

## 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:

---

Yan Yan

---

Date

# Transcription factors and supercoiling establish DNA topology that influences transcription

By

Yan Yan  
Doctor of Philosophy

Physics

---

Laura Finzi  
Advisor

---

William Kelly  
Committee Member

---

Harold Kim  
Committee Member

---

Minsu Kim  
Committee Member

---

Eric Weeks  
Committee Member

Accepted:

---

Lisa A. Tedesco, Ph.D.  
Dean of the James T. Laney School of Graduate Studies

---

Date

# **Transcription factors and supercoiling establish DNA topology that influences transcription**

By

Yan Yan  
B. S., Shandong University, China, 2011

Advisor: Laura Finzi, Ph. D.

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  
**2018**

# Abstract

## Transcription factors and supercoiling establish DNA topology that influences transcription

By Yan Yan

Protein-mediated DNA looping is ubiquitous in chromatin organization and gene regulation both in eukaryotes and prokaryotes. It occurs when one or more protein(s) bridge two distant sites on the double helix. In principle, different protein-mediated loops may, interact giving rise to more complex regulation patterns. Using the tethered particle motion technique, the probability of looping by the lac repressor protein (LacI) was shown to depend on LacI concentration, loop size and binding affinity of operators. Furthermore, it was shown that DNA loops mediated by LacI and the  $\lambda$  bacteriophage repressor protein,  $\lambda$  CI, interact differently depending on the topological arrangement: side-by-side loops do not affect each other, nested loops assist each other's formation, while alternating loops inhibit each other's formation. These observations provided clear support for the loop domain model for insulation.

Nucleoid associated protein (NAPs) and/or supercoiling could both be responsible for the fact that in vivo levels of LacI-mediated looping are higher than those measured in vitro for a range of large loop sizes in identical DNA templates. Using magnetic tweezers, physiological levels of negative supercoiling were shown to drive the looping probability from 0 to 100 % under slight tension that likely exists in vivo. In contrast, even saturating (micromolar) concentrations of HU couldn't raise the looping probability above 30 % in similarly stretched DNA or 80% in DNA without tension. Furthermore, it was shown that loops that formed in supercoiled DNA create topological domains that may exceed the loop segment length (distance between protein binding sites). This is relevant to regulation by distant elements.

Magnetic tweezers were also used to show that RNAP halts for several minutes upon encountering a LacI bound to a single operator. The average pause lifetime is compatible with RNAP waiting for LacI dissociation. Puzzlingly, RNAP seems to slide back to the promoter immediately after encountering a protein and reinitiate elongation. These observations still need to be understood.

# **Transcription factors and supercoiling establish DNA topology that influences transcription**

By

Yan Yan  
B. S., Shandong University, China, 2011

Advisor: Laura Finzi, Ph. D.

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  
**2018**

# Acknowledgement

I owe my deepest gratitude to my advisor, Prof. Laura Finzi, for her generous encouragement, guidance and support throughout my graduate academic studies. These pages would be impossible without her kind mentorship and friendship over my long stay in the Physics Department. I also sincerely thank Dr. David Dunlap for the kind advising and patient teaching.

I am heartily thankful to my Ph. D. advising committee members: Dr. William Kelly, Dr. Harold Kim, Dr. Minsu Kim, Dr. Eric Weeks for their encouraging words, suggestions for my research, thoughtful criticism, and for attending my annual meetings.

I am grateful to my previous coworkers, Dr. Sandip Kumar and Dr. Mónica Fernández. They greatly helped me mastering a variety of lab techniques and developing critical thinking with constructive discussion during my first year of research in the lab.

I would like to thank the group members, Dr. Suleyman Ucuncuoglu, Dr. Yue Ding, Dr. Diane Wiener, Dr. Daniel Kovari, Dr. Zsuzsanna Vörös, Wenxuan Xu for thoughtful discussions and for their friendship.

Many thanks also to the undergraduate students who worked with me, Alex Zhang, James Boyle, Hayes Burdette-Sapp. They facilitated my work with sample preparations, experiments, and data analysis. They all did a great job.

I am indebted to our collaborators Dr. Fenfei Leng, for generously supplying the HU protein and plasmids, and Dr. Kathleen Matthews for generously supplying the LacI protein.

I would also like to thank my parents, my father Chengchun Yan, and my mother Chongzhen Bai. They always trust, support, and encourage me to overcome difficulties in my life.

I owe much to my advisors in China, Prof. Ming Li, Prof. Chunhua Xu and their previous graduate students Dr. Zhe Pang and Dr. Shuang Wang. They are the first people who lead me into this field and taught me a lot in my early scientific research. I also thank them for their collaboration on my qualifier proposal which is not included in this dissertation.

I received guidance and assistance from many people and friends over the years. It is impossible to list and thank them all. This dissertation is dedicated to all of them as well.

# Table of contents

<b>Chapter 1 Introduction</b> .....	1
§ 1.1 DNA structure.....	1
§ 1.1.1 DNA primary structure.....	1
§ 1.1.2 DNA secondary structure and polymorphism of DNA structure.....	3
§ 1.1.3 DNA tertiary structure, DNA supercoiling.....	4
§ 1.1.4 DNA quaternary structure.....	5
§ 1.2 Introduction of HU protein.....	6
§ 1.3 Introduction of transcription.....	7
§ 1.3.1 Central dogma of molecular biology.....	7
§ 1.3.2 Components and function of RNA polymerase.....	9
§ 1.3.3 Transcription steps: initiation, elongation and termination.....	10
§ 1.3.4 Transcriptional regulation.....	14
§ 1.4 The lac repressor protein, LacI, and LacI-mediated DNA looping.....	14
§ 1.4.1 Lac Operon and structure of the LacI protein.....	14
§ 1.4.2 LacI mediated DNA looping study <i>in vivo</i> and <i>in vitro</i> .....	16
§ 1.5 DNA topological domains.....	20
§ 1.6 Experimental techniques.....	23
§ 1.6.1 Single molecule techniques.....	23
§ 1.6.2 Tethered-particle motion.....	23
§ 1.6.3 Magnetic tweezers.....	26
<b>Chapter 2 Material and Methods</b> .....	29
§ 2.1 Preparation of DNA constructs.....	29
§ 2.2 Buffer recipes.....	32
§ 2.3 Microchamber Preparation.....	34

§ 2.4 Tethered Particle Motion (TPM) experiments .....	35
§ 2.5 Magnetic tweezers (MTs) experiments.....	37
§ 2.6 Estimation of the HU concentration in an <i>E. coli</i> cell.....	41
§ 2.7 Exclusion of artifacts in the study of topological domains.....	41
§ 2.8 Measuring the coiling and extent of topological domains.....	43
§ 2.9 Obtaining the gyre size.....	43

**Chapter 3 protein concentration, operator affinity and loop size dependence of LacI-mediated DNA looping.....44**

§ 3.1 Introduction .....	44
§ 3.2 Results.....	47
§ 3.2.1 Two looped states are observed in the experimental trace .....	47
§ 3.2.2 LacI concentration dependence of protein mediated DNA looping .....	48
§ 3.2.3 Operator affinity dependence of protein mediated DNA looping.....	50
§ 3.2.4 Loop size dependence of protein mediated DNA looping .....	50
§ 3.3 Discussion.....	52

**Chapter 4 Quantitation of interactions between two DNA loops demonstrates loop domain insulation in *E. coli* cells .....57**

§ 4.1 Introduction.....	57
§ 4.2 Results.....	61
§ 4.2.1 Quantitation of interactions between DNA loops.....	61
§ 4.2.2 Alternating loops give loop interference.....	61
§ 4.2.3 Nested loops give loop assistance.....	64
§ 4.2.4 Side-by-side loops do not interact.....	64
§ 4.2.5 Model to quantitate loop interactions.....	65
§ 4.3 Discussion.....	66



<b>Chapter 5 Protein-mediated looping of DNA under slight tension requires supercoiling</b> .....	69
§ 5.1 Introduction.....	69
§ 5.2 Results.....	72
§ 5.2.1 HU decreased tether lengths and promoted LacI-mediated looping.....	72
§ 5.2.2 Supercoiling decreased tether lengths and induced LacI-mediated looping.....	76
§ 5.3 Discussion and conclusion.....	80
§ 5.3.1 Contraction, especially with writhe, enhances loop formation.....	80
§ 5.3.2 Supercoiling lowers the free energy of DNA looping in a protein-dependent manner.....	82
§ 5.3.3 Comparing the effects of HU and supercoiling on DNA looping via $J_{loop}$ factor.....	83
§ 5.4 Conclusion.....	84

<b>Chapter 6 Protein-mediated loops in supercoiled DNA create large topological domain</b> .....	86
§ 6.1 Introduction.....	86
§ 6.2 Results.....	88
§ 6.2.1 Loop-securing LacI resists torsional stress in dynamically supercoiled DNA.....	88
§ 6.2.2 LacI-mediated loops preferentially trap negative supercoiling at higher tension.....	91
§ 6.2.3 LacI-mediated loops dynamically trap large topological domains in supercoiled DNA.....	93
§ 6.2.4 Exclusion of artifacts.....	95
§ 6.3 Discussion.....	97

<b>Chapter 7 RNA polymerase pauses at <i>lac</i> repressor obstacles</b> .....	99
§ 7.1 Introduction.....	99
§ 7.2 Results.....	101
§ 7.2.1 RNA polymerase overcomes pausing in magnetic tweezer experiments.....	101
§ 7.2.2 RNA polymerase may shuttling backward once encounter obstacle.....	104
§ 7.2.3 DNA extension can reach an extension beyond the end of DNA template.....	106
§ 7.3 Discussion and Conclusion.....	107
<b>References</b> .....	109

# List of Figures

## Chapter 1 Introduction

<b>Figure 1.1</b> (A) GC and AT paring. (B) Primary structure of a 3 base-long single-stranded DNA with sequence GAC. ....	2
<b>Figure 1.2</b> Side and top view of (A) B-DNA, (B) A-DNA, (C) Z-DNA.....	3
<b>Figure 1.3</b> Conformation of DNA supercoiling.....	5
<b>Figure 1.4</b> Structural basis for DNA bending by HU protein.....	6
<b>Figure 1.5</b> Illustration of the central dogma of molecular biology.....	8
<b>Figure 1.6</b> Promoter sequence.....	11
<b>Figure 1.7</b> Transcription of two divergent genes.....	12
<b>Figure 1.8</b> Gene regulation in <i>lac</i> operon.....	15
<b>Figure 1.9</b> Examples of topological domains.....	20
<b>Figure 1.10</b> Tethered particle motion and the change of excursion length when bead shows looping transition.....	24
<b>Figure 1.11</b> Magnetic tweezers setup.....	26

## Chapter 2 Material and Methods

<b>Figure 2.1</b> The probability of labeling along the anchorage fragment as a function of the distance from the ligation junction with the main fragment. The anchorage fragments are prepared by digesting with XmaI or ApaI approximately 2 kbp-long, biotin- or digoxigenin-labeled DNA amplicons centered on the multicloning site of pBluescriptKS+.....	31
---	----

## Chapter 3 protein concentration, operator affinity and loop size dependence of LacI-mediated DNA looping

<b>Figure 3.1</b> Experimental schematics, representative data statistical result.....	46
<b>Figure 3.2</b> Calculation of $J_{loop}$ factor.....	52
<b>Figure 3.3</b> Comparison of $J_{loop}$ factor values with previous literature.....	55

## **Chapter 4 Quantitation of interactions between two DNA loops demonstrates loop domain insulation in *E. coli* cells**

<b>Figure 4.1</b> Interactions between DNA loops.....	59
<b>Figure 4.2</b> TPM analysis of alternating and nested loops.....	63

## **Chapter 5 Protein-mediated looping of DNA under slight tension requires supercoiling**

<b>Figure 5.1</b> Experimental schematics and representative data.....	73
<b>Figure 5.2</b> HU promotes looping.....	75
<b>Figure 5.3</b> Supercoiling promotes looping.....	77
<b>Figure 5.4</b> $J_{loop}$ factors decrease as DNA tethers shorten especially when the DNA becomes negative supercoiled.....	85

## **Chapter 6 Protein-mediated loops in supercoiled DNA create large topological domain.....86**

<b>Figure 6.1</b> LacI-mediated loops trap supercoiling and resist torsion.....	90
<b>Figure 6.2</b> Loops that trapped high levels of supercoiling greatly reduced the extension of tethers under tension of 0.25 (left) and 0.45 (right) pN.....	93
<b>Figure 6.3</b> Different loop topologies may trap flanking DNA.....	94

## **Chapter 7 RNA polymerase pauses at *lac* repressor obstacles**

<b>Figure 7.1</b> A schematic diagram of the magnetic tweezers transcription assay.....	101
<b>Figure 7.2</b> LacI bound to an O1 operator pauses transcription.....	103
<b>Figure 7.3</b> TEC pauses and expected operator occupancy.....	104
<b>Figure 7.4</b> RNA polymerase (RNAP) show transcription forward and shuttling back in the absence of LacI (black) and in the presence of LacI (blue).....	105
<b>Figure 7.5</b> RNA polymerase (RNAP) show transcription forward and shuttling back to an extension exceed the upstream end of the DNA template.....	106
<b>Figure 7.6</b> Possible mechanism for RNAP shuttling.....	107

# Chapter 1 Introduction

## § 1.1 DNA structure

DNA (deoxyribonucleic acid), a polymer of deoxyribonucleotides, is the repository of genetic information in cells. It has been over 60 years since Watson and Crick published the right handed helical structure of DNA in 1953 (6). The unique structure of DNA enables genetic and epi-genetic information to be stored and regulated in prokaryotes and eukaryotes. Prokaryotes are single cell organisms that lack a cell nucleus, and their DNA is compacted by supercoiling and a host of DNA organizing proteins. Eukaryotic cells, in contrast, have intracellular organelles that organize the cellular contents into compartments. In these cells, DNA and its associated proteins are contained in the nucleus. The human nucleus is approximately  $10^7 \mu\text{m}^3$  (7) and contains about three billion base pairs (bp) (about 1.8 m long when fully extended). Clearly, DNA needs to be tightly packaged to fit in the nucleus. To better understand this process, it is useful to review the four different levels of DNA structure: primary, secondary, tertiary and quaternary.

