible progress that RNAP might make as inside a closed loop. Since a loop is a torsionally constrained region, an RNAP might proceed furthest within a loop if it is in the middle of the loop at the moment the loop closes and constrains torsion. At this

location in which the length of DNA flanking RNAP in both direction within the loop is maximal, the torque in the flanking DNA can be expressed as: $\tau \approx \frac{2\pi k_B T C_{eff}}{L_s} \Delta Lk$ [147] (table 1, equation 7), in which L_s is the contour length of the flanking DNA. C_{eff} is the effective twist persistence length (twist modulus/k_BT) determined as $C_{eff} = C(1 - \frac{C}{4A}\sqrt{\frac{k_BT}{Af}})$ [147], where C = 100nm, A = 50nm, and the tension, *f*, in this case is 0.45 pN. ΔLk is the linking number change in the DNA. In this study, $L_s \approx 200\ \text{bp} * \frac{3.4}{10.5}\frac{\text{nm}}{\text{bp}} = 64.8\ \text{nm}$, $C_{eff} \approx 63$ nm, $k_BT \approx 4.1$ pN·nm. Therefore, the maximum change in linking number that RNAP can induce before stalling ($|\tau| \approx 11$ pN·nm), is $\Delta Lk = \frac{|\tau| L_s}{2\pi k_B T C_{eff}} = \frac{11*64.8}{2\pi*4.1*80} \approx 0.44$ turns, which corresponds to $\Delta Lk * 10.5\frac{\text{bp}}{\text{turn}} \approx 5$ bp. Thus, 5 bp is the furthest that RNAP should be expected to transcribe before stalling when it operates inside a ~400 bp loop, and this maximum is expected when RNAP is halfway between the two operators. According to this calculation, RNAP can stall very quickly after loop closure which corresponds to the random pauses in interval (IV). Elongation resumed after loop breakdown and RNAP transcribed to O_{dist}. In interval (V), RNAP paused at O_{dist} indicating that LacI was still associated with this operator, then continued transcription in interval (VI), finally reaching the end of the template. This record illustrates how transcription through a loop can be monitored using TPM.

§5.2.2 RNAP pauses longer entering than exiting LacI loops

To measure the time required for transcription of the loop and determine whether looping affected the passage of RNAP through bound LacI, the pause times of RNAP at O1 or O2 binding sites in the O_{prox} positions in unlooped templates (Figure 5.1B, rightmost) were compared with those in looped templates (Figure 5.1C, interval III). On unlooped templates, pauses at proximal or distal O1 (O2) operators occupied by LacI were similar and their pause time were aggregated in Figure 5.2. A loop had profound effects on RNAP elongation. RNAP paused longer at promoter-proximal LacI-O1 or LacI-O2 obstacles that were part of a loop (Figure 5.2A, left). Pauses at LacI-$O1_{prox}$ obstacles were 77 ± 7 s (N=215) without a loop, but 199

± 25 s (N=92) with a loop. Pauses at LacI-$O2_{prox}$ obstacles were 40 ± 5 s (N=104) without a loop, but 79 ± 14 s (N=58) with a loop. A likely reason for the increases is that the secondary, distal binding site increases the local concentration of LacI and the effective affinity for the proximal binding site [27]. The steric hindrance caused by the loop itself may also contribute to the lengthening of the pause at the proximal binding site, although this effect is likely to be small [148]. It is also informative that looped LacI-$O1_{prox}$ obstacles obstructed transcription for a considerably longer time (199 s / 77 s) than looped LacI-$O2_{prox}$ obstacles (79 s / 40 s) with respect to unlooped templates. Since LacI has higher affinity for the O1 than the O2 binding site, LacI is more likely to remain at O1 than at O2 after the loop break down, increasing the probability that an elongating RNAP pauses at O1.

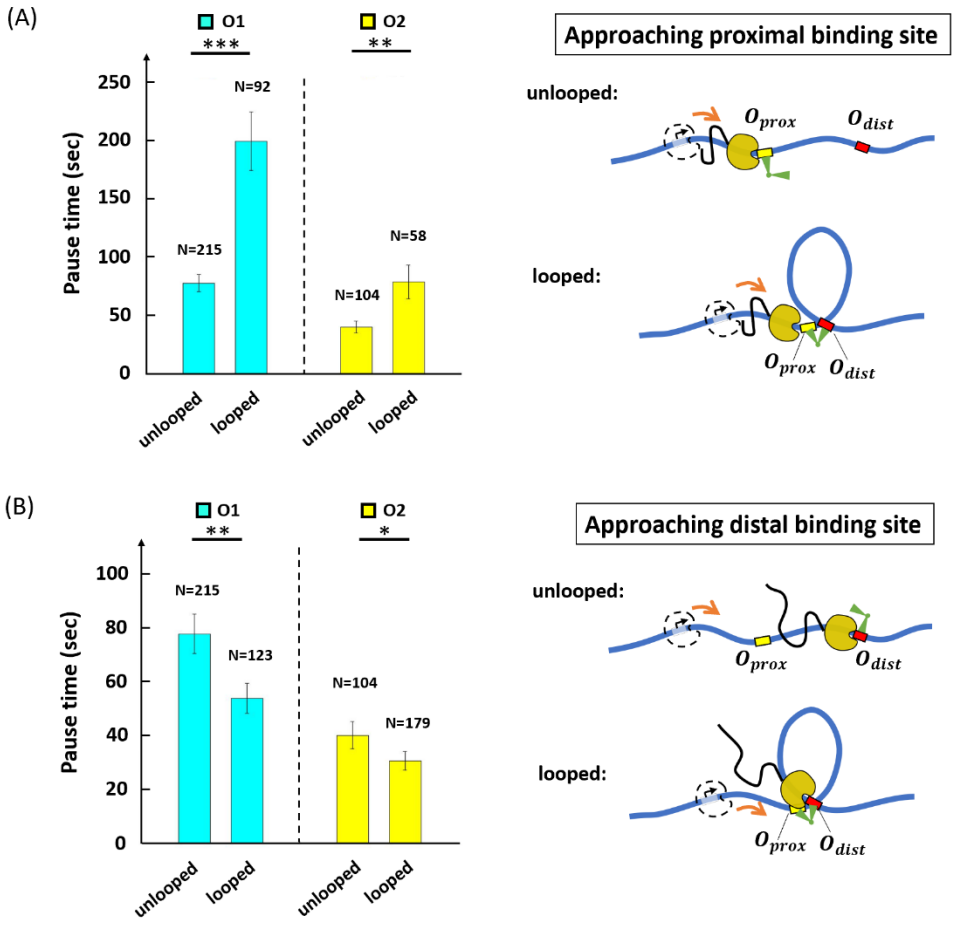


Figure 5. 2. The LacI-mediated loop enhances and attenuates RNAP pausing at the proximal and distal binding sites respectively. (A) Average pause durations were longer at the proximal operator

consisting of either O1 (cyan) or O2 (yellow), in the looped with respect to the unlooped conformations depicted at right (***$\boldsymbol{p \leq 0.001}$, **$\boldsymbol{p \leq 0.01}$ for two-sample t-tests). (B) Average pause durations were shorter at the distal operator consisting of either O1 (cyan) or O2 (yellow), in the looped with respect to the unlooped conformations depicted at right (**$\boldsymbol{p \leq 0.01}$, *$\boldsymbol{p \leq 0.12}$ for two-sample t-tests). Standard errors and sample size, *N*, are indicated.

I next investigated RNAP approaching the promoter-distal operator, O_{dist}. When RNAP is at the distal operator, any loops that form change the configuration of the already transcribed DNA segment but do not change the length of the untranscribed DNA ahead of RNAP (Figure 2.2 and cartoon (V) on the right of Figure 5.1C). I used traces such as those in the rightmost panel of Figure 5.1B and in interval (V) of Figure 5.1C to compare pause durations at O_{dist} with or without looping. Under the 0.2 nM LacI concentration used, 41% of the LacI obstacles at O_{dist} can be assumed to secure loops (Figure 2.7). Surprisingly, and in contrast to O_{prox}, I observed that under these conditions, the average pause time of RNAP at LacI-$O1_{dist}$ and LacI-$O2_{dist}$ was shortened with respect to obstacles on unlooped templates to 54 ± 6 s (N=123) and 30 ± 4 s (N=179), respectively. Assuming that pauses result from LacI bound to either looped or unlooped templates and 44% looped obstacles, RNAP inside the loop pauses at LacI for approximately $\frac{54 - 0.59 * 77}{0.44} \approx$ 19 s at the distal O1 operator, and for $\frac{30 - 0.59 * 40}{0.44} \approx 15$ s at the distal O2 operator. These results suggest that transcription within the loop promotes the release of LacI from a distal operator site, despite the locally increased concentration of LacI and regardless of the operator affinity.

§5.3 Supercoiling plays a key role in regulating RNAP elongations through loop

§5.3.1 Transcription of looped segments is slower

Inspection of interval IV in Figure 5.1C suggests that a protein-mediated loop can significantly delay transcription by RNAP. According to the twin-supercoiled-domain model [1], rotation of the DNA template unwinds DNA behind the transcribing RNAP, generating negative supercoiling (Figure 5.3A, red DNA), and winds DNA ahead, generating positive supercoiling (Figure 5.3A, yellow DNA). Furthermore, since a LacI-mediated loop constitutes a topological domain [146], transcription within the loop will generate torsional stress. Within a 400 bp-long loop, the torsional stress created by a transcribing RNAP can quickly accumulate to +11 pN·nm ahead or -11 pN·nm behind, stalling RNAP progress[111]. I estimate that RNAP might translocate as few as five bp within the 400 bp loop before stalling (last paragraph of §5.2.2). Stalled RNAP is prone to backtracking, and recovery from the backtracked state delays RNAP progress. Thus, I measured the total time required to transcribe the loop segment in each trace (duration of interval IV in Figure 5.1C), averaged it over all traces in the presence of LacI/looping and compared it with the average time required to transcribe between the two operators in the absence of LacI/looping (Figure 5.1B, black control). The average transcription time for looped O1-O2 and O2-O1 segments were 192 $\pm$ 31 s (N=104) and 185 $\pm$ 29 s (N=86), respectively. Both these times were much longer than the average time without looping (32 $\pm$ 5 s; N=35) (Figure 5.3B). I, therefore, conclude that the LacI loop can significantly hinder RNAP progress by generating torsional stress.

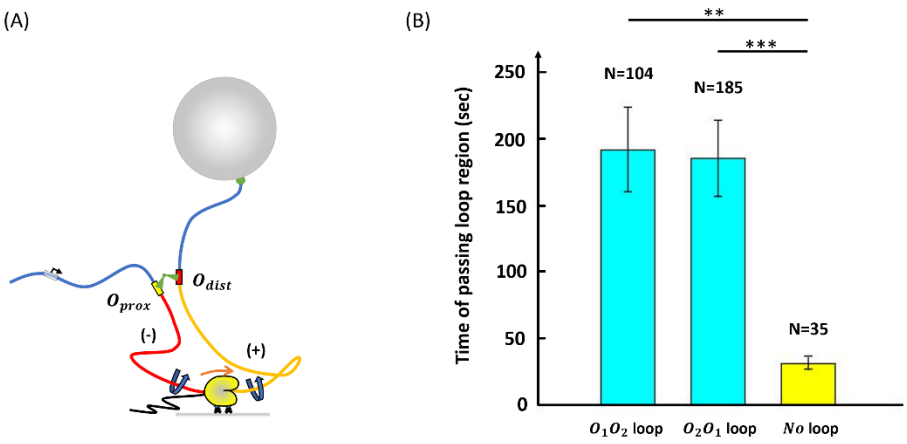


Figure 5. 3. RNAP transcribes a loop more slowly. (A) A cartoon depicting RNAP transcribing a loop. The right-angle black arrow indicates the promoter. The blue-colored DNA segments are torsionally relaxed. The red-colored DNA segment is unwound, while the yellow DNA segment is overwound by RNAP. Nascent RNA is the thin, black line emerging from RNAP. (B) RNAP requires almost tenfold more time to transcribe a looped segment (cyan) compared to the same unlooped segment (yellow). $***\boldsymbol{p \leq 0.001}$, $**\boldsymbol{p \leq 0.01}$ (two-sample t-tests). Standard errors and numbers of samples, *N*, are indicated.

§5.3.2 RNAP surpasses LacI obstacles faster on positively supercoiled templates

Because RNAP was still inside the loop when it encountered the promoter-distal site, I hypothesized that the shorter pauses at the distal operator might result from destabilization of LacI-O_{dist} complexes by transcription-generated positive supercoiling. To test this hypothesis, I used magnetic tweezers (MTs) to follow elongation of an RNAP transcribing toward a LacI-O1 obstacle on a construct where the DNA ahead could be positively supercoiled (Figure 5.4A, Figure 2.9B). In this experiment, the segment between RNAP and the tethered bead was rotationally immobilized by multiple biotin-streptavidin linkages to the bead. The promoter of this template was 253 base pairs (approximately 24 turns) upstream from the O1 binding site (Figure 2.2). To create positive supercoiling just as RNAP arrived at the LacI obstacle, the DNA template was preloaded with negative turns (Figure 5.4A) under forces ranging between approximately 0.25

and 0.8 pN. As RNAP transcribes, the DNA tether length would be expected to change as depicted in Figure 5.4A. I first verified that RNAP could transcribe a tether preloaded with -24 turns (gray and black trace in Figure 5.4B). After introducing NTPs (blue arrow in Figure 5.4B), the tether extension increased, due to annihilation of the negative supercoiling by the positive supercoiling generated by the elongating RNAP, until the DNA tether became torsionally relaxed. After that, RNAP continued to wind the DNA tether introducing positive plectonemes until the bead was drawn to the surface of the flow-chamber, or RNAP stalled due to either steric hindrance by the plectonemes, or perhaps by large torsional stress. Once the ability of RNAP to transcribe a negatively supercoiled template by several turns was verified (Figure 5.4B, grey and black curves), transcription was recorded in the presence of LacI after pre-loading the template with -19 turns. RNAP was expected to reach the O1 binding site after having supercoiled the DNA template to $\Delta Lk = +5$, according to extension versus twist curves (Figure 2.14, left), in which the tether length with +5 turns can be clearly distinguished from that of a relaxed tether ($\Delta Lk = 0$). The blue and red traces in Figure 5.4B are the raw data and a 4 s moving average, respectively, of a measurement with 10 nm LacI. Although RNAP paused also at random positions along the trace, the expected pause at O1 was clearly distinguishable (Figure 5.4B, between vertical, black, dashed lines). For comparison, I acquired MT measurements of the RNAP pause in front of the LacI-O1 obstacle on torsionally relaxed (nicked) DNA under 0.45 pN tension (Figure 5.4C right).