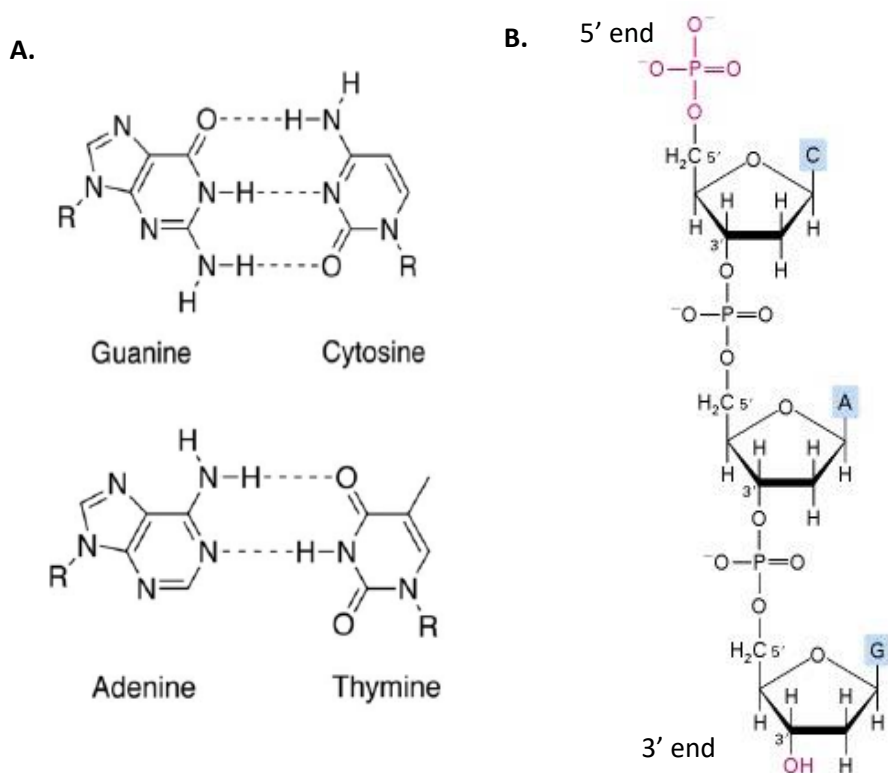
### § 1.1.1 DNA primary structure

DNA's primary structure is the linear sequence of nucleotides linked together by phosphodiester bonds. Deoxyribonucleotides consist of a pentose sugar (deoxyribose), a phosphate group, and a nitrogenous base. There are four different bases: Adenine(A),

Guanine(G), Thymine (T) and Cytosine(G). A and G are purines, C and T are pyrimidines.

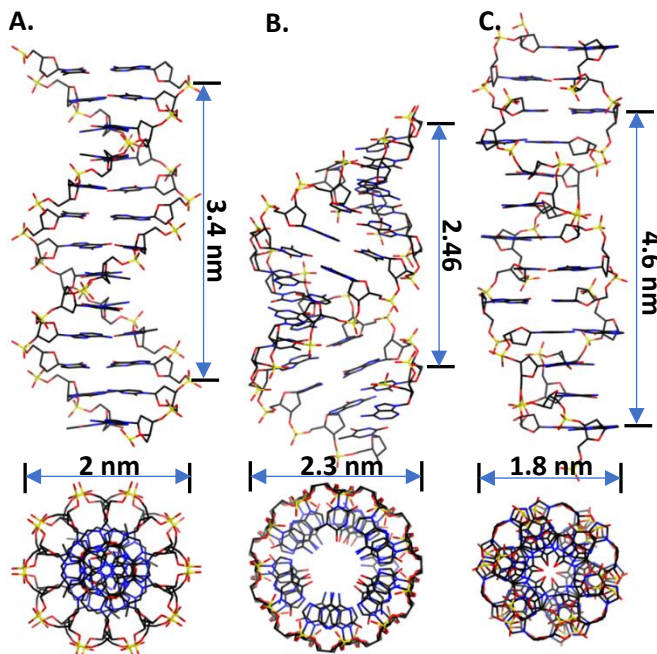
Bases can interact with each other specifically: A and T form two hydrogen bonds (H bonds), while G and C form three H bonds (Fig. 1.1, A). For both the purine and pyrimidine bases, the phosphate group forms a bond with the deoxyribose sugar.

Nucleic acids are formed when several nucleotides come together through phosphodiester linkages between the 5' and 3' carbon atoms (Fig. 1.1, B).



**Fig. 1.1.** (A) GC and AT pairing. (B) Primary structure of a 3 base-long single-stranded DNA with sequence GAC.

### § 1.1.2 DNA secondary structure and polymorphism of DNA structure



**Fig. 1.2.** Side and top view of (A) B-DNA, (B) A-DNA, (C) Z-DNA

When base pairing occurs, DNA adopts a specific secondary structure: that of a double helix. It was first determined, in what is now known the B form, by Watson and Crick in 1953 (6). DNA in cells typically consists of two complementary anti-parallel (sense and anti-sense) polynucleotide chains that are interwound. The coiling occurs in such a way that it forms two types of grooves, one major (wider) and one minor (narrower) groove (8). However, there are several forms in which DNA can occur characterized by 4 main parameters: 1. Handedness – right or left; 2. Pitch of the helix; 3. diameter ; 4. Relative size of the major and minor grooves. As a consequence, double stranded DNA can adopt three major stable conformations (9): B-DNA, A-DNA, and Z-DNA. B-DNA is the most common configuration under physiological conditions (Fig. 1.2, A). B-DNA is a right-handed helix with a pitch of 10.4 bp per helical turn (3.4 nm, each base pair separated

by 0.34 nm) and a diameter of 2 nm. The A- form is common under dehydrating conditions. It's a right-handed double helix with 2.46 nm helical pitch, 0.23 nm base pair separation, 11 bp per turn and 2.3 nm in width (Fig. 1.2, B). Z-DNA is a relatively rare left-handed double-helix with 4.6 nm helical pitch, 0.38 nm base pairs separation, 12 bp in one helical turns, and 1.8 nm (9) (Fig. 1.2, C) which can be induced when DNA is unwound and subjected to tension (10).

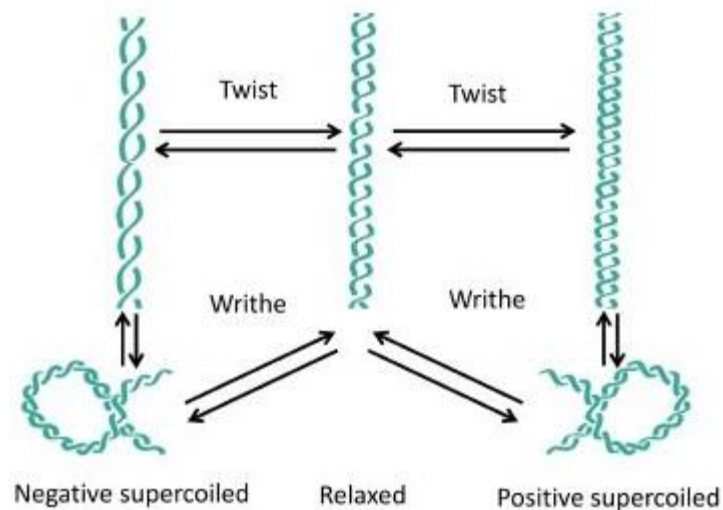
### § 1.1.3 DNA tertiary structure, DNA supercoiling

Tertiary structure is defined by the large-scale, 3-dimensional folding of DNA. Different tertiary structures of DNA include Holiday junctions, G-quadruplexes and supercoiling.

DNA supercoiling most commonly refers to the formation of gyres of double-stranded (ds) DNA whereby the longitudinal axis of the molecule crosses itself. More precisely, the torsional state of DNA is characterized by the ds-helix's linking number (Lk). B-form DNA, as found in the crystals studied by Rosalind Franklin and Maurice Wilkins, is considered to be torsionally relaxed with Lk number,  $Lk_0$ , given by the number of base pairs divided by 10.4 bp/turn,  $Lk_0 = N/10.4$ . Thus, the linking number represents the number of times that the two DNA strands are interwound, independent of their exact configuration. In vivo, the linking number of DNA dynamically changes. It is important in many biological process, such as DNA compaction and transcription regulation. Additionally, certain enzymes such as topoisomerases are able to change DNA topology to facilitate functions such as DNA replication and transcription (11). The supercoiling density of a DNA is defined as  $\sigma = (Lk - Lk_0)/Lk_0 = n/Lk_0$ , where  $Lk$  and  $Lk_0$  are



the linking numbers of twisted and relaxed DNA respectively, and  $n$  is the number of mechanically introduced twists. Two parameters that describe changes in linking number are twist ( $Tw$ ) and writhe ( $Wr$ ). Twist refers to the helical winding of the DNA strands around each other, whereas writhe measures the number of times the double helix crosses itself. In the absence of supercoiling,  $Lk_0$  equals to  $Tw_0$ . Winding or unwinding the DNA molecule will change  $Lk$  and may introduce writhe, as described in Fig. 1.3. In general, the change in linking number is given by:  $\Delta Lk = Lk - Lk_0 = Tw - Tw_0 + Wr = \Delta Tw + Wr$ .



**Fig. 1.3** Conformation of DNA supercoiling. DNA can exist as a relaxed molecule (center) which have a right-handed chirality without any supercoiling. Winding DNA can over-twist B-DNA into a tighter right-handed helix (right top) or form a right-handed plectonemic structure (right bottom). Unwinding DNA will form a helix with a longer pitch (left top) and eventually form a left-handed plectoneme (left bottom).

#### § 1.1.4 DNA quaternary structure

DNA quaternary structure refers to the helix 3-D configurations induced by the interaction with other nucleic acid molecules, or proteins. Usually, such interactions

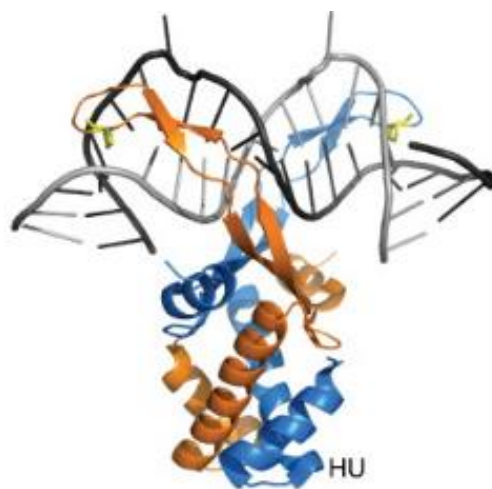
compact DNA. An obvious example is DNA compaction into nucleosomes/chromatin by histones (12). The quaternary structure of DNA strongly affects the accessibility of DNA to transcription machinery for gene expression.

## § 1.2 Introduction of HU protein

The HU protein (histone-like protein from *Escherichia coli* strain U93) is a small, histone-like, nucleoid-associated protein (NAP) which exists abundantly in bacteria (13, 14). HU is a nonspecific, DNA-bending protein, which plays an essential

architectural role in packaging DNA (15) (Fig. 1.4). In most bacteria, HU exists as an 18-kDa homodimer. In Enterobacteria,

including *Escherichia coli* (*E. coli*), there are two variants of similar HU subunits, HU $\alpha$  and HU $\beta$ , each of them with molecular weight 9-kDa. They can associate to form either homodimers (HU- $\alpha$ 2 and HU- $\beta$ 2) or heterodimers (HU-  $\alpha\beta$ ). The heterodimer is the most common form of HU in *E. coli* (16). *In vivo* HU concentrations vary between a high level of 30,000 copies per cell in exponential phase and a low level of 7500 in stationary phase (17, 18). HU restrains supercoiling and participates in specific control functions in DNA transactions like replication, transcription, recombination, and DNA repair (13, 15, 19-21). Although HU has a high non-specific affinity for DNA, it bind with high affinity to



**Fig. 1.4** Structural basis for DNA bending by HU protein (5). Cartoon representation of crystal structure of HU protein bound to a short piece of DNA (PDB:1P78).

certain types of DNA structures, such as nicked, bent, gapped, three- or four-way junction, and AT-rich DNA (22-27).

The DNA-binding behavior of HU depends on salt concentration (28). Single DNA stretching experiments have shown that at low salt concentration (below 100 mM), low HU concentration caused a reduction in extension of DNA, relative to the extension of naked DNA (29-31). However, at higher HU concentrations bending was no longer observed (32) and DNA became stiffer (29, 30). Further increasing, within physiological levels, the salt concentration (150 -200 mM), HU only compacts DNA into a flexible HU-DNA filament (28). Finally, DNA compaction is eliminated at 300 mM, or higher, salt concentration (28).