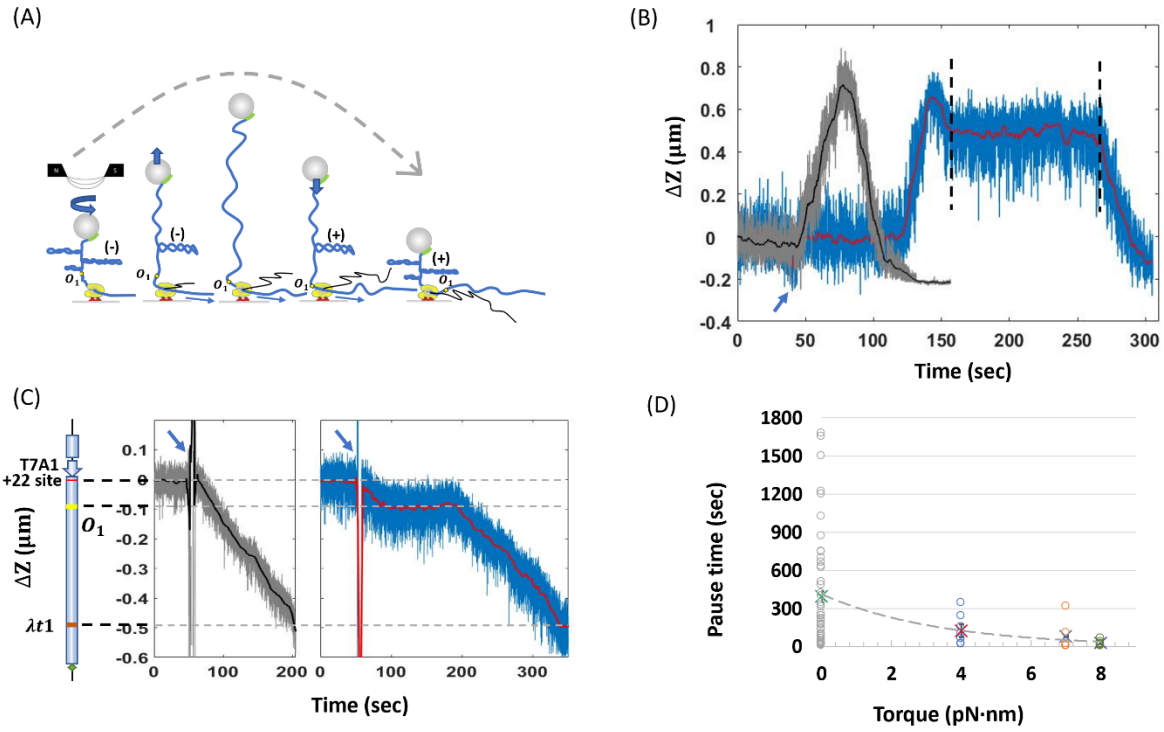


Figure 5. 4. Comparison of RNAP pause times at O1 with and without positive supercoiling. (A) A DNA tether was mechanically unwound forming plectonemes prior to the addition of NTPs. Subsequent transcription introduced positive supercoils that annihilated the mechanically induced, negative supercoils and lengthened the tether to a maximum. Further transcription and positive supercoiling eventually produced plectonemes that contracted the tether length. The dashed gray curve indicates the extension of the DNA tether during progressive conversion of negative to positive plectonemes due to transcription by RNAP. (B) Representative observations of extension versus time were recorded during transcription without (gray) and with (blue and red) LacI on a template pre-loaded with negative supercoiling. The blue arrow indicates the time at which all four NTPs or NTPs+LacI were introduced. The two vertical black dashed lines circumscribe a pause by RNAP at the LacI-O1 operator complex. (C) Representative observations of elongation were recorded in the absence of pre-loaded supercoiling without (left panel, gray and black) or with (right panel, blue and red) LacI. The cartoon on the left of the two traces is a schematic representation of the DNA template (D) Pause times by RNAP at the LacI obstacle varied as a function of torque on the DNA. Torque values were calculated using expression in table 1 equation 7, except that C_{eff}

was estimated in 100 mM [Na+] instead of 50 mM [K+]. Circles represent measured pause times, while crosses represent the average pause times.

Figure 5.4D shows how pause times changed with different torque on the DNA. The average RNAP pause at the LacI-O1 obstacle with no torque ($\tau_1 = 0$ pN·nm) (data: gray circles, average: green cross) was 393 ± 64 s (N=49), much longer than the average pause time with positive torques $\tau_2 \sim 4$ pN·nm (data: blue circles, average: red cross), $\tau_3 \sim 7$ pN·nm (data: orange circles, average blue cross) and $\tau_4 \sim 8$ pN·nm (data: green circles, average: purple cross), which pause RNAP for 125 ± 42 s (N=8), 82 ± 50 s (N=6) and 26 ± 8 s (N=8), respectively. Thus, positive supercoiling significantly facilitated transcription through the Lac-O1 obstacle and likely disrupted the LacI-O2 obstacle as well. The decrease in pause duration as positive torque on the DNA increased suggests that positive supercoiling weakens LacI binding and transcriptional roadblocking. Positive supercoiling generated by RNAP translocation is also known to destabilize nucleosomes [110, 149]; it may, therefore, represent one means by which RNAP removes protein roadblocks along the DNA.

§5.4 Discussion and conclusion

With this work I obtained evidence that the ability of DNA-bound LacI to act as barrier for the transcribing RNAP is strongly dependent on DNA topology. The TPM measurements show that a DNA loop formed by LacI bridging two operator sequences alters the roadblocking effect of LacI-operator obstacles in opposite ways depending on their position relative to the promoter (proximal vs. distal). Approaching a loop from the outside, RNAP paused in front of LacI-proximal operator roadblock longer than it did when there was no loop. Such increases in pause duration are likely due to the fact that the loop effectively increases the LacI concentration in the vicinity of the operator [27]. This increases the occupancy of the O_{prox} binding site and may sterically hinder approaching RNAP [148]. Once RNAP clears the proximal operator, it is

dramatically decelerated within the 400 bp looped segment, as compared to transcription of the same DNA in an unlooped configuration, likely as a result of torsionally induced stalling. Yet, once RNAP transcribes toward the end of the loop, it clears the distal LacI roadblock faster than the same roadblock in the absence of DNA loop.

Torsional disruption of LacI-DNA complexes might explain the shorter pauses at distal LacI securing a loop. Indeed, using magnetic tweezers to arrange positive supercoiling of the DNA to coincide with the RNAP arrival at the distal LacI obstacle shortened pauses by RNAP. This is strong evidence that positive supercoiling generated by transcription facilitates clearance of the LacI obstacle. In general, accumulated positive supercoiling ahead of RNAP may accelerate protein dissociation from DNA and shorten pauses at protein-mediated loops or other DNA structures.

Further control experiments using topoisomerase IB, or a nicking enzyme targeted to the loop region to artificially release the accumulating torsion, were not productive due to the inability to synchronize activities of those enzymes with RNAP elongation and nicking of either DNA strand induced RNAP pausing, or undesired transcription initiation, at nicks. Nonetheless, the results reported in this work strongly indicate that small loops of few hundred base pairs, such as the one considered here and those induced by many prokaryotic regulators, significantly slow transcription by RNAP, and that the positive torsional stress accumulated ahead of a transcription elongation complex may help to clear the path.