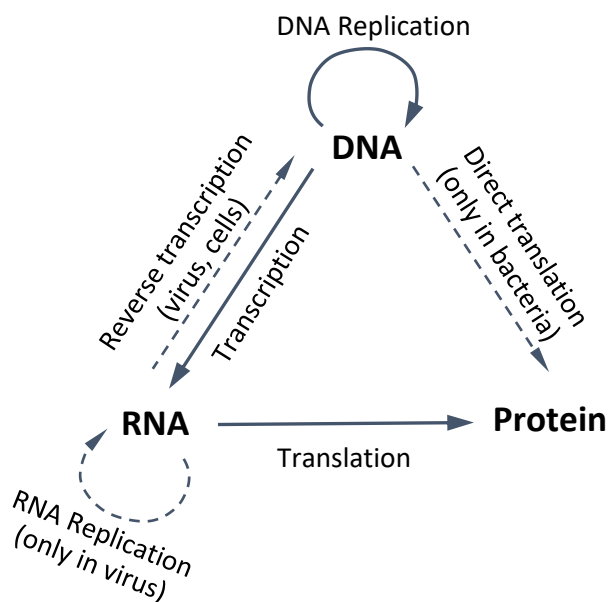
HU potently changes the bacterial transcriptome (14), and binding correlates with negative supercoiling in stationary phase *E. coli* (33).

## **§ 1.3 Introduction of transcription**

### **§ 1.3.1 Central dogma of molecular biology**

The central dogma of molecular biology is an explanation of the flow of genetic information within a biological system. It is a framework for understanding the transfer of the sequence information between information-carrying biopolymers in living organisms. It was first stated by Francis Crick in 1958 (34), and re-stated in a paper published in 1970 (35). There are three major classes of such biopolymers: DNA, RNA (both nucleic acids), and proteins. There are  $3 \times 3 = 9$  conceivable direct transfers of

information that occurs between these (Fig. 1.5). The dogma classifies these into 3 groups of 3: three general transfers (believed to occur normally in most cells) (Fig. 1.5 solid arrows), three special transfers (known to occur, but only under specific conditions in case of some viruses or in a laboratory) (Fig. 1.5 dotted arrows), and three transfers believed never to occur.



**Fig. 1.5** Illustration of the central dogma of molecular biology. Solid arrows show general transfer; dotted arrows show special transfer. The absent arrows are the undetected transfers specified by the central dogma.

The general transfers describe the normal flow of biological information: DNA can be copied into new DNA through DNA replication (DNA  $\rightarrow$  DNA). DNA codes for RNA through transcription (DNA  $\rightarrow$  RNA). There are three types of RNA: ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA). The information carried by mRNA is used by ribosomes, large molecular machines, to synthesize proteins through the process of translation (RNA  $\rightarrow$  Protein). rRNA is a component of the ribosome, and

is essential for protein synthesis, tRNA carries amino acid residues to be incorporated into mRNA to the ribosome.

The special transfers describe: RNA being copied into RNA through RNA replication (RNA  $\rightarrow$  RNA). This process occurs because of the existence of RNA virus. DNA being synthesized using an RNA template by reverse transfer (DNA  $\rightarrow$  RNA), and protein being translated directly from a DNA template without the mRNA (DNA  $\rightarrow$  Protein). The unknown transfers describe: a protein being copied from a protein (Protein  $\rightarrow$  Protein), RNA synthesized using a protein as a template (Protein  $\rightarrow$  RNA), and DNA synthesized using the protein as a template (Protein  $\rightarrow$  DNA) (35).

### § 1.3.2 Components and function of RNA polymerase

RNA polymerase (RNAP) is the enzyme that catalyzes the copying of the information of a DNA into an RNA sequence during the process of transcription. It has been found in all species, but the number and composition of proteins vary cross taxa. In most prokaryotes, a single RNA polymerase species transcribes all types of RNA. *E. Coli* RNA polymerase consists of five different subunit types. The beta ( $\beta$ ) subunit with molecular weight of 150-kDa, the 155-kDa beta prime ( $\beta'$ ), the 10-kDa omega subunit ( $\omega$ ), two 36-kDa alpha ( $\alpha$ ) subunits, and the sigma factor ( $\sigma$ ). There are different types of  $\sigma$  factors, one of these is the  $\sigma^{70}$  factor with a molecular weight of 70-kDa. The  $\sigma$  subunit is recruited for promoter recognition and transcription initiation, it can dissociate from the rest of the complex, leaving the core enzyme ( $\alpha_2\beta\beta'\omega$ , 11 Å (36)). Only as a

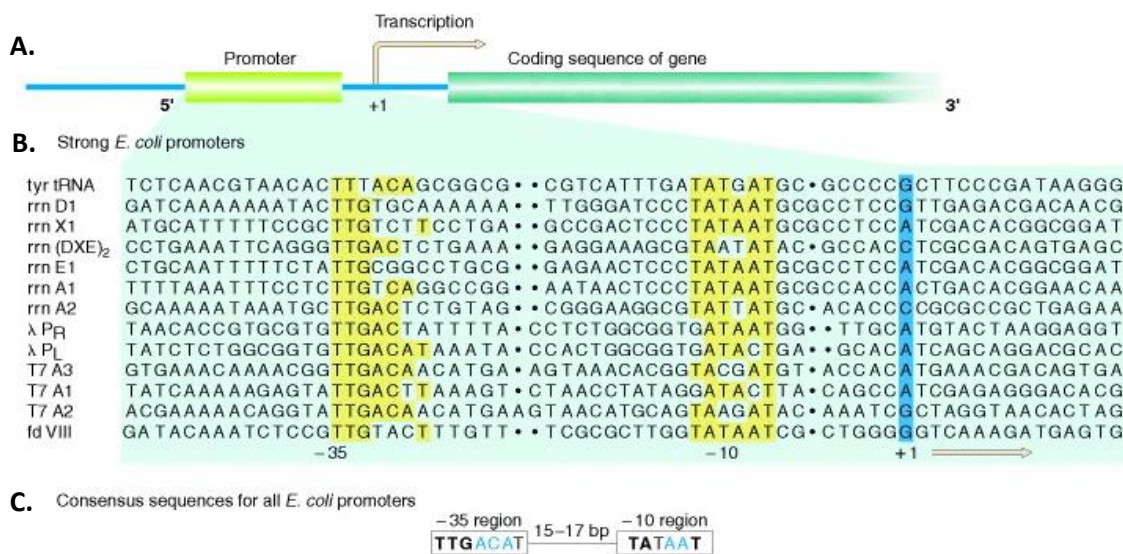
holoenzyme ( $\alpha_2\beta\beta'\omega\sigma^{70}$ , 27 Å (37)) RNAP can specifically initiate transcription (38). The  $\alpha$  subunits form a dimer ( $\alpha_2$ ) that acts as the platform onto which the  $\beta$  and  $\beta'$  subunits bind, and which plays a role in transcription activation (39). The  $\beta$  and  $\beta'$  subunits together form the catalytic center of the enzyme (36). The physiological role of the small  $\omega$  subunit is not yet completely understood in bacteria, a recent research indicate that in cyanobacteria the  $\omega$  subunit facilitates the association of the primary  $\sigma$  factor with the RNAP core, thus controls transcription efficiency (40).

### § 1.3.3 Transcription steps: initiation, elongation and termination

Transcription is the first step of gene expression. During transcription, the information of DNA is copied into RNA by RNA polymerase. Transcription in *E. coli* comprises three steps: initiation, elongation and termination.

**Initiation**, which is typically the rate-limiting and the most regulated stage of transcription, is by itself a complex, multistep process consisting of the following steps (41, 42). (i) RNAP holoenzyme binds to the -10 and -35 DNA elements in the promoter (Fig. 1.6) recognition sequence (PRS), upstream to the transcription start site (TSS) to form a closed promoter complex ( $RP_c$ ). (ii)  $RP_c$  isomerizes through multiple intermediates into an open complex ( $RP_o$ ), in which a ~12-bp DNA stretch (bases at positions -10 to +2) is melted to form a transcription bubble, the template DNA strand ( $3' \rightarrow 5'$  strand) is inserted into the RNAP major cleft, the base at +1 of the TSS is placed in register with the active site, the non-template strand ( $5' \rightarrow 3'$  strand) is tightly bound

to  $\sigma^{70}$  and the downstream DNA duplex (bases +3 up to +20) is loaded into RNA  $\beta'$  DNA-



**Fig. 1.6** Promoter sequence. (A) The promoter lies “upstream” of the initiation point and coding sequences. (B) Promoter sites have regions of similar sequences, as indicated by the yellow region in the 13 different promoter sequences in *E. coli*. Spaces (dots) are included to maximize homology at the consensus sequence. The gene governed by each promoter sequence is indicated on the left. Numbering is according to the number of bases before (-) or after (+) the RNA synthesis initiation point. (C) Color coding in the consensus sequences for all *E. coli* promoters is as follows: blue letters, >75%; boldface black letters, 50-75%; black letters, 40-50% (3).

binding clamp. (iii) An initial transcribing complex (RP<sub>ITC</sub>) is formed after binding of nucleoside triphosphates (NTPs), and the start of RNA synthesis during an abortive initiation step (AI), followed by RNAP cycling through multiple polymerization trials via a DNA scrunching mechanism (43, 44), release of short “abortive transcripts” and return to an RP<sub>O</sub> for a new synthesis trial (45-47); and finally, (iv) RNAP promoter escape, which occurs when enough strain is built in the enzyme, the  $\sigma^{70}$  undergoes structural transition to relieve blockage of the RNA exit channel in RNAP and loses its grip on the PRS, nascent RNA enters the RNA exit channel and transcription enters the elongation stage. In AI, the interactions between  $\sigma^{70}$  and the PRS limit the lengths of abortive transcripts.

The stronger these interactions are, the longer the time RNAP will spend cycling in AI (48) and the longer the lengths of abortive transcripts will be (49). For these reasons, transcription initiation is rather slow after establishing tight promoter interactions.

**Elongation** is a very efficient and fast process (50-53). Shortly after initiation, the  $\sigma$  factor most often dissociates from the RNAP. The RNA is always synthesized in the 5'  $\rightarrow$  3' direction (Fig. 1.7), with triphosphates (NTPs) acting as substrates for the enzyme.

The addition of each ribonucleotide can be represented by the following chemical

equation:  $+(NMP)_n \xrightarrow{Mg^{2+}, RNAP} (NTP)_{n+1} + PP_i$ . The energy for elongation is provided

by the splitting of high-energy triphosphate into monophosphate and release of

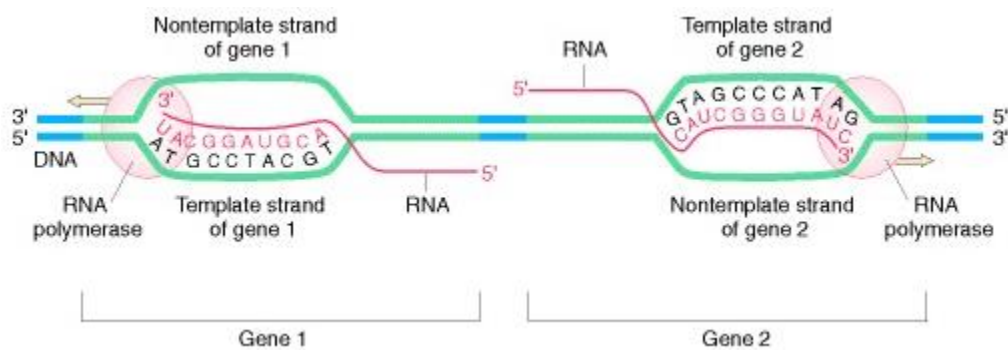
inorganic diphosphate ( $PP_i$ ). The so called “transcription bubble” must be maintained

during elongation and it moves along the DNA double strands. As RNAP moves along the

helical groove of DNA, it generate positive (+) DNA supercoiling ahead and negative (-)

DNA supercoiling behind. This is the “twin supercoiled domain model” (54). During

elongation, RNAP may encounter specific sequences that cause transcriptional pausing



**Fig. 1.7** Transcription of two divergent genes. RNAP moves from the 3' end of the template strand, creating an RNA strand that grows in a 5'  $\rightarrow$  3' direction. Some genes are transcribed from one strand of the DNA double helix, other genes use the other strand as the template. (Figure from An Introduction to Genetic Analysis, 7<sup>th</sup> edition).



or stalling, sometimes accompanied by backtracking of the RNA chain pushing its 3' end into the RNAP secondary channel (55), which becomes critical for termination of transcription.

**Termination** is the process that ends transcription, which includes the release of the nascent RNA and the enzyme from the template. There are two major mechanisms for termination in *E. coli*: i) intrinsic termination (or Rho-independent transcription termination), and ii) Rho-dependent termination.

The first type involves termination sequences within the RNA that signal the RNA polymerase to stop. The termination sequence on the DNA template usually contains about 40 bp, ending in a GC-rich stretch that is followed by a run of six or more A's on the template strand. The corresponding GC sequences on the RNA are arranged so that the transcript in this region can form complementary bonds with itself (a palindromic sequence) that forms a stem-loop hairpin structure. It is formed by the terminal run of U's that correspond to the A residues on the DNA template. The RNA hairpin loop and section of U residues appear to serve as a signal that leads to the dissociation of the RNAP from the DNA template and terminate the transcription.

The second mechanism uses a termination factor called  $\rho$  factor (rho factor) which is a protein that stop RNA synthesis at specific sites. This protein binds at a rho utilization site on the nascent RNA strand and runs along the mRNA towards the RNAP. mRNA with rho-dependent termination signals do not have the string of U residues at the end of the RNA and usually do not have hairpin loops. When the  $\rho$ -factor, bound on

RNA, reaches the RNAP, it causes RNAP to dissociate from the DNA, terminating transcription (56).

### **§ 1.3.4 Transcriptional regulation**

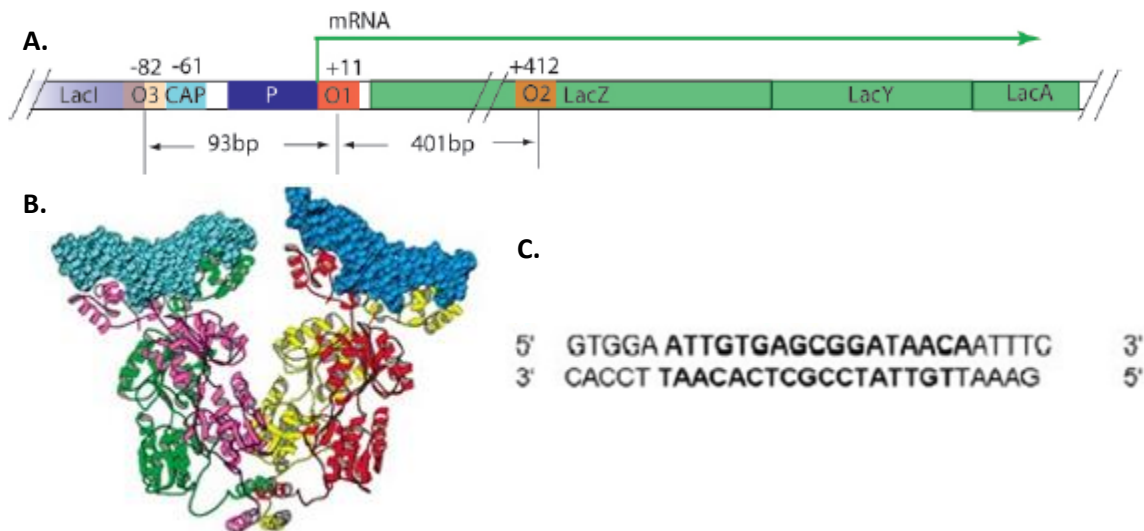
Transcriptional regulation in organisms is accomplished through the sequence specific binding of proteins (transcription factors, or TFs) that activate or inhibit transcription. Transcription factors may act as activators, repressors, or both. Repressors often act by preventing RNAP from forming a productive complex with the promoter, or by binding downstream of the promoter and acting as roadblocks to elongation. Activators, instead, facilitate formation of a productive complex. Furthermore, DNA motifs induced by TFs have been shown to be predictive of epigenomic modifications, suggesting that transcription factors play a role in regulating the epigenome (57).

## **§ 1.4 The lac repressor protein, LacI, and LacI-mediated DNA looping**

### **§ 1.4.1 Lac Operon and structure of the LacI protein**

The *lac* system in *E. coli* bacteria has been a classic model for both *in vivo* and *in vitro* investigation of gene regulation. The *lac* operon refers to the genes responsible for the lactose transport and metabolism in bacteria (58). It consists of three adjacent structural genes, LacZ (coding for  $\beta$ -galactosidase), LacY (coding for lactose permease), and LacA (coding for transacetylase), under the control of the same promoter (Fig. 1.8, A). The *lac* operon is regulated by several factors, including the presence of glucose and lactose.

The response to lactose requires the regulatory protein Lac repressor (LacI) which is encoded by the nearby LacI gene. In the absence of lactose, LacI binds to one or two specific binding sites, operators, and prevents RNAP from transcribing, either by inhibiting the binding of RNAP (59-61) or stopping its entry into the processive elongation phase (62, 63). Thus, the operon is “off”. In the presence of lactose without glucose, LacI undergoes a conformational change which dramatically decreases the affinity of LacI for its operators. In this case, RNAP that recognize the promoter can transcribe the operon’s structural genes into mRNA and the operon is “on”.



**Fig. 1.8** Gene regulation in *lac* Operon. (A) Schematic of *lac* Operon. (B) A ribbon diagram of the quaternary structure of the Lac repressor complexed to DNA, Lac repressor is a tetrameric structure where each monomer shows in one color. (C) Sequence of DNA operator, the bases in bold font indicate the minimal requirement for specific binding of an operator.

LacI is a “V-shape” homo-tetramer (Fig. 1.8, B) with two DNA binding domains.

The binding of a LacI to two distinct operators along a DNA molecule induces the looping out of the intervening DNA sequence. Operators are normally DNA sequences of approximately 27 base pairs; the minimal operator required for specific binding is 17

base pairs-long (Fig. 1.8, C) (64). There are three wildtype LacI operators (Fig. 1.8, A).

The primary operator O1 is located at position 11 relative to the promoter. Two auxiliary operators, O2 and O3, lie within the LacI and LacZ gene, respectively. O3 is 93 bp upstream of O1 while O2 is 401 bp downstream of O1. In addition, there is an engineered, symmetric operator Os. The affinity of LacI to these four operators is:  $O_s > O_1 > O_2 > O_3$ . Maximally efficient repression is achieved via DNA looping by the simultaneous binding of LacI protein to two operators (58, 65). This is because DNA looping significantly enhances protein association to the lower-affinity site due to the tethering effect of DNA looping, and increases LacI concentration in the proximity of the operators (66, 67).

#### **§ 1.4.2 LacI mediated DNA looping study *in vivo* and *in vitro***

Protein mediated DNA looping is important in all aspects of DNA metabolism, such as transcription, replication and recombination. To understand the physical mechanism of protein-induced DNA loop formation, and its physiological role, it is necessary to conduct experiments both *in vitro* and *in vivo*, as well as theoretical analyses.

The probability of loop formation is set by the LacI tetramer concentration, binding affinities of operators, and the loop size (4, 68). *In vitro* experiment shown that loop formation by the LacI protein is concentration dependent. For any given DNA template with two LacI operators, when the concentration is low, the probability of forming a loop is small and will increase while the LacI concentration increase. However,

if the concentration of LacI is too high, the probability of forming a loop is low because the two operators are occupied by separate LacI tetramers. At intermediate concentrations, the looping state has its highest probability of occurrence (69).

Loop size is another factor that influence the looping probability. For small loop sizes, looping needs to overcome bending stiffness (70, 71). In this regime, looping and the gene expression level are oscillating with the length of the loop with a periodicity of approximately 11 bp, which is consistent with the DNA helical repeat (58, 72-75). The characteristic oscillation in gene repression amplitude in the *lac* operon is characterized at the single base-pair level. Large loop formation is suppressed by entropy. The free energy of looping measures the energy involved in forming the looped configuration as a function of the inter-operator distance. It is an important parameter that can be extracted using thermodynamic models (76, 77). The minimum in looping free energy is around 70 bp *in vivo*, which is about half the persistence length of DNA (76, 78-80). *In vitro*, theoretical work suggests that approximately 500 bp long DNA segments most easily form LacI mediated loops (71). This difference has been suggested to arise from the greater effective flexibility of *in vivo* versus *in vitro* DNA. The mechanics of DNA inside living cells is considerably more complicated than *in vitro*. Besides the flexibility of naked DNA, a number of other factors should be considered including (i) the geometry and flexibility of the looping protein (61, 81), (ii) the existence of DNA supercoiling in the cell (82) and (iii) the presence of structural proteins such as HU, IHF and H-NS in prokaryotes (74), which are abundant nucleoid-associated proteins. The contribution and importance of each factor need to be decoupled to analyze the role of DNA looping

in transcriptional regulation. To examine each factor separately, *in vitro* single molecule biophysics have been employed.

Bulk binding assays, such as electrophoretic mobility shift assays and filter binding assays, are traditionally applied to measure the affinity of proteins to their DNA targets. These techniques offer the opportunity for systematically varying parameters, such as DNA length, the degree of supercoiling, or presence of structural proteins, thereby decoupling the various effects contributing to looping processes *in vivo*. Previous research using electrophoretic mobility shift assay (EMSA) found that the looping probability decreases as the inter-operator spacing is reduced from 210 bp to 60 bp (83), which agrees with the quantitative observations using the filter binding assay (84), and demonstrated an increase in the affinity of LacI binding to a single operator site. Their results suggest that negative supercoiling favors the association of LacI and its target DNA sites, resulting in an increase in looping probability and looping stability. Furthermore, this probability is found to remain relatively constant over loop lengths between 100 and 500 bp (85). It was also found that supercoiling could shift the optimal spacing for loop stabilization depending on the degree of supercoiling, suggesting the distance of helical repeat is changing (86). Although such studies demonstrate the role of supercoiling and inner-operator spacing in DNA looping, they lack either quantitative kinetic information or systematic investigation.

Another important class of experiments that have shed light on the mechanics of DNA looping *in vitro* are single-molecule measurements using the Tethered Particle Motion (TPM) method (87, 88). Finzi and Gelles were the first who applied TPM to

measure LacI mediated loop formation and breakdown, and to elucidate the kinetic information of such process (89). In this method, looped and unlooped states can be distinguished by the excursion of the tether (90-98). Thus, modulations in its motion reflect conformational changes in the tethering molecule. This method has recently revealed the presence of two-looped states in LacI mediated DNA loops (69, 99-103), consistently with the presence of multiple configurations of the LacI-DNA nucleoprotein complex observed using FRET (104, 105), electron microscopy studies (106) and suggested by x-ray crystallography studies (61). The different excursion may be explained as characteristic of either the parallel vs. anti-parallel arrangement of the DNA in the nucleoprotein complex (69), or open vs closed LacI tetramer conformation (100). AFM imaging has directly observed the open and closed states of LacI protein (2). Another study using TPM provided evidence for at least three distinct loop structures contributing to LacI-mediated looping *in vitro* when the loop length is short (on the order of the DNA persistence length), more than previously reported (107).

When both DNA and LacI are present in solution, five states are possible: bare DNA, DNA with one LacI bound to a single operator, DNA with two distinct LacI tetramers, each bound to a separate operator, DNA looped by LacI (101). The weight of each configuration as shown in (101). This model states that if the operators have dissociation constant  $K_{weak}$  and  $K_{strong}$  for weak and strong operator respectively. The interning DNA has looping  $J$ -factor  $J_{loop}$ , this term applied to quantify the effective concentration of LacI near weak operator when the strong operator is occupied by a LacI protein.

Thus, the expression of looping probability  $P_{loop}$  will have to consider the weight of the five possible states, and will be given by the following equation,

$$P_{loop} = \frac{\frac{1}{2} \frac{[R]J_{loop}}{K_{weak}K_{strong}}}{1 + \frac{[R]}{K_{weak}} + \frac{[R]}{K_{strong}} + \frac{[R]}{K_{weak}} \frac{[R]}{K_{strong}} + \frac{1}{2} \frac{[R]J_{loop}}{K_{weak}K_{strong}}}$$

Where  $[R]$  is the concentration of LacI, and  $J_{loop}$  is the  $J$ -factor, which refers to the probability of forming a loop from one operator bound by LacI. The  $J$ -factor depends on the length, flexibility of the DNA and the phasing of the operators (108-110).

The  $J$ -factor can be calculated from the ratio of the unlooped to the looped state probabilities (101),  $\frac{P_{loop}}{P_{unloop}}$ , the dissociation rate from the weak operator  $K_{weak}$ , and the concentration of LacI,  $[R]$ , using the equation:  $J = 2 \frac{P_{loop}}{P_{unloop}} ([R] + K_{weak})$ .

## § 1.5 DNA topological domains

Topological domains, which are operational units in practically all DNA genomes, are DNA segments with constrained ends, this means that the free rotation of the ends is inhibited and the supercoiling density within such a topological domain is invariable.



**Fig. 1.9** Examples of topological domains. (A) Circular DNA, (B) linear DNA attached to other cellular structure, (C) chromosomal DNA loops.



The ends of DNA are constrained in circular DNA (Fig. 1.9, A), as typically found in bacteria, mitochondria, chloroplasts, viruses, etc. In this case, there are obviously no DNA ends, since both DNA strands are covalently closed. In eukaryotic chromosomes, instead, the ends of the chromosomal DNA molecule (or of parts of it) may attach to other cellular structures such as the nuclear membrane (Fig. 1.9, B), or by protein-mediated DNA loops (Fig. 1.9, C)(111).

There is both *in vitro* and *in vivo* evidence for the existence of topological domains. Chromosomes isolated from *E. coli* cells required a number of nicks instead of a single nick to relax the supercoiling completely (112). The whole chromosome is divided into 12-80 domains. This is in agreement with other types of *in vivo* measurements, where the average size of the domains was estimated to be ~ 100 kb in length, so that each genome would be divided into ~50 domains (113). Electron microscopy (114, 115) provides morphological evidence for the existence of topological domains. Isolated *E. coli* chromosomes display many individually supercoiled loops emanating from a central region. These supercoiled loops were hypothesized to be topological domains. This was supported by the finding that a single loop could be relaxed while the rest of the loops remained supercoiled that in average some domains were relaxed and some would not (116). Moreover, the number of loops visible by electron microscopy are estimated between 65 and 200 per nucleoid (114, 115). Later experiments have suggested that this domain size is too large. A less invasive approach to studying domains *in vivo* focused on the study of the domain structure in a region of the *Salmonella enterica* chromosome exploiting the requirement for negative

supercoiling of the  $\gamma\delta$  resolvase DNA substrate. The study concluded that the topological domains are variable in both size and location of topological constraints, and that domains average 25 kb in length (117). Gene microarrays were used to examine the entire chromosome at once to determine how far DNA relaxation extends from double strand breaks generated *in vivo* by a restriction enzyme. Electron microscopy was independently used to directly measure the size the supercoiled domains. Both methods gave an average size of domain  $\sim 10$  kb (118). Monte Carlo simulations using varying mean domain lengths concluded that domain barriers are not located at fixed sites on the chromosome but instead are randomly distributed (118).