Generally, destabilization of DNA roadblocks as positive supercoiling accumulates is likely to enhance RNAP progression along DNA with bound proteins. Simultaneously, negative supercoiling trailing the transcription complex may help dislodged proteins rebind and/or may stabilize proteins behind the complex. *In vivo*, transcription would generate supercoiling at rates of 3.9 - 5.5 turns/sec (39 – 55 bp/s) [150], a potent source of supercoiling for topoisomerases to manage. Looping transcription factors that can shift between sites ahead and behind transcription complexes would maintain at least one connection to the DNA as RNAP passes and avoid diffusing away from their binding sites. This adds a fine level of control to that exerted by the overall concentration of the transcription factors.

In vivo, the ability of protein-mediated loops to hinder RNAP elongation may be a critical factor in the regulation of transcription at the local level. In the eukaryotic organism *Drosophila melanogaster*, it was observed, using 4C-seq assays, that RNA polymerase II was often paused near promoters involved in long-range interactions via several kilobase pair-long loops with enhancers. The authors hypothesized that since promoter-proximal complexes can exert enhancer-blocking activity [151], the presence of paused polymerase could safeguard against premature transcriptional activation, and yet keep the system poised for activation [152]. It is possible that in the case of regulatory loops found along the template during elongation, RNAP pausing, either in front or within, is part of a mechanism to (i) wait for additional factors to resolve the loop, or relieve supercoiling, as a signal to restart transcription, or (ii) avoid over transcription of a gene. Given the ubiquity of looping in any genome, I propose that RNAP may temporarily stall inside loops where it may be able to respond to regulatory factors and eventually transit the loop segment dispersing protein-DNA obstacles through the generation of positive supercoiling. In this way, RNAP can carve its path through DNA structures and DNA-bound proteins that are necessary for genome maintenance and regulation of gene expression in all life.

Conclusion

My work focused on the interaction between supercoiling, transcription and LacI mediated looping with the ultimate goal of addressing the questions of how RNA polymerase travels along roadblock-coated and topologically complex DNA.

The first project did not involve transcription but focused on the effect of supercoiling on LacI-mediated looping. This work was in anticipation of the following investigation of transcription through loops. It showed that negative supercoiling plays a key role in regulating looping activity and makes it deterministic within the lifetime of the cell under various conditions of tension and torsion that likely develop during the cell cycle. The stable loop and unloop ratio that DNA unwinding ensures can also tune gene expression in vivo since protein-mediated DNA looping is part of the network of signals that regulates all aspects of DNA metabolism, including transcription. As one of the most common DNA topologies, looping can, however, also be a transcription roadblock. The following investigation of transcription through loops provided insight into the (i) strength of a protein roadblock which secures a loop, (ii) challenges of transcribing within a 400 bp, LacI-mediated loop and (iii) mechanism by which RNAP can exit the loop. The data indicate that RNAP progress inside the loop region is delayed not by a slower transcription rate but by frequent and rather long pauses likely due to the accumulation of positive supercoiling ahead of RNAP.

In summary, I first reviewed the current knowledge on the energetics of different DNA configurations that may be commonly encountered in vivo and summarized it in table format. This will facilitate the quantitative understanding of the relationship between genomic structure and function from the standpoint of energy cost.

Then I found that negative supercoiling stabilizes the LacI-mediated loop and turns it into a strong roadblock, while positive supercoiling destabilizes LacI binding to DNA facilitating the exit of RNAP trapped inside a loop. Both positive and negative supercoiling, therefore, play key roles in regulating how RNAP navigates complex topologies, such as protein-mediated loops. In this way, my work links enzyme

activities, DNA conformational changes and protein binding/regulation together. Additionally, my work is the first where a single RNAP has been tracked in real time while it transcribed through a DNA loop and allowed quantification of the dynamics of this process. Thus, it contributes to the understanding of the mechanism of gene expression and regulation.

References

[1] Liu, L. F., and Wang, J. C. (1987) Supercoiling of the DNA template during transcription, *Proc Natl Acad Sci U S A 84*, 7024-7027.

[2] Cournac, A., and Plumbridge, J. (2013) DNA looping in prokaryotes: experimental and theoretical approaches, *Journal of bacteriology 195*, 1109-1119.

[3] Hao, N., Sullivan, A. E., Shearwin, K. E., and Dodd, I. B. (2021) The loopometer: a quantitative in vivo assay for DNA-looping proteins, *Nucleic acids research 49*, e39-e39.

[4] Schleif, R. (1992) DNA LOOPING, *Annual Review of Biochemistry 61*, 199-223.

[5] Dodd, I. B., and Egan, J. B. (2002) Action at a distance in CI repressor regulation of the bacteriophage 186 genetic switch, *Molecular microbiology 45*, 697-710.

[6] Dodd, I. B., and Egan, J. B. (1996) DNA binding by the coliphage 186 repressor protein CI, *J Biol Chem 271*, 11532-11540.

[7] Whitson, P. A., Hsieh, W. T., Wells, R. D., and Matthews, K. S. (1987) Supercoiling facilitates lac operator-repressor-pseudooperator interactions, *J Biol Chem 262*, 4943-4946.

[8] Kannan, S., Kohlhoff, K., and Zacharias, M. (2006) B-DNA under stress: over- and untwisting of DNA during molecular dynamics simulations, *Biophys J 91*, 2956-2965.

[9] Essevaz-Roulet, B., Bockelmann, U., and Heslot, F. (1997) Mechanical separation of the complementary strands of DNA, *Proc Natl Acad Sci U S A 94*, 11935-11940.

[10] Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B., and Mann, R. S. (2010) Origins of specificity in protein-DNA recognition, *Annu Rev Biochem 79*, 233-269.

[11] Wang, A. H., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., and Rich, A. (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution, *Nature 282*, 680-686.

[12] Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V., and Kopka, M. L. (1982) The anatomy of A-, B-, and Z-DNA, *Science 216*, 475-485.

[13] Pohl, F. M., and Jovin, T. M. (1972) Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC), *J Mol Biol 67*, 375-396.

[14] Nordheim, A., Lafer, E. M., Peck, L. J., Wang, J. C., Stollar, B. D., and Rich, A. (1982) Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding, *Cell 31*, 309-318.

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

[16] Wang, H. Y., Elston, T., Mogilner, A., and Oster, G. (1998) Force generation in RNA polymerase, *Biophys J 74*, 1186-1202.

[17] Kornberg, A., and Baker, T. A. (1992) *DNA replication*, Vol. 3, Wh Freeman New York.

[18] Morin, J. A., Cao, F. J., Lazaro, J. M., Arias-Gonzalez, J. R., Valpuesta, J. M., Carrascosa, J. L., Salas, M., and Ibarra, B. (2015) Mechano-chemical kinetics of DNA replication: identification of the translocation step of a replicative DNA polymerase, *Nucleic acids research 43*, 3643-3652.

[19] White, J. H. (1969) Self-linking and the Gauss integral in higher dimensions, *American journal of mathematics 91*, 693-728.

[20] Naughton, C., Avlonitis, N., Corless, S., Prendergast, J. G., Mati, I. K., Eijk, P. P., Cockroft, S. L., Bradley, M., Ylstra, B., and Gilbert, N. (2013) Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures, *Nature Structural & Molecular Biology 20*, 387-395.