There are significant advantages to having topological domains in a chromosome. First, topological domains will contribute substantially to the compaction of the chromosome. The division of a replicated chromosome into  $\sim 500$  loops would decrease its radius of gyration from  $\sim 10 \mu\text{m}$  to  $<0.5 \mu\text{m}$  (119). Second, supercoiling is a vulnerable state; without topological domains, a single nick or break would cause the supercoiling relaxation of the whole genome and lead to cell death. Thus, the existence of domains reduces the amount of DNA that is relaxed by DNA nicking or breakage. This is important for bacteria, which require negative supercoiling for viability (120, 121). Third, small domains will greatly simplify the problem of catenane and precatenane resolution following replication (122, 123). Small domains concentrate catenane links, making it easier for the decatenase, topoisomerase IV, to find the links between enormous chromosomes (124). Concentration of catenane links will also increase the free energy of catenanes and, therefore, will drive decatenation toward completion

(125). Forth, organization of the chromosome into small domains will greatly aid in the repair of double-strand breaks by keeping the two ends to be jointed in proximity to each other. Proper localization of the ends would be particularly important when multiple double-strand breaks are repaired by nonhomologous end-joining. Finally, domains partition the genome into active and non-active regions.

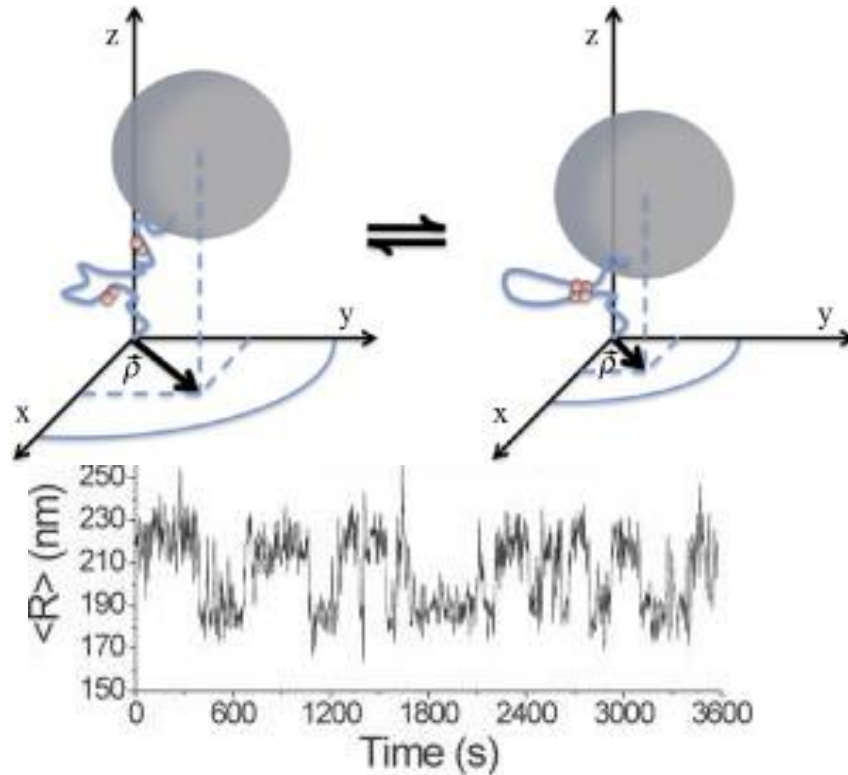
## **§ 1.6 Experimental techniques**

### **§ 1.6.1 Single molecule techniques**

Single molecule techniques are techniques that investigate the properties and behavior of individual molecules, in contrast to more traditional bulk techniques that measure the ensemble average of a given parameter. Single molecule techniques include single molecule fluorescence imaging and single molecule manipulation methods. In the research described here, the tethered-particle motion (TPM) technique and magnetic tweezers (MTs) were used to study protein-induced DNA looping.

### **§ 1.6.2 Tethered-particle motion**

The tethered particle motion (TPM) technique is an optical microscopy method that is used for studying various biopolymers, such as DNA, and their interaction with other entities, such as proteins. TPM was first introduced by Schafer, Gelles, Sheetz and Landick in 1991 (87). In this method, a biopolymer is tethered between a stationary substrate, such as the microscope flow-chamber surface, and a micrometer-scale sphere



**Fig. 1.10** Tethered particle motion and the change of excursion length when bead shows looping transition.

(bead), which is large enough to be imaged with conventional optical microscopy (Fig. 1.10). The constrained Brownian motion of the bead serves as a reporter of the underlying macromolecular dynamics of the tethering DNA. Thus, the bead position is tracked in time. The projection of the bead position in the x-y plane at any given point in time is recorded and, over a sufficiently long period of time, the distribution of positions in the 2-D plane will be circular. The center of the circle defines the tether's anchoring point with coordinates  $(\bar{x}_t, \bar{y}_t)$  and the distance between the anchoring point and the position of the bead at any time,  $t$ , defines the excursion length, given by  $\rho_t = \sqrt{(x - \bar{x}_t)^2 + (y - \bar{y}_t)^2}$ , where  $(x, y)$  are the coordinates of the bead at any given point. Changes in the extent of the average motion (the "average excursion",  $\langle \rho \rangle$ )

reflect conformational changes of the tethering molecule. Such changes may be caused, for example, by DNA looping, DNA hybridization or DNA bending (89, 90, 96-98, 126-128). The time necessary to obtain a mean square displacement depends on the viscosity of the buffer (129).

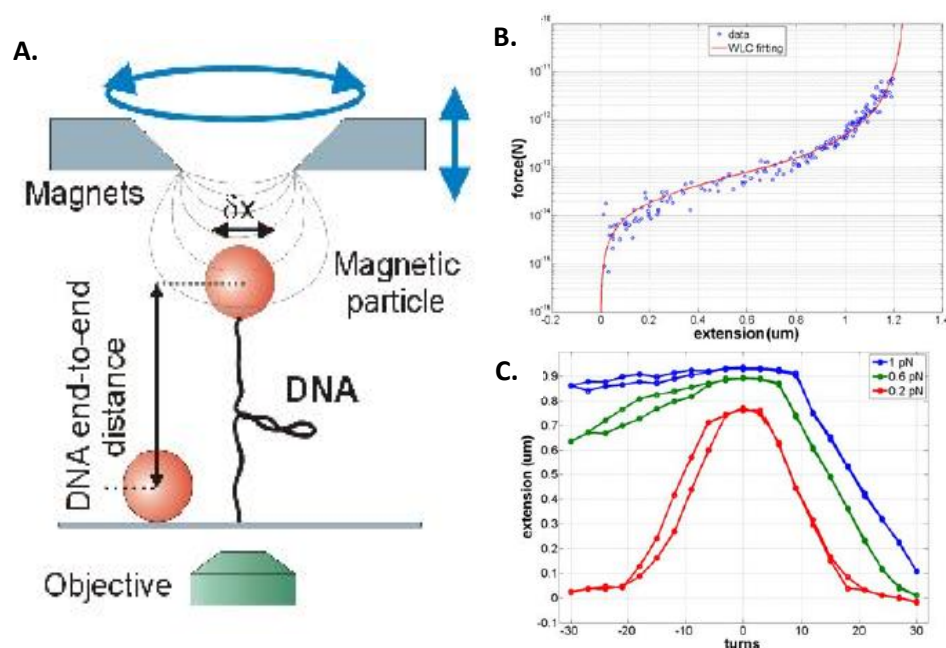
The motion of the bead is influenced by the particle size, the polymer length, the physical and chemical properties of the polymer, the solution in which the bead and the polymer diffuse, and the nearby surface. The volume-exclusion effect due to the proximity of the bead to the wall can be characterized by an excursion number ( $N$ ) to indicate whether the motion of the bead is dominated by diffusion of the DNA ( $N < 1$ ) or diffusion of the bead ( $N > 1$ ) (94).

The interpretation of TPM measurements relies on an appropriate calibration curve which relates the excursion (in terms of  $\langle \rho \rangle$ ) and a measure of the end-to-end length of the DNA tether, to the contour length of the tether. Calibration curves vary depending on the buffer, viscosity and bead size.  $\langle \rho \rangle$  depends non-linearly on the DNA contour length, and the slope of the curve decreases as the DNA contour length increases (95, 129). In contrast, the mean-square excursion,  $\rho^2 = (x - \bar{x})^2 + (y - \bar{y})^2$ , depends linearly on the contour length of DNA from 0.1 kbp to 3.5 kbp (129), and the calibration curve using a 160 nm radius bead in  $\lambda$  buffer (buffer composition see section 2.2) is  $\langle \rho^2 \rangle = 100.89 \times L + 3445$  (129).

Although TPM is a relatively simple technique, there are subtleties to consider. For example, multiply-tethered beads should be discarded, and DNA non-specific

absorption, bead transient sticking to the surface and dissociation of the tether should all be minimized (92, 93, 95, 98, 130). Furthermore, the total observation, exposure and intrinsic diffusive times of the tethered particle need to be considered.

### § 1.6.3 Magnetic tweezers



**Fig. 1.11** Magnetic tweezers setup (A) and DNA behavior under tension (B) and torsion (C). (A) Schematic experimental setup; (B) Force vs. DNA extension (open circle) and WLC fitting of the data (red line); (C) DNA extension vs. magnet turns curve (hat curves) under different tension.

The application of magnetic tweezers (MT) to pull on single DNA molecule was first introduced by Smith, Finzi and Bustamante in 1992 (131). In 1996, Strick, Allemand, Bensimon and Croquette published a different implementation of MT with which to simultaneously pull and rotate DNA, which is widely used today (132). The basic setup of MT (Fig. 1.11, A) is similar to that used in TPM, a DNA molecule tethers a magnetic bead

to the glass surface of a microscope flow-chamber. Above the stage, there is a pair of permanent magnets which attract the magnetic bead, thus applying tension to the DNA, the magnets can also be rotated to twist the DNA molecule.

DNA in vivo is under tension and torsional stress. Therefore, understanding DNA behavior in the presence of these parameters is very important to the understanding of DNA transactions. Tension on the DNA can be calculated using the equation  $F = k_B T \frac{l}{\delta x^2}$ , where  $k_B$  is the Boltzman constant,  $T$  is the temperature,  $l$  is DNA extension,  $\delta x^2$  is the mean square displacement of the bead in the x direction which can be measured experimentally by tracking the center of the bead (132). By monitoring the position or fluctuation of the tethered bead, the behavior of DNA under tension or torsion can be revealed. The dependence of the DNA end-to-end distance (DNA extension) on external forces below the phase transition point (60 pN) is shown in Fig. 1.11 (B). The force vs. extension curve of DNA molecule follow Worm-like-chain model

$$(133), F = \frac{k_B T}{P} \left[ \frac{1}{4 \left(1 - \frac{l}{L}\right)^2} - \frac{1}{4} + \frac{l}{L} \right],$$

$l$  is the extension of DNA (end to end distance),  $F$  is the tension applied on single DN molecule. One of the fitting parameters,  $P$ , is the persistence length, which represents the stiffness of the DNA molecule, the longer the persistence length, the stiffer the DNA molecule. The persistence length of B-form DNA is approximately 50 nm (134). Another fitting parameter,  $L$ , is the contour length of DNA.

DNA can also be twisted with magnetic tweezers. The DNA extension vs. turns curve show different behavior under different tension (Fig. 1.11, C). Since B-form DNA is

a right-handed molecule, when the tension is high enough, unwound DNA does not buckle, but is thought to denature (135) or switch to left-handed DNA (10). In this situation, the extension remains unchanged, because no DNA fragment is absorbed to the plectonemic form (132). Overwound DNA adopts a plectonemic form that smoothly reduces the DNA extension. Under low tension, either wound or unwound DNA molecule will introduce plectonemic structure, result a decrease DNA extension, lead to a symmetric hat curve. Thus, by investigating DNA extension under tension or torsion, not only one can characterize its mechanical properties, such as its persistence length (bending rigidity) (133) and its torsional persistence length (torsional rigidity) (136), but one can follow conformational changes induced by different external agents, such as proteins.



## Chapter 2 Material and Methods

### § 2.1 Preparation of DNA constructs

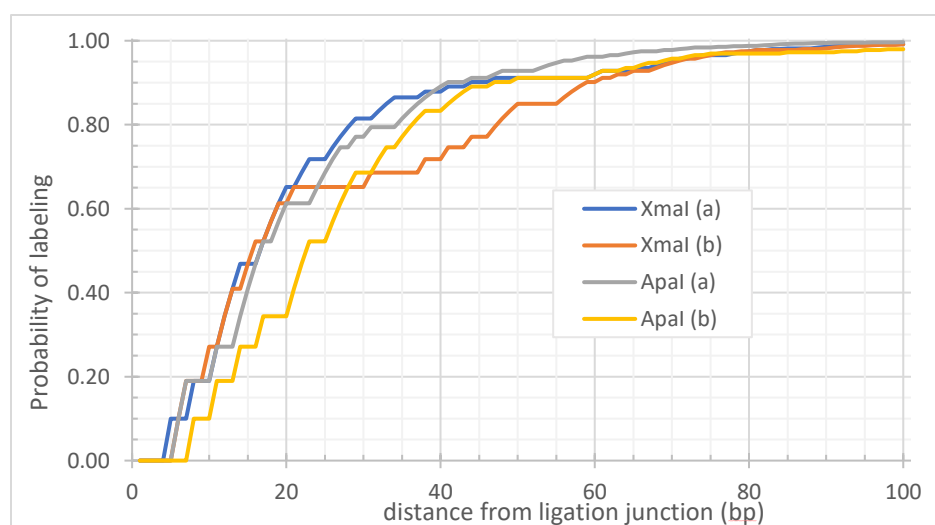
All DNA fragments for TPM experiments were amplicons of PCR reactions with plasmid DNA (pUC-HF-LL18 (4), pYY\_I1\_400\_BstEII (137), or pO1O2 (138)) as templates. The plasmid pYY\_I1\_400\_BstEII was made by using a Gibson assembly kit (New England BioLabs, Ipswich, MA). By this method, five pieces of DNA were inserted into the pBR322. The first fragment contained the promoter, the second contained the near operator (O<sub>s</sub>), the third the spacer, the fourth the far operator (O<sub>1</sub>), and the fifth the tail with the terminator sequence. Each consecutive fragment had approximately 25 bp overlap. Other reagents were obtained from the following sources: dNTPs (Fermentas-Thermo Fisher Scientific Inc., Pittsburgh, PA), digoxigenin- and biotin- labeled primers (Integrated DNA Technologies, Coralville, IA, or Invitrogen, Life Technologies, Grand Island, NY) (4, 139) and Taq Polymerase (New England BioLabs, Ipswich, MA). The total lengths of the DNA constructs named O<sub>s</sub>-900-O<sub>1</sub>, O<sub>s</sub>-400-O<sub>1</sub> and O<sub>1</sub>-400-O<sub>2</sub> (written as O<sub>strong</sub>-LoopSize (bp)-O<sub>weak</sub>) were 1632, 909, and 831 bp respectively, with centrally located LacI-inducible loops.