[21] Lal, A., Dhar, A., Trostel, A., Kouzine, F., Seshasayee, A. S., and Adhya, S. (2016) Genome scale patterns of supercoiling in a bacterial chromosome, *Nat. Commun. 7*, 11055.

[22] Fulcrand, G., Dages, S., Zhi, X., Chapagain, P., Gerstman, B. S., Dunlap, D., and Leng, F. (2016) DNA supercoiling, a critical signal regulating the basal expression of the lac operon in Escherichia coli, *Sci Rep 6*, 19243.
[23] Zachmann, D. W. (1979) Nonlinear analysis of a twisted axially loaded elastic rod, *Quarterly of Applied Mathematics 37*, 67-72.
[24] Vinograd, J., Lebowitz, J., Radloff, R., Watson, R., and Laipis, P. (1965) The twisted circular form of polyoma viral DNA, *Proc Natl Acad Sci U S A 53*, 1104-1111.
[25] Kimura, K., and Hirano, T. (1997) ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation, *Cell 90*, 625-634.
[26] Parker, C. N., and Halford, S. E. (1991) Dynamics of long-range interactions on DNA: the speed of synapsis during site-specific recombination by resolvase, *Cell 66*, 781-791.
[27] Priest, D. G., Cui, L., Kumar, S., Dunlap, D. D., Dodd, I. B., and Shearwin, K. E. (2014) Quantitation of the DNA tethering effect in long-range DNA looping in vivo and in vitro using the Lac and lambda repressors, *Proc Natl Acad Sci U S A 111*, 349-354.
[28] Yan, Y., Ding, Y., Leng, F., Dunlap, D., and Finzi, L. (2018) Protein-mediated loops in supercoiled DNA create large topological domains, *Nucleic acids research 46*, 4417-4424.
[29] Yan, Y., Leng, F., Finzi, L., and Dunlap, D. (2018) Protein-mediated looping of DNA under tension requires supercoiling, *Nucleic acids research 46*, 2370-2379.
[30] Marko, J. F., and Neukirch, S. (2012) Competition between curls and plectonemes near the buckling transition of stretched supercoiled DNA, *Physical Review E 85*.
[31] van Loenhout, M. T. J., de Grunt, M. V., and Dekker, C. (2012) Dynamics of DNA Supercoils, *Science 338*, 94-97.
[32] Kim, S. H., Ganji, M., Kim, E., van der Torre, J., Abbondanzieri, E., and Dekker, C. (2018) DNA sequence encodes the position of DNA supercoils, *Elife 7*.
[33] Brutzer, H., Luzzietti, N., Klaue, D., and Seidel, R. (2010) Energetics at the DNA Supercoiling Transition, *Biophysical Journal 98*, 1267-1276.
[34] Desai, P. R., Brahmachari, S., Marko, J. F., Das, S., and Neuman, K. C. (2020) Coarse-grained modelling of DNA plectoneme pinning in the presence of base-pair mismatches, *Nucleic acids research 48*, 10713-10725.
[35] Lee, J. Y., Kim, Y. J., Lee, C., Lee, J. G., Yagyu, H., Tabata, O., and Kim, D. N. (2019) Investigating the sequence-dependent mechanical properties of DNA nicks for applications in twisted DNA nanostructure design, *Nucleic acids research 47*, 93-102.
[36] Dittmore, A., Brahmachari, S., Takagi, Y., Marko, J. F., and Neuman, K. C. (2017) Supercoiling DNA Locates Mismatches, *Phys Rev Lett 119*, 147801.
[37] Lionberger, T. A., Demurtas, D., Witz, G., Dorier, J., Lillian, T., Meyhöfer, E., and Stasiak, A. (2011) Cooperative kinking at distant sites in mechanically stressed DNA, *Nucleic acids research 39*, 9820-9832.
[38] Marko, J. F., and Neukirch, S. (2013) Global force-torque phase diagram for the DNA double helix: structural transitions, triple points, and collapsed plectonemes, *Phys Rev E Stat Nonlin Soft Matter Phys 88*, 062722.
[39] Strick, T. R., Allemand, J. F., Bensimon, D., and Croquette, V. (1998) Behavior of supercoiled DNA, *Biophys J 74*, 2016-2028.
[40] Jeon, J.-H., Adamcik, J., Dietler, G., and Metzler, R. (2010) Supercoiling Induces Denaturation Bubbles in Circular DNA, *Physical Review Letters 105*, 208101.
[41] Wittig, B., Wölfl, S., Dorbic, T., Vahrson, W., and Rich, A. (1992) Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene, *The EMBO journal 11*, 4653-4663.
[42] Maruyama, A., Mimura, J., Harada, N., and Itoh, K. (2013) Nrf2 activation is associated with Z-DNA formation in the human HO-1 promoter, *Nucleic acids research 41*, 5223-5234.

[43] Balandina, A., Kamashev, D., and Rouviere-Yaniv, J. (2002) The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. DNA, RNA, and their hybrids, *J Biol Chem 277*, 27622-27628.
[44] Rouviere-Yaniv, J., Yaniv, M., and Germond, J. E. (1979) E. coli DNA binding protein HU forms nucleosomelike structure with circular double-stranded DNA, *Cell 17*, 265-274.
[45] Bensaid, A., Almeida, A., Drlica, K., and Rouviere-Yaniv, J. (1996) Cross-talk between topoisomerase I and HU in Escherichia coli, *J Mol Biol 256*, 292-300.
[46] Lavoie, B. D., Shaw, G. S., Millner, A., and Chaconas, G. (1996) Anatomy of a flexer-DNA complex inside a higher-order transposition intermediate, *Cell 85*, 761-771.
[47] Xiao, B., Johnson, R. C., and Marko, J. F. (2010) Modulation of HU-DNA interactions by salt concentration and applied force, *Nucleic acids research 38*, 6176-6185.
[48] Aki, T., Choy, H. E., and Adhya, S. (1996) Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor, *Genes Cells 1*, 179-188.
[49] Haykinson, M. J., and Johnson, R. C. (1993) DNA looping and the helical repeat in vitro and in vivo: effect of HU protein and enhancer location on Hin invertasome assembly, *EMBO J 12*, 2503-2512.
[50] Chintakayala, K., Singh, S. S., Rossiter, A. E., Shahapure, R., Dame, R. T., and Grainger, D. C. (2013) E. coli Fis protein insulates the cbpA gene from uncontrolled transcription, *PLoS Genet 9*, e1003152.
[51] Guajardo, R., and Sousa, R. (1999) Characterization of the effects of Escherichia coli replication terminator protein (Tus) on transcription reveals dynamic nature of the Tus block to transcription complex progression, *Nucleic acids research 27*, 2814-2824.
[52] Kassavetis, G. A., Kaya, K. M., and Chamberlin, M. J. (1978) Escherichia coli RNA polymerase-rifampicin complexes bound at promoter sites block RNA chain elongation by Escherichia coli/RNA polymerase and T7-specific RNA polymerase, *Biochemistry 17*, 5798-5804.
[53] Widom, J. R., Rai, V., Rohlman, C. E., and Walter, N. G. (2019) Versatile transcription control based on reversible dCas9 binding, *RNA 25*, 1457-1469.
[54] Killian, J. L., Li, M., Sheinin, M. Y., and Wang, M. D. (2012) Recent advances in single molecule studies of nucleosomes, *Current Opinion In Structural Biology 22*, 80-87.
[55] Kim, S., Brostromer, E., Xing, D., Jin, J. S., Chong, S. S., Ge, H., Wang, S. Y., Gu, C., Yang, L. J., Gao, Y. Q., Su, X. D., Sun, Y. J., and Xie, X. S. (2013) Probing Allostery Through DNA, *Science 339*, 816-819.
[56] Dillon, S. C., and Dorman, C. J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression, *Nature Reviews Microbiology 8*, 185-195.
[57] Krogh, T. J., Møller-Jensen, J., and Kaleta, C. (2018) Impact of Chromosomal Architecture on the Function and Evolution of Bacterial Genomes, *Front Microbiol 9*, 2019-2019.
[58] Verma, S. C., Qian, Z., and Adhya, S. L. (2019) Architecture of the Escherichia coli nucleoid, *PLOS Genetics 15*, e1008456.
[59] Qian, J., Dunlap, D., and Finzi, L. (2021) Basic mechanisms and kinetics of pause-interspersed transcript elongation, *Nucleic acids research 49*, 15-24.
[60] Lewis, M. (2005) The lac repressor, *C R Biol 328*, 521-548.
[61] Oehler, S., Eismann, E. R., Kramer, H., and Muller-Hill, B. (1990) The three operators of the lac operon cooperate in repression, *EMBO J 9*, 973-979.
[62] Han, L., Lui, B. H., Blumberg, S., Beausang, J. F., Nelson, P. C., and Phillips, R. (2009) Calibration of Tethered Particle Motion Experiements, In *Mathematics of DNA Structure, Function and Interactions* (Benham, C. J., Harvey, S., Olson, W. K., Sumners, D. W. L., and Swigon, D., Eds.), pp 123-138, Springer, New York.
[63] Kumar, S., Manzo, C., Zurla, C., Ucuncuoglu, S., Finzi, L., and Dunlap, D. (2014) Enhanced Tethered-Particle Motion Analysis Reveals Viscous Effects, *Biophysical Journal 106*, 399-409.