DNA tethers for DNA looping experiments in the magnetic-tweezers were generated by ligating a 2115 bp (0.7 μm) main construct containing the O<sub>1</sub>-400-O<sub>2</sub> sequence with a centrally positioned, 400 bp, LacI-inducible loop (139). This construct was produced by PCR using plasmid pO1O2\_401 (140) as a template for the

*S/pO102\_401/929\_Xmal* (tgCCCGGGaccggaagacatgc) and *A/pO102\_401/3043\_Apal* (ctGggCCCggtgaatccgtagcga) primers. The main fragment so obtained was then digested and ligated with T4 DNA ligase (New England BioLabs, Ipswich, MA) to a digoxigenin- or a biotin- labeled DNA tail at each end. The tails were approximately 1 kbp in length. After labeling the final DNA molecules with a streptavidin-labeled bead and attaching them to an anti-digoxigenin-coated coverglass they were torsionally constrained (141). The main construct was a PCR amplicon produced with an equimolar dNTP mix and primers including *Apal* or *Xmal* (New England BioLabs, Ipswich, MA) restriction digest sites. The nearly 1 kbp biotin- or digoxigenin- labeled DNA fragments for anchorage were generated by *Apal* or *Xmal* restriction digests of approximately 2 kbp PCR amplicons obtained from *pBluKSP+* and *S/pUC19/2412* and *A/pUC19/1435* primers, using dATP, dCTP, dGTP, dTTP (Fermentas-Thermo Fisher Scientific Inc., Pittsburgh, PA) and digoxigenin-11-dUTP (Roche Life Science, Indianapolis, IN) or biotin-11-dUTP (Invitrogen, Life Technologies, Grand Island, NY) in a molar ratio of 1:1:1:0.9:0.1. After digestion and purification with silica-membrane-based kits (Qiagen, Germantown, MD), the main and the anchorage fragments were ligated using T4 DNA ligase. Incorporation of the biotin- and digoxigenin-dUTP labels was random and approximately 10% of dTTP sites in the sequences. Since the labels are attached to dUTP and the sequences are known, the probability of incorporating labeled nucleotides can be calculated as a function of the distance from the ligation junction ( $x$ ) as the complement of the probability of not incorporating any label within that span. Since the fraction of labeled nucleotides was 0.1, the probability of labeling at each adenine

nucleotide was 0.1. The probability that not even one label would have been incorporated within a segment  $x$  nucleotides long was  $P_{unlabeled} = 0.9^{N_{A,x}}$ , where  $N_{A,x}$  is the number of adenines in the segment. Then the probability that at least one label will be incorporated in the segment is  $P_{label} = 1 - 0.9^{N_{A,x}}$ . There is about a 0.95 probability that at least one label will be incorporated within 6 to 71 bp from the junction (Fig. 2.1, Appendix 1). The variation of the effective tether length due to labeling will range from 14 to 127 bp or about 38 nm. Multiple labels along the anchorage segments are necessary to torsionally restrain the DNA tethers.

The DNA molecules generated for transcription elongation measurements in



**Fig. 2.1** The probability of labeling along the anchorage fragment as a function of the distance from the ligation junction with the main fragment. The anchorage fragments are prepared by digesting with XmaI or ApaI approximately 2 kbp-long, biotin- or digoxigenin-labeled DNA amplicons centered on the multi-cloning site of pBluescriptKS+.

magnetic tweezers were 3025 bp in length and were produced by PCR using plasmids pYY\_N1400\_BstEII (2) or pRS\_1N400\_BstEII, with an unlabeled forward primer and a biotin-labeled reverse primer. The amplicon was purified using a QIAQuick PCR Cleanup

kit (Qiagen, Germantown, MD). The plasmids were generated with Q5 Site-directed mutagenesis kit (New England BioLabs, Ipswich, MA) to mutate, in one case, the operator sequence of the  $O_{near}$  of pYY\_I1\_400\_BstEII from  $O_s$  to  $O_{3-}$  (a null operator,  $O_{null}$ ) to obtain pYY\_N1\_400\_BstEII plasmid. In a second case, the kit was used to mutate the  $O_{near}$  from  $O_s$  to  $O_1$  and  $O_{far}$  from  $O_1$  to  $O_{3-}$  to obtain the pRS\_1N\_400\_BstEII plasmid. The fragment generated by PCR contained the T7A1 promoter close to the upstream end, a  $O_1$  operator located either 669 or 261 bp downstream, the lambda t1 terminator 1298 bp downstream, and a biotin at the far, downstream end for attachment to a 1  $\mu$ m diameter streptavidin-coated bead (Dynabead MyOne Streptavidin T1, Invitrogen, Life Technologies, Grand Island, NY). The DNA template encoded only A, G and U ribonucleotides up to position +22, in order to be able to stall the transcription complex (TEC) by withholding CTP.

## § 2.2 Buffer recipes

### **PBS (Phosphate Buffer Saline) buffer**

10 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KH}_2\text{PO}_4$ , 137 mM NaCl, 2.7 mM KCl, pH=7.4

### **Prewash buffer**

10 mM Tris-HCl (pH=7.4), 200 mM KCl, 0.5 mg/ml  $\alpha$ -casein

**DNA incubation buffer**

10 mM Tris-HCl (pH=7.4), 200 mM KCl

**Lambda ( $\lambda$ ) buffer**

10 mM Tris-HCl (pH=7.4), 200 mM KCl, 5% DMSO, 0.1 mM EDTA, 0.2 mM dTT and 0.1 mg/ml  $\alpha$ -casein

**Basic-wash buffer (BWB)**

Glutamate BWB: 20 mM Tris-Glutamate (pH=8.0), 50 mM K[Glu], 1 mM dTT

Chloride BWB: 25 mM Tris-HCl (pH=7.4), 100 mM KCl, 1 mM dTT

**Transcription buffer (TXB)**

20 mM Tris-Glutamate (pH=8.0), 10 mM Mg[Glu]<sub>2</sub>, 50 mM K[Glu], 1 mM dTT, 0.2 mg/mL  $\alpha$ -casein

**§ 2.3 Microchamber Preparation**

Microchambers were prepared similarly to what has been previously described (142, 143). In brief, a microchamber with ~30  $\mu$ L volume was created between two glass

slides, 22 × 22 mm, No.1 and 24 × 50 mm, No. 1 for magnetic tweezers experiment, 22 × 22 mm, No.1 and 3" × 1" × 1 mm for tethered-particle motion experiment (Fisherbrand, Thermo Fisher Scientific, Waltham, MA) separated by a parafilm gasket with a narrow inlet and outlet to reduce evaporation of the reaction buffer, and a wide, central observation area (139, 144).

For DNA looping experiments using either TPM or MTs, DNA tethers were attached through a single digoxigenin at one end to the chamber bottom passivated with antidigoxigenin (Roche Life Science, Indianapolis, IN), 4 µg/mL at room temperature for 1 hr, or at 4 °C overnight, and at the other end to either a 320-nm-diameter streptavidin-coated polystyrene bead (Spherotech, Lake Forest, IL) (TPM experiment), or a 1.0-µm-diameter streptavidin-coated paramagnetic bead (Dynabead MyOne Streptavidin T1, Invitrogen, Life Technologies, Grand Island, NY). The chamber surfaces were passivated with 0.1 mg/mL BSA or 0.5 mg/mL α-casein solution to prevent non-specific sticking of the DNA and beads to the surface. The chamber was filled with a prewash buffer and stored in a sealed box to maintain high humidity box at 4°C. Before use the chamber was flushed with 200 µL of λ buffer.

The chamber for transcription experiments using magnetic tweezers was incubated with 5 or 10 µg/mL purified Anti-HA 11 Epitope tag antibody (16B12, monoclonal, Biolegend, San Diego, CA) in basic-wash buffer (BWB) at 4 °C overnight (≤ 24 hrs) or at room temperature for 1 hr. Then, the surface was passivated with 3 or 6 mg/mL α-casein (Sigma-Aldrich, St. Louis, MO) in BWB at room temperature for 1 hr, or 4 °C, overnight (≤ 24 hrs). Stalled elongation complexes (SECs) were produced by

incubating 25 nM doubly-HA tagged *E.coli* RNA polymerase (Karen Adelman Laboratory, NIH), approximately 10 nM DNA template, 50  $\mu$ M GpA RNA dinucleotide (initiating dinucleotide, Trilink BioTechnologies, San Diego, CA) in DEPC-treated water, and 10  $\mu$ M ATP/UTP/GTP in TXB at 37 °C for 30 mins. SECs were then drawn into the chamber and incubated 30 mins at room temperature to let the HA-labeled RNA polymerase bind to the anti-HA-coated surface. The far end of the DNA from the promoter was then labeled with a 1.0  $\mu$ m diameter, streptavidin-coated paramagnetic bead (Dynabead MyOne Streptavidin T1, Invitrogen, Grand Island, NY) by incubating it with 20  $\mu$ g/mL for 7 or 10 mins. The extension of the DNA tether was monitored after introducing 1 mM NTPs with/without 1 nM Lacl (Kathleen Matthews Laboratory, Rice University) in TXB.

## § 2.4 Tethered Particle Motion (TPM) experiments

All Lacl and HU titration TPM experiments were conducted in  $\lambda$  buffer at room temperature. For HU titration experiments, the concentrations of Lacl were fixed at 20 nM, 5 nM or 2.5 nM for Os-900-O1, Os-400-O1 and O1-400-O2 DNA, to give an approximately 50% looping probability for Os-900-O1 and 25% looping probability for both Os-400-O1 and O1-400-O2 with no HU.

A Leica DM LB-100 microscope (Leica Microsystems, Wetzlar, Germany) with an oil-immersion objective (63  $\times$ , NA 0.6-1.4) and differential interference contrast (DIC) was used to observe tethered beads through a CV-A60 video camera (JAI, Copenhagen, Denmark). The beads were tracked at 50 Hz using custom Labview (National

Instruments, Austin, TX) software, to record a time series of absolute XY positions of each bead. Vibrational or mechanical drift in the position of each tethered bead was removed by subtracting the position of one, or the average position of multiple, stuck bead(s) within the same field of view (145, 146). The DNA excursion for each was calculated as  $\langle \rho \rangle_{8s} = \langle \sqrt{(x - \langle x \rangle_{8s})^2 + (y - \langle y \rangle_{8s})^2} \rangle_{8s}$  with  $(x, y)$  representing the momentary position of a given bead and  $(\langle x \rangle_{8s}, \langle y \rangle_{8s})$  representing an eight-second moving average, effectively the anchor point of the bead. The mean square excursion of the bead was calculated using the formula  $\langle \rho^2 \rangle_{8s} = \langle (x - \langle x \rangle_{8s})^2 + (y - \langle y \rangle_{8s})^2 \rangle_{8s}$ . Changes in the extent of the motion, “excursion”, reflect conformational transformations of the tethered molecule (143, 146, 147).

Then, beads that exhibited clouds of  $(x, y)$  positions for which the ratio of the major and minor axes of an elliptical fitting was greater than 1.07, were discarded, since they were likely to have had multiple DNA tethers. Furthermore, beads with anomalously low excursion values were discarded to exclude tethers which became frequently stuck on chamber surface or did not exhibit the entire free range of motion. Seventy-four percent of the beads recorded exhibited symmetrical motion of the proper magnitude and were included in looping analyses (Appendix 2).

The time-series data for selected beads from the same experimental condition were pooled to generate a histogram of the observed excursions. The histogram was fit with up to three Gaussian distributions, representing the excursion values of one unlooped and up to two looped states mediated by LacI. Looping probabilities for



different conditions were calculated by dividing the area under Gaussian distribution of looped states by the total area under all three Gaussians.

## § 2.5 Magnetic tweezers (MTs) experiments

A magnetic tweezer (MT) was used to supercoil DNA to assay the dynamics of LacI-mediated loop formation, and to study the road block effect of LacI bound O1 operator to transcription. The MT consists of permanent magnets that can be translated along and rotated about the optical axis of the microscope to vary the strength of a laterally oriented magnetic field in the microchamber (148, 149). Single DNA tethers were identified through extension vs. twist measurements. Under a high tension of 2 pN, overwound DNA adopts a plectonemic form that smoothly reduces the DNA extension. In contrast, nicked DNA swivels about single bonds to relax any applied torsion, fails to form plectonemes, and the extension does not decrease. Stretching DNA tethers under high tension ( $\sim 2$  pN) detaches any DNA segments that are non-specifically stuck on the surface of the chamber or bead and extends a 2115 bp DNA template to approximately 0.7 microns.

For the measurements on the effect of DNA supercoiling on loop formation and those on topological domains, an established magnetic tweezers setup was used to record the  $x$ ,  $y$ ,  $z$  and  $t$  coordinates of one tethered bead and one non-specifically stuck bead using a custom MatLab (Mathworks, Natick, MA) routine at 10 Hz. In the DNA looping experiment, three-minute recordings under three levels of tension (0.25, 0.45,

or 0.75 pN) and a series of sigma values were measured to determine the DNA extension in each condition, before adding the LacI protein. After adding 1 nM LacI protein (provided by Kathleen Matthews, Rice University),  $x$ ,  $y$ ,  $z$  and  $t$  data were recorded for 20 mins at a selected tension and twist settings. The DNA extension vs. time data were then analyzed to identify probable looping events. A custom MatLab “change point” algorithm followed by an expectation-maximization routine (150, 151) was applied to determine the duration of looped ( $\tau_L$ ) and unlooped ( $\tau_U$ ) states, and the looping probability was calculated as:  $\sum \tau_L / (\sum \tau_L + \sum \tau_U)$ . The free energy change was calculated using  $\Delta G/kT = -\ln(\tau_L/\tau_U)$ .

DNA topology experiments were conducted recording DNA extension vs. turn curves (hat curves). DNA molecules without nicks formed plectonemes that reduced extension upon under- or over-winding at low tension (<0.4 pN) and were selected for analysis. The supercoiling of a DNA molecule can be quantified using the change in the linking number,  $\Delta Lk = Lk - Lk_0$  in which  $Lk$  is the measured linking number. The linking number of torsionally relaxed DNA,  $Lk_0$ , equals the number of base-pairs divided by the helical pitch (~10.4 bp/turn for B DNA). The change in the linking number includes both the excess twist and writhe,  $\Delta Lk = \Delta Tw + \Delta Wr$ , compared to torsionally relaxed DNA. DNA molecules were repeatedly wound and unwound, in steps of 2 turns, to change the linking number by  $\pm 10\%$  ( $\pm 20$  turns/(2115 bp/10.4 bp/turn)) at tensions of 0.25 or 0.45 pN, values that are estimated to be relevant for genomic DNA in physiological conditions (152, 153). Under 0.25 pN tension, the extension-versus-turns curves were symmetric. At 0.45 pN, extension-versus-turns curves were slightly asymmetric, due to partial conversion of

unwound DNA to left-handed forms (10, 154). The resulting extension-versus turns data was recorded, with and without 1 nM Lacl (10, 154, 155).

The real-time transcription experiments were conducted using a custom-built inverted microscope equipped with multiplexed magnetic tweezers. The microscope is similar in construction to several previously described (156, 157). The microscope was assembled using a Nikon plan 100 $\times$ /1.25 oil immersion objective (Nikon Instruments Inc. Melville, NY), P-721 piezo flexure objective scanner (PI Physik Instrumente LP Auburn, MA), an  $f = 160$  mm tube lens (Thorlabs Inc. Newton, NJ), and a Basler acA2000-185  $\mu$ m camera (IVS Imaging, Coppel, TX). Samples were illuminated using a custom LED (Luxeon Star LEDs, Quadica Developments Inc. Brantford, ON, Canada), brighted illuminator. The magnetic field generated by two 1/2'' $\times$ 1/4'' $\times$ 1/8'' Neodymium N52 grade magnets (K&J Magnetics Inc, Pipersville, PA), spaced 1 mm apart, attached to a steel hub, mounted on a vertical translation and rotation stage (custom design) along the bright-field beam path.

Real-time 3D particle tracking was implemented following a previously published scheme (158). The  $XY$  – location of each particle was tracked using a radial symmetry detection algorithm (159). The combination of objective, lens, and camera yielded a pixel resolution of 72.5 nm/pixel. With moderate image noise the radial symmetry algorithm localized particles to within 5-10% of a pixel, yielding an effective lateral accuracy of around 3-7 nm.  $Z$  – 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 ( $\text{argmin} \left[ \sum_r (I_r[z] - \hat{I}_r)^2 \right]$ ). For these experiments a

finite sampled lookup table ( $I_r[z_k]$ , with  $k=1,2,\dots$ ) was used. The sub-step height was calculated by fitting the squared intensity differences to a parabola and using the vertex as the best estimate of  $Z$ . Calibration experiments revealed that this scheme yielded a depth resolution of 10-20 nm. Microscope controls and 3D tracking software were written in MATLAB (Mathworks Natick, MA) and utilize Moco-manager ([www.micromanager.org](http://www.micromanager.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>. Extension-versus-time data were acquired at 164 Hz using a custom-built instrument. The single biotin label at the end far from the T7A1 promoter acted as a swivel to torsionally relax the tether during transcription. Before adding NTPs, turbulence lasting almost one minute produced spurious length measurements. When the turbulence subsides, many tethers returned to the previously measured extension value and shortly thereafter transcription elongation resumed, and the DNA extension decreased. 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. Pausing times were estimated by fitting sections of the time series with linear functions representing transcription before pausing, pausing and transcription after pausing. The duration of a pause was estimated as the distance between the intersections of each two fitting lines.

## § 2.6 Estimation of the HU concentration in an *E. coli* cell

*E. coli* is a rod-shaped bacterium measuring approximately 1  $\mu\text{m}$  in diameter and 2  $\mu\text{m}$  in length. The volume of a cylindrical *E. coli* cell is therefore

$$V = \pi r^2 L = 3.14 \times (0.5 \times 10^{-6})^2 \times (2 \times 10^{-6}) \text{ m}^3 = 1.57 \times 10^{-18} \text{ m}^3 \\ = 1.57 \times 10^{-15} \text{ L}.$$

The number of HU proteins in an *E. coli* cell (in logarithmic phase) has been estimated to range between 30,000 to 55,000 copies (17, 160). The number of moles of HU protein in an *E. coli* cell with 50,000 copies equals

$$5 \times 10^4 \text{ molecules} / 6.02 \times 10^{23} \text{ molecules} \cdot \text{mole}^{-1} = 8.3 \times 10^{-20} \text{ mole}$$

and the concentration of HU protein =  $\frac{8.3 \times 10^{-20} \text{ mole}}{1.57 \times 10^{-15} \text{ L}} = 5.29 \times 10^{-5} \text{ M} = 52.9 \mu\text{M}$ . Thus, the concentration range of HU protein is approximately 30 to 60  $\mu\text{M}$ .

## § 2.7 Exclusion of artifacts in the study of topological domains

Non-specific interactions could in principle decrease the length of the DNA tether and can be of three different types: 1) bead to surface, 2) DNA to surface or bead, 3) DNA-bound protein to surface or bead. We have ruled out the possibility that any of these could be the cause of the DNA length shortening we observed in this study. First, non-specific sticking of a DNA-tethered bead to the surface of the flow chamber would produce tether length measurements equal to zero. Such tethers were discarded from the analysis. Second, several observations indicated that transient, non-specific interactions between

DNA and either the surface of the flow-chamber or the bead did not occur in the analyzed data: i) control measurements in the absence of LacI (looping protein) showed only one state, such that the extension-versus-turns curves recorded in this condition had reproducible center and height values; ii) decreases in the extension of the tethers were greater than or equal to the expected length of the loop segment,  $\Delta L$ , under the applied tensions. Random, nonspecific sticking would not be expected to exclude smaller decreases; iii) The clear correlation between the tether length decrease and number of turns trapped in the LacI-mediated loop would not result from non-specific interactions that caused temporary, sticking of random segments of the DNA to either the surface of the flow-chamber, or the bead. Third, no interaction has ever been reported between LacI and streptavidin, or antidigoxigenin, and observations that allow us to exclude DNA sticking to surface artifacts (ii and iii) allow us to exclude protein-mediated DNA sticking as well.

Finally, non-specific binding of LacI to DNA is negligible (if not non-existent) in the buffer condition and LacI concentration used, as seen by atomic force imaging (161, 162) and tethered particle microscopy assays where, in the absence of supercoiling, the DNA tether length in the presence of LacI remained constant in time (89, 90, 139, 163, 164).

## **§ 2.8 Measuring the coiling and extent of topological domains**

The number of turns trapped by the LacI-mediated DNA loop was indicated by the relative shift of the peak in the extension-versus-turns curves recorded with and without LacI

protein (Fig. 6.1 A) (165). The size of the topological domains was calculated as the difference,  $\Delta L$ , between the peak height of the extension-versus-turns curve with and without protein (Fig. 6.1 A).

## § 2.9 Obtaining the gyre size

The slope of extension-versus-turns curves depends on the applied tension. As DNA is twisted beyond the buckling transition, the point at which it has absorbed the maximum torque and buckles into a form with writhe, a portion becomes plectonemic and the extension linearly decreases until the entire tether is plectonemic. After this point, the curve will flatten, because the proximity of the microsphere to the surface prevents detection of any further length changes. To analyze the dependence of  $\Delta L$ , (the difference between the peak height of the extension-versus-turn curve with and without protein-mediated loop) on trapped turns, the gyre size,  $l_{gyre}$ , needed to be determined. This was done by dividing the maximum extension in the hat curve by the number of turns introduced.  $l_{gyre}$  was then used to estimate the values of minimum  $\Delta L$  according to equation 6.1.

# Chapter 3 Protein concentration, operator affinity and loop size dependence of LacI-mediated DNA looping

## § 3.1 Introduction

Interactions between proteins bound to separate sites on the same DNA molecules induce DNA looping, which are critical in gene regulation and other (166-169). The formation of loop is influenced by protein concentration, binding affinities of protein-DNA interacting sites (operators), and the loop size (4, 68). In vivo, the separation of interacting sites ranges from a few base pairs to hundreds of kilobase pairs. At short distances can compare to DNA persistence length, DNA bending affinity inhibit the protein at one operator site to find the other site in 3D space, thus DNA tethering is weak (70, 71). Also, when the separation between the sites is too large, the two sites need to overcome entropy meet each other in space, so the DNA tethering is also low (76, 77). Only in an intermittent loop size, two operators can meet each other. The effect of DNA tethering can be quantified by  $J$ -factor  $J_{loop}$ , the effective molar concentration of looping protein near one operator site when the other site is occupied by the looping protein, or as the free energy of looping reaction  $\Delta G_{loop}$  (71, 110, 170, 171). These two parameters are interconvertible with  $\Delta G_{loop} = -RT \cdot \ln J_{loop}$ . The formation of naked DNA loop is an energetically unfavorable reaction under

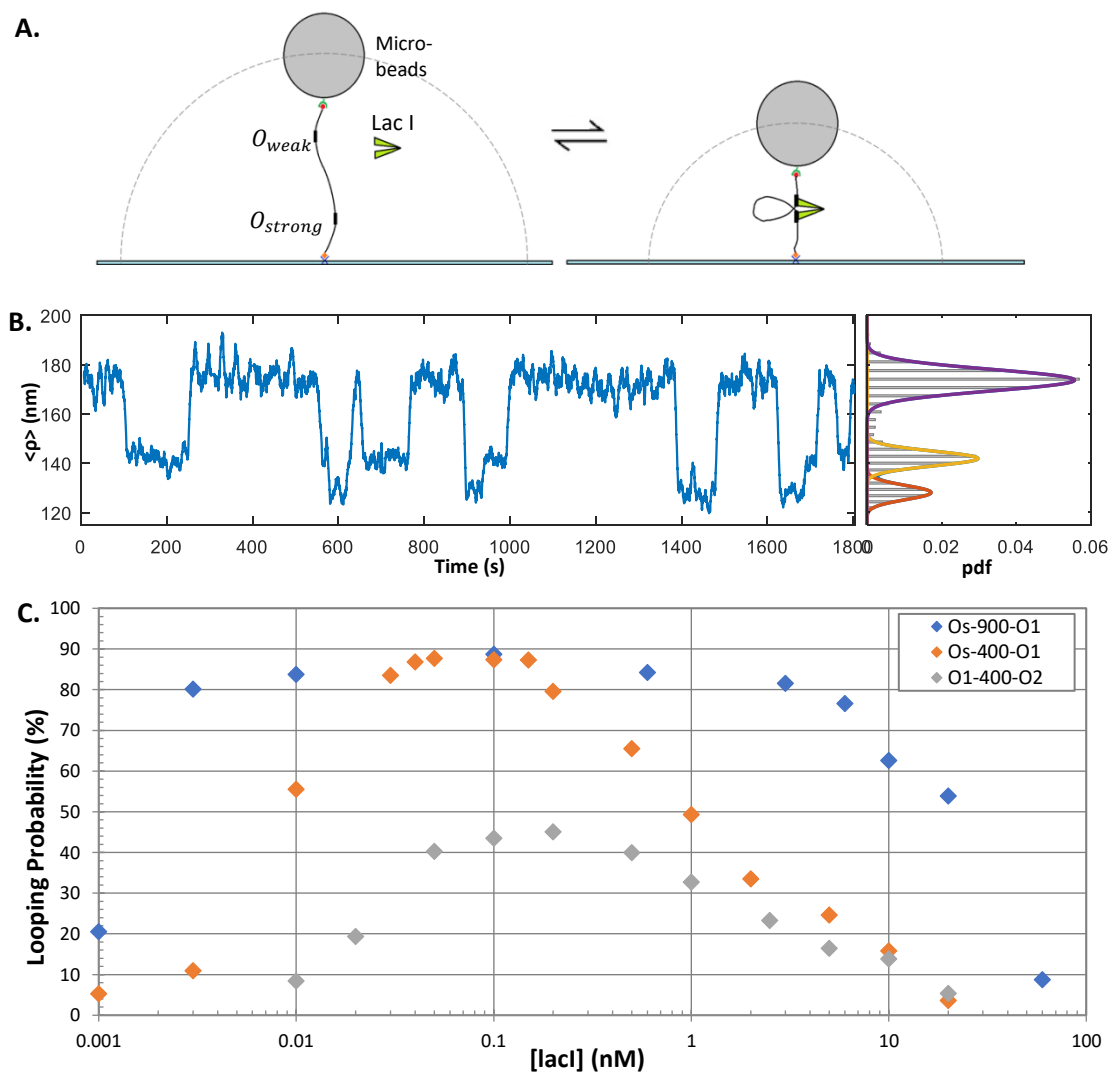


physiological condition due to the enthalpic cost of DNA bending and twisting (particularly important for short DNA segments) and the entropic cost of restricted configurational freedom of the DNA (the major limitation for long DNA loops).