[64] Kovari, D. T., Yan, Y., Finzi, L., and Dunlap, D. (2018) Tethered Particle Motion: An Easy Technique for Probing DNA Topology and Interactions with Transcription Factors, In *Single Molecule Analysis: Methods and Protocols, 2nd Edition* (Peterman, E. J. G., Ed.), pp 317-340.
[65] Priest, D. G., Kumar, S., Yan, Y., Dunlap, D. D., Dodd, I. B., and Shearwin, K. E. (2014) Quantitation of interactions between two DNA loops demonstrates loop domain insulation in E. coli cells, *Proceedings of the National Academy of Sciences of the United States of America 111*, E4449-E4457.
[66] Ucuncuoglu, S., Schneider, D. A., Weeks, E. R., Dunlap, D., and Finzi, L. (2017) Chapter Sixteen - Multiplexed, Tethered Particle Microscopy for Studies of DNA-Enzyme Dynamics, In *Methods in Enzymology* (Spies, M., and Chemla, Y. R., Eds.), pp 415-435, Academic Press.
[67] Normanno, D., Vanzi, F., and Pavone, F. S. (2008) Single-molecule manipulation reveals supercoiling-dependent modulation of lac repressor-mediated DNA looping, *Nucleic Acids Research 36*, 2505-2513.
[68] Lia, G., Bensimon, D., Croquette, V., Allemand, J. F., Dunlap, D., Lewis, D. E. A., Adhya, S. C., and Finzi, L. (2003) Supercoiling and denaturation in Gal repressor/heat unstable nucleoid protein (HU)-mediated DNA looping, *Proceedings Of The National Academy Of Sciences Of The United States Of America 100*, 11373-11377.
[69] Lia, G., Praly, E., Ferreira, H., Stockdale, C., Tse-Dinh, Y. C., Dunlap, D., Croquette, V., Bensimon, D., and Owen-Hughes, T. (2006) Direct observation of DNA distortion by the RSC complex, *Molecular Cell 21*, 417-425.
[70] Lia, G., Semsey, S., Lewis, D. E., Adhya, S., Bensimon, D., Dunlap, D., and Finzi, L. (2008) The antiparallel loops in gal DNA, *Nucleic Acids Res 36*, 4204-4210.
[71] Gosse, C., and Croquette, V. (2002) Magnetic tweezers: micromanipulation and force measurement at the molecular level, *Biophys J 82*, 3314-3329.
[72] Parthasarathy, R. (2012) Rapid, accurate particle tracking by calculation of radial symmetry centers, *Nat Methods 9*, 724-726.
[73] Hatfield, G. W., and Benham, C. J. (2002) DNA topology-mediated control of global gene expression in Escherichia coli, *Annual Review Of Genetics 36*, 175-203.
[74] Xu, W., Dunlap, D., and Finzi, L. (2021) Energetics of twisted DNA topologies, *Biophys J 120*, 3242-3252.
[75] Sheinin, M. Y., Forth, S., Marko, J. F., and Wang, M. D. (2011) Underwound DNA under tension: structure, elasticity, and sequence-dependent behaviors, *Phys Rev Lett 107*, 108102.
[76] Oberstrass, F. C., Fernandes, L. E., and Bryant, Z. (2012) Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA, *Proc Natl Acad Sci U S A 109*, 6106-6111.
[77] SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proc Natl Acad Sci U S A 95*, 1460-1465.
[78] Fuertes, M. A., Cepeda, V., Alonso, C., and Perez, J. M. (2006) Molecular mechanisms for the B-Z transition in the example of poly d(G-C)center dot d(G-C) polymers. A critical review, *Chemical Reviews 106*, 2045-2064.
[79] Jost, D., and Everaers, R. (2009) A unified Poland-Scheraga model of oligo- and polynucleotide DNA melting: salt effects and predictive power, *Biophys J 96*, 1056-1067.
[80] Bauer, W. R., and Benham, C. J. (1993) The free energy, enthalpy and entropy of native and of partially denatured closed circular DNA, *Journal of Molecular Biology 234*, 1184-1196.
[81] Benham, C. J. (1992) Energetics of the Strand Separation Transition in Superhelical DNA *Journal of Molecular Biology 225*, 835-847.
[82] Fye, R. M., and Benham, C. J. (1999) Exact method for numerically analyzing a model of local denaturation in superhelically stressed DNA, *Physical Review E 59*, 3408.
[83] Ho, P. S. (1994) The non-B-DNA Structure of d(CA/TG)(N) does not Differ from that of Z-DNA, *Proceedings of the National Academy of Sciences of the United States of America 91*, 9549-9553.

[84] Peck, L. J., and Wang, J. C. (1983) Energetics of B-to-Z Transition in DNA, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 80*, 6206-6210.
[85] Oberstrass, F. C., Fernandes, L. E., Lebel, P., and Bryant, Z. (2013) Torque spectroscopy of DNA: base-pair stability, boundary effects, backbending, and breathing dynamics, *Phys Rev Lett 110*, 178103.
[86] Kim, S. H., Lim, S. H., Lee, A. R., Kwon, D. H., Song, H. K., Lee, J. H., Cho, M., Johner, A., Lee, N. K., and Hong, S. C. (2018) Unveiling the pathway to Z-DNA in the protein-induced B-Z transition, *Nucleic acids research 46*, 4129-4137.
[87] Marko, J. F., and Siggia, E. D. J. M. (1994) Bending and twisting elasticity of DNA, *27*, 981-988.
[88] Skoruppa, E., Nomidis, S. K., Marko, J. F., and Carlon, E. (2018) Bend-Induced Twist Waves and the Structure of Nucleosomal DNA, *Phys Rev Lett 121*, 088101.
[89] Skoruppa, E., Laleman, M., Nomidis, S. K., and Carlon, E. (2017) DNA elasticity from coarse-grained simulations: The effect of groove asymmetry, *J Chem Phys 146*, 214902.
[90] Nomidis, S. K., Kriegel, F., Vanderlinden, W., Lipfert, J., and Carlon, E. (2017) Twist-Bend Coupling and the Torsional Response of Double-Stranded DNA, *Phys Rev Lett 118*, 217801.
[91] Nomidis, S. K., Skoruppa, E., Carlon, E., and Marko, J. F. (2019) Twist-bend coupling and the statistical mechanics of the twistable wormlike-chain model of DNA: Perturbation theory and beyond, *Phys Rev E 99*, 032414.
[92] Marko, J. F. (1998) DNA under high tension: overstretching, undertwisting, and relaxation dynamics, *Physical Review E 57*, 2134.
[93] Gore, J., Bryant, Z., Nollmann, M., Le, M. U., Cozzarelli, N. R., and Bustamante, C. (2006) DNA overwinds when stretched, *Nature 442*, 836-839.
[94] Sheinin, M. Y., and Wang, M. D. (2009) Twist-stretch coupling and phase transition during DNA supercoiling, *Phys Chem Chem Phys 11*, 4800-4803.
[95] Liebl, K., Drsata, T., Lankas, F., Lipfert, J., and Zacharias, M. (2015) Explaining the striking difference in twist-stretch coupling between DNA and RNA: A comparative molecular dynamics analysis, *Nucleic acids research 43*, 10143-10156.
[96] Lipfert, J., Skinner, G. M., Keegstra, J. M., Hensgens, T., Jager, T., Dulin, D., Kober, M., Yu, Z., Donkers, S. P., Chou, F. C., Das, R., and Dekker, N. H. (2014) Double-stranded RNA under force and torque: similarities to and striking differences from double-stranded DNA, *Proc Natl Acad Sci U S A 111*, 15408-15413.
[97] Geggier, S., and Vologodskii, A. (2010) Sequence dependence of DNA bending rigidity, *Proc Natl Acad Sci U S A 107*, 15421-15426.
[98] Reymer, A., Zakrzewska, K., and Lavery, R. (2018) Sequence-dependent response of DNA to torsional stress: a potential biological regulation mechanism, *Nucleic acids research 46*, 1684-1694.
[99] Balasubramanian, S., Xu, F., and Olson, W. K. (2009) DNA sequence-directed organization of chromatin: structure-based computational analysis of nucleosome-binding sequences, *Biophys J 96*, 2245-2260.
[100] Morozov, A. V., Fortney, K., Gaykalova, D. A., Studitsky, V. M., Widom, J., and Siggia, E. D. (2009) Using DNA mechanics to predict in vitro nucleosome positions and formation energies, *Nucleic acids research 37*, 4707-4722.
[101] Daniels, B. C., and Sethna, J. P. (2011) Nucleation at the DNA supercoiling transition, *Physical Review E 83*, 041924.
[102] Dittmore, A., Silver, J., and Neuman, K. C. (2018) Kinetic Pathway of Torsional DNA Buckling, *J Phys Chem B 122*, 11561-11570.
[103] Walker, P. U., Vanderlinden, W., and Lipfert, J. (2018) Dynamics and energy landscape of DNA plectoneme nucleation, *Physical Review E 98*.
[104] Bustamante, C., Bryant, Z., and Smith, S. B. (2003) Ten years of tension: single-molecule DNA mechanics, *Nature 421*, 423-427.

[105] Moroz, J. D., and Nelson, P. (1997) Torsional directed walks, entropic elasticity, and DNA twist stiffness, *Proceedings of the National Academy of Sciences of the United States of America 94*, 14418-14422.
[106] Matek, C., Ouldridge, T. E., Doye, J. P., and Louis, A. A. (2015) Plectoneme tip bubbles: coupled denaturation and writhing in supercoiled DNA, *Sci Rep 5*, 7655.
[107] Crothers, D. M., and Spatz, H. C. (1971) Theory of friction-limited DNA unwinding, *Biopolymers 10*, 1949-1972.
[108] Thomas, T. J., and Bloomfield, V. A. (1983) Chain flexibility and hydrodynamics of the B and Z forms of poly(dG-dC).poly(dG-dC), *Nucleic acids research 11*, 1919-1930.
[109] Oberstrass, F. C., Fernandes, L. E., and Bryant, Z. (2012) Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA, *Proceedings Of The National Academy Of Sciences Of The United States Of America 109*, 6106-6111.
[110] Sheinin, M. Y., Li, M., Soltani, M., Luger, K., and Wang, M. D. (2013) Torque modulates nucleosome stability and facilitates H2A/H2B dimer loss, *Nat Commun 4*, 2579.
[111] Ma, J., Bai, L., and Wang, M. D. (2013) Transcription under torsion, *Science 340*, 1580-1583.
[112] Champoux, J. J. (2001) DNA topoisomerases: Structure, function, and mechanism, *Annual Review of Biochemistry 70*, 369-413.
[113] Corbett, K. D., and Berger, J. M. (2004) Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases, *Annual Review of Biophysics and Biomolecular Structure 33*, 95-118.
[114] Gellert, M., Mizuuchi, K., Odea, M. H., and Nash, H. A. (1976) DNA Gyrase - Enzyme that Introduces Superhelical Turns into DNA, *Proceedings of the National Academy of Sciences of the United States of America 73*, 3872-3876.
[115] Heberling, T., Davis, L., Gedeon, J., Morgan, C., and Gedeon, T. (2016) A Mechanistic Model for Cooperative Behavior of Co-transcribing RNA Polymerases, *PLoS Comput Biol 12*, e1005069.
[116] Kim, S., Beltran, B., Irnov, I., and Jacobs-Wagner, C. (2019) Long-Distance Cooperative and Antagonistic RNA Polymerase Dynamics via DNA Supercoiling, *Cell 179*, 106-119 e116.
[117] Kouzine, F., Liu, J., Sanford, S., Chung, H. J., and Levens, D. (2004) The dynamic response of upstream DNA to transcription-generated torsional stress, *Nat Struct Mol Biol 11*, 1092-1100.
[118] El Hanafi, D., and Bossi, L. (2000) Activation and silencing of leu-500 promoter by transcription-induced DNA supercoiling in the Salmonella chromosome, *Mol Microbiol 37*, 583-594.
[119] Tabuchi, H., Handa, H., and Hirose, S. (1993) Underwinding of DNA on binding of yeast TFIID to the TATA element, *Biochem Biophys Res Commun 192*, 1432-1438.
[120] Lankas, F., Sponer, J., Langowski, J., and Cheatham, T. E., 3rd. (2003) DNA basepair step deformability inferred from molecular dynamics simulations, *Biophys J 85*, 2872-2883.
[121] Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M., and Zhurkin, V. B. (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes, *Proc Natl Acad Sci U S A 95*, 11163-11168.
[122] Friedel, M., Nikolajewa, S., Sühnel, J., and Wilhelm, T. (2009) DiProDB: a database for dinucleotide properties, *Nucleic acids research 37*, D37-D40.
[123] Shahmuradov, I. A., Mohamad Razali, R., Bougouffa, S., Radovanovic, A., and Bajic, V. B. (2017) bTSSfinder: a novel tool for the prediction of promoters in cyanobacteria and Escherichia coli, *Bioinformatics 33*, 334-340.
[124] Ravichandran, S., Subramani, V. K., and Kim, K. K. (2019) Z-DNA in the genome: from structure to disease, *Biophysical Reviews 11*, 383-387.
[125] Lafer, E. M., Valle, R. P., Möller, A., Nordheim, A., Schur, P. H., Rich, A., and Stollar, B. D. (1983) Z-DNA-specific antibodies in human systemic lupus erythematosus, *The Journal of Clinical Investigation 71*, 314-321.
[126] Wang, G., Christensen, L. A., and Vasquez, K. M. (2006) Z-DNA-forming sequences generate large-scale deletions in mammalian cells, *Proc Natl Acad Sci U S A 103*, 2677-2682.

[127] Wong, B., Chen, S., Kwon, J.-A., and Rich, A. (2007) Characterization of Z-DNA as a nucleosome-boundary element in yeast Saccharomyces cerevisiae, *Proceedings of the National Academy of Sciences 104*, 2229-2234.
[128] Marshall, P. R., Zhao, Q., Li, X., Wei, W., Periyakaruppiah, A., Zajaczkowski, E. L., Leighton, L. J., Madugalle, S. U., Basic, D., Wang, Z., Yin, J., Liau, W. S., Gupte, A., Walkley, C. R., and Bredy, T. W. (2020) Publisher Correction: Dynamic regulation of Z-DNA in the mouse prefrontal cortex by the RNA-editing enzyme Adar1 is required for fear extinction (Nature Neuroscience, (2020), 23, 6, (718-729), 10.1038/s41593-020-0627-5), *Nature Neuroscience 23*, 1034.
[129] Yan, Y., Xu, W., Kumar, S., Zhang, A., Leng, F., Dunlap, D., and Finzi, L. (2021) Negative DNA supercoiling makes protein-mediated looping deterministic and ergodic within the bacterial doubling time, *Nucleic acids research 49*, 11550-11559.
[130] Finzi, L., and Gelles, J. (1995) Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules, *Science 267*, 378-380.
[131] Han, L., Garcia, H. G., Blumberg, S., Towles, K. B., Beausang, J. F., Nelson, P. C., and Phillips, R. (2009) Concentration and length dependence of DNA looping in transcriptional regulation, *PLoS ONE 4*, e5621.
[132] Berger, M., Gerganova, V., Berger, P., Rapiteanu, R., Lisicovas, V., and Dobrindt, U. (2016) Genes on a Wire: The Nucleoid-Associated Protein HU Insulates Transcription Units in Escherichia coli, *Sci Rep 6*, 12.
[133] Morales, P., Rouviere-Yaniv, J., and Dreyfus, M. (2002) The histone-like protein HU does not obstruct movement of T7 RNA polymerase in Escherichia coli cells but stimulates its activity, *Journal of Bacteriology 184*, 1565-1570.
[134] Bahloul, A., Boubrik, F., and Rouviere-Yaniv, J. (2001) Roles of Escherichia coli histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of dnaA46ts, *Biochimie 83*, 219-229.
[135] Nir, G., Lindner, M., Dietrich, H. R. C., Girshevitz, O., Vorgias, C. E., and Garini, Y. (2011) HU Protein Induces Incoherent DNA Persistence Length, *Biophysical Journal 100*, 784-790.
[136] Sagi, D., Friedman, N., Vorgias, C., Oppenheim, A. B., and Stavans, J. (2004) Modulation of DNA conformations through the formation of alternative high-order HU-DNA complexes, *Journal of Molecular Biology 341*, 419-428.
[137] van Noort, J., Verbrugge, S., Goosen, N., Dekker, C., and Dame, R. T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments, *Proc Natl Acad Sci U S A 101*, 6969-6974.
[138] Kouzine, F., Gupta, A., Baranello, L., Wojtowicz, D., Ben-Aissa, K., Liu, J. H., Przytycka, T. M., and Levens, D. (2013) Transcription-dependent dynamic supercoiling is a short-range genomic force, *Nature Structural & Molecular Biology 20*, 396-403.
[139] Finzi, L., and Dunlap, D. D. (2010) Single-molecule approaches to probe the structure, kinetics, and thermodynamics of nucleoprotein complexes that regulate transcription, *J Biol Chem 285*, 18973-18978.
[140] Gibson, B., Wilson, D. J., Feil, E., and Eyre-Walker, A. (2018) The distribution of bacterial doubling times in the wild, *Proc Biol Sci 285*, 20180789.
[141] Guo, F., and Adhya, S. (2007) Spiral structure of Escherichia coli HUab provides foundation for DNA supercoiling, *Proc Natl Acad Sci U S A 104*, 4309-4314.
[142] Lipfert, J., Klijnhout, S., and Dekker, N. H. (2010) Torsional sensing of small-molecule binding using magnetic tweezers, *Nucleic acids research 38*, 7122-7132.
[143] Salerno, D., Brogioli, D., Cassina, V., Turchi, D., Beretta, G. L., Seruggia, D., Ziano, R., Zunino, F., and Mantegazza, F. (2010) Magnetic tweezers measurements of the nanomechanical properties of DNA in the presence of drugs, *Nucleic acids research 38*, 7089–7099.
[144] Vlijm, R., Kim, S. H., De Zwart, P. L., Dalal, Y., and Dekker, C. (2017) The supercoiling state of DNA determines the handedness of both H3 and CENP-A nucleosomes, *Nanoscale 9*, 1862-1870.

[145] Marklund, E., van Oosten, B., Mao, G., Amselem, E., Kipper, K., Sabantsev, A., Emmerich, A., Globisch, D., Zheng, X., Lehmann, L. C., Berg, O. G., Johansson, M., Elf, J., and Deindl, S. (2020) DNA surface exploration and operator bypassing during target search, *Nature 583*, 858-861.

[146] Leng, F., Chen, B., and Dunlap, D. D. (2011) Dividing a supercoiled DNA molecule into two independent topological domains, *Proceedings of the National Academy of Sciences 108*, 19973-19978.

[147] Moroz, J. D., and Nelson, P. (1997) Torsional directed walks, entropic elasticity, and DNA twist stiffness, *Proc Natl Acad Sci U S A 94*, 14418-14422.

[148] Voros, Z., Yan, Y., Kovari, D. T., Finzi, L., and Dunlap, D. (2017) Proteins mediating DNA loops effectively block transcription, *Protein Sci 26*, 1427-1438.

[149] Teves, S. S., and Henikoff, S. (2014) Transcription-generated torsional stress destabilizes nucleosomes, *Nat Struct Mol Biol 21*, 88-94.

[150] Vogel, U., and Jensen, K. F. (1994) The RNA chain elongation rate in Escherichia coli depends on the growth rate, *J Bacteriol 176*, 2807-2813.

[151] Chopra, V. S., Cande, J., Hong, J. W., and Levine, M. (2009) Stalled Hox promoters as chromosomal boundaries, *Genes & Development 23*, 1505-1509.

[152] Ghavi-Helm, Y., Klein, F. A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E. E. (2014) Enhancer loops appear stable during development and are associated with paused polymerase, *Nature 512*, 96-100.