As mentioned in section 1.4, the Lac repressor (LacI) may bind to one of several operator DNA sequences, known as wild type operators O1, O2, O3, and to the engineered operator Os. The approximate binding affinity of Os is 5 times as O1, O1 is 4 times as O2, O2 is 67 times as O3 (172). Since a LacI tetramer contains two DNA-binding domains, it can induce DNA looping by simultaneously binding to two operators. Since O1 overlaps with the promoter for the Lac operon, and it's the strongest operator, the binding of LacI to O1 most significantly blocks transcription. When LacI dissociates from O1, instead of diffusing away, it can bind to either O2 or O3 and close back to O1 reforming the loop. Thus, the existence of O2 and O3 increase the local concentration of LacI near O1.

The probability of forming a loop will increase as the Lac I concentration increase. However, if the concentration of Lac I is too high, the probability of forming a loop is low because the two operators are occupied by separate transcriptional factors. At intermediate concentration, the looping state has its highest probability (69, 101).

In this chapter, I investigate the effect of different factors, LacI concentration, operator affinity and loop size on LacI mediated DNA looping. In particular, I used the tethered particle motion (TPM) to measure loop formation and breakdown over time (Fig. 3.1 A, B). The looping probability can be calculated by dividing the time spend in the



**Fig. 3.1** Experimental schematics, representative data and statistical result. (A) A schematic representation of tether changes due to LacI-induced DNA looping during a TOM experiment (not to scale). (B) A representative recording of excursion versus time shows intermittent changes in the average excursion of the bead as LacI-induced looping changes the overall tether length in a 909 bp Os-400-O1 DNA tether. At right, a histogram of the excursion values for the recordings exhibits three peaks which represent the two looped states and the unlooped state fitted with Gaussian curves, shown in red, yellow, and purple respectively. Similar data were recorded for 1632-bp-long, and 831-bp-long O1-400-O2 tethers (not shown). (C) The looping probability was measured as a function of LacI concentration for Os-900-O1, Os-400-O1, and O1-400-O2, Os-O1 looping probability peaked at 87% while O1-O2 only reached 45%.

looped state with the overall recording time. Furthermore, I calculated the  $J$ -factor  $J_{loop}$  and compared our data with experimental results from a collaborator's group and with a theoretical estimation (71). My observations were closer to the theoretical than the experimental result.

## § 3.2 Results

### § 3.2.1 Two looped states are observed in the experimental trace

In TPM, the amplitude of the Brownian motion of a bead tethered by a single DNA molecule to the glass surface of a microscope flow-chamber (Fig. 3.1, A) is monitored over time. A typical experimental recording from TPM assays for three different constructs, Os-900-O1, Os-400-O1, O1-400-O2 (written as  $O_{strong}$  – Loop Size (bp) –  $O_{weak}$ ; tether length 1632, 909 and 831 bp, with centrally located LacI-inducible loops) exhibited intermittent switching of the excursion of beads between low and high levels (Fig. 3.1, B). In agreement with previous results (4, 173, 174), these corresponded to looped and unlooped DNA configurations (Fig. 3.1, A). As seen in the figure, there are two clearly distinct looped states as seen both in the trajectory and the histogram (Fig. 3.1, B), which are similar to previous experimental result (4, 99, 101, 175, 176). Control experiments done by the Phillips group with only one of two binding sites showed only the highest peak, which supports the idea that the two lower peaks indicate looped configurations (101). One hypothesis is that these two looped states reflect the “open” and “closed” configuration of the Lac repressor molecule. Direct interconversion between the two looped species suggested the two distinct looped states are indeed

due to different conformations of Lac repressor protein (Fig. 3.1, B) (99). An alternate hypothesis is that the two peaks correspond to different DNA topologies, parallel and anti-parallel of DNA phase (61, 65, 69).

### § 3.2.2 LacI concentration dependence of protein mediated DNA looping

In order to extract quantities such as the free energy of looping associated with repressor binding (or equivalently, a  $J_{loop}$  for looping) and to examine how the propensity for looping depends upon the number of repressors, we needed looping data at a number of different concentrations.

One way to characterize the looping probability as a function of concentration is shown in Fig. 3.1 C. There are various ways to obtain data of the sort displayed in this plot. First, by examining the trajectories, we can simply compute the fraction of time that the DNA spends in each of the different states, with the looping probability given by the ratio of the time spent in either of the looped states to the total elapsed time. To compute the time spent in each state, a threshold has to be established to decide when each transition has occurred. This can be ambiguous, because trajectories sometimes undergo rapid jumps back and forth between different states; it is not always unequivocally clear when an apparent transition is real, and when it is a random fluctuation without change of looping state. A second way of obtaining the looping probability is to make a histogram of the time trajectories, fit it with three Gaussian and compute the areas under the different peaks and use the ratios of areas as a measure of

looping probability. This method, however, does not properly account for possible variation between different beads and the life time of looped and unlooped states, because they are all added up into one histogram. A third alternative is to obtain the looping probability for each individual bead, by plotting its histogram and calculating the area under that subset of the histogram corresponding to the looped states, this method can calculate the mean looping probability and the standard error for each construct, but this method is critical for the length of each measurement, because the short trace may not be sufficient to obtain a result with statistical significance. In this chapter, looping probability were analyzed according to the second method described above (also described in methods) to determine the fractions of time spent in looped and unlooped states.

LacI titration measurements exhibit three looping probability regimes. At very low concentration, we expect that there will be negligible looping because neither of the operators will be bound by Lac repressor. At intermediate concentrations, the equilibrium will be dominated by states in which a single repressor tetramer is bound to the DNA at the strong operator, punctuated by transient looping events. In the very high concentration limit, each operator will be occupied by a separate tetramer, making the formation of a loop nearly impossible (Fig. 3.1, C).

### **§ 3.2.3 Operator affinity dependence of protein mediated DNA looping**

The maximum looping probability in each DNA construct occurred at intermediate LacI concentration, the values depend on operator affinity. For both the Os-900-O1 and the Os-400-O1 tethers, the looping probability reached a maximum of almost 90 %, at a LacI concentration of 0.1 nM. While for O1-400-O2 construct, the maximum probability of looping, which occurred at approximately 0.2 nM (higher concentration than Os-O1 operators) of LacI concentration, was 45 % (Fig. 3.1, C). In the following paragraph, we compute the  $J_{loop}$  factor of a certain loop, which is a parameter to describe how easy a loop form based on a piece of polymer. The  $J_{loop}$  factors for Os-400-O1 and O1-400-O2 are the same. Thus, the difference of maximum looping probability for different polymer is due to the difference of  $K_d$  values for Os and O1 compared to O1 and O2, not the loop size, nor the different sequence of the loop fragment. When the LacI concentration is high, the titration curve for both Os-400-O1 and O1-400-O2 overlap.

#### **§ 3.2.4 Loop size dependence of protein mediated DNA looping**

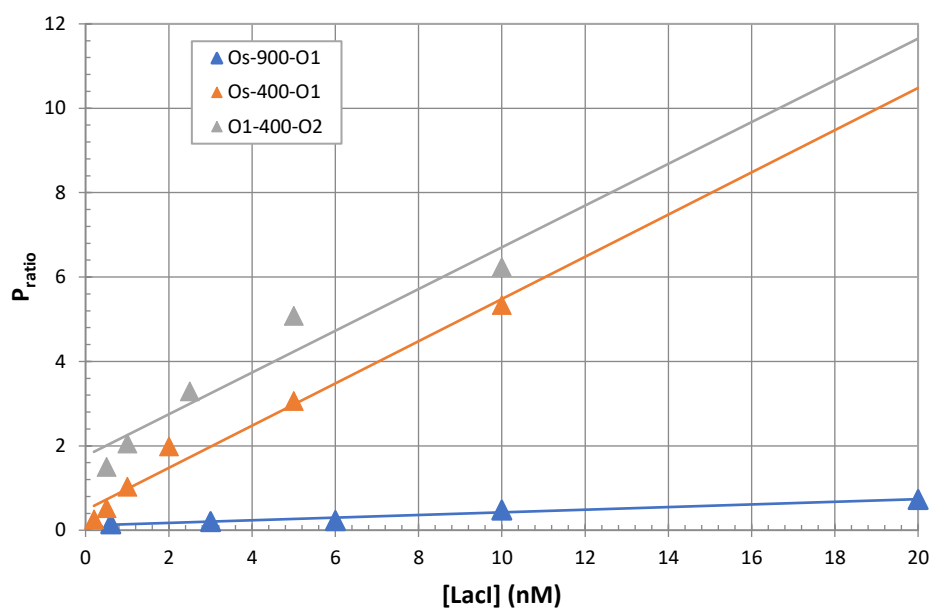
The fraction of time spent in the looped configuration is controlled by several competing effects. Suppose that a repressor tetramer is bound to the stronger operator; shortening the inter-operator spacing reduces the volume over which the other operator wanders relative to the second binding site on the repressor, increases the apparent local “concentration” of free operator in the neighborhood of that binding site, and hence enhances looping. But decreasing the inter-operator spacing also has the effect of discouraging looping, due to the large elastic energy cost of forming a shorter loop.

Moreover, a shorter overall DNA construct increases the entropic force exerted by bead-wall avoidance, again discouraging looping (94).

In our measurement, the range of high looping probability for the constructs varied according to loop size. The Os-900-O1 construct looped with 90% probability across four decades of LacI concentration, while Os-400-O1 reached similar level across only two decades (Fig. 3.1, C). This is counterintuitive to the theoretical result that loops of 400 bp form more easily than longer ones such as 900 bp (71). Perhaps, it results from longer sequences flanking the operator, as in the 900 bp loop case, which have been to enhance the operator binding affinity (177). When a LacI dissociates from the operator, instead of diffusing into space, a long flanking sequence help to keep LacI in the vicinity of the operator and helps the LacI protein rebind to the operator.

To see what our measurement of this looping equilibrium tells us, we therefore needed to calculate in some detail the expected local concentration of operator (the “looping  $J_{loop}$  factor” based on a particular mathematical model of DNA elasticity. Note that the  $J_{loop}$  factor is that it is independent of the particular binding strengths of the different operators. The value of the  $J_{loop}$  factor can be determined by fitting the decreasing section of the looping probability versus [LacI] curve at LacI concentrations with the equation  $P_{ratio} = \frac{p_{unlooped}}{p_{looped}} = \frac{2K_{weak}}{J_{loop}} + \frac{2[LacI]}{J_{loop}}$ , where  $p_{unlooped}$  is the fraction of time the DNA is unlooped,  $p_{looped}$  is the fraction of time the DNA is looped, and  $K_{weak}$  is the dissociation rate of the weak operator. The slope of each  $\frac{p_{unlooped}}{p_{looped}}$  vs [LacI] plot equals  $\frac{2}{J_{loop}}$  (4). Figure 3.2 shows the fitting, from the slope we can calculate the  $J_{loop}$

values for three tethers are 64 nM, 4 nM and 4 nM for Os-900-O1, Os-400-O1 and O1-400-O2, respectively.



**Fig. 3.2** Calculation of  $J_{loop}$  factor.  $J_{loop}$  for Os-900-O1, Os-400-O1, O1-400-O2 were determined from the ratio of fraction of time the DNA is unlooped ( $p_{unlooped}$ ) versus the fraction of time the DNA is looped ( $p_{looped}$ ) at each  $[LacI]$ . According to the equation of  $\frac{p_{unlooped}}{p_{looped}} = \frac{2K_{weak}}{J_{loop}} + \frac{2[LacI]}{J_{loop}}$  (4), the slope of the fitting plot value of  $2/J_{loop}$ .

### § 3.3 Discussion

The existence of two looped states may be interpreted with two different hypotheses, one is that these two looped states reflect the “open” and “closed” configurations of the Lac repressor molecule. Direct interconversion between the two looped species suggested the two distinct looped states are indeed due to different conformations of the Lac repressor protein (Fig. 3.1, B) (99). The “open” and “close” configuration of LacI was observed by AFM imaging, which allowed to determine the size of an “open” LacI



protein to be about 18 nm and of the “close” configuration about 13 nm (2). However, using a calibration equation (129), we obtained that the distance of two looped peaks is 40 nm. Thus the Lacl conformation alone cannot explain the two looped states observed. Instead, the two peaks may correspond to the parallel and anti-parallel DNA loop topologies (61, 65, 69), or to a combination of protein configuration and loop topology. In fact, the change of Lacl configuration may lead to the change of DNA topology, and vice versa.

We have shown here that the looping probability depends on the operator binding affinity. As previous (69) theoretical predictions and experimental results had suggested, the maximum value of the Lacl concentration titration curve shifts to the right as the dissociation constant from operators increases. Changing the strength of the operators should change both the concentration at which looping is maximal, and the amount of looping at the maximum. This observation can be formalized by applying the

equation 
$$P_{loop} = \frac{\frac{1}{2K_{weak}K_{strong}} [R]J_{loop}}{1 + \frac{[R]}{K_{weak}} + \frac{[R]}{K_{strong}} + \frac{[R]}{K_{weak}K_{strong}} + \frac{[R]}{2K_{weak}K_{strong}} + \frac{1}{2K_{weak}K_{strong}} [R]J_{loop}}$$
 (see introduction 1.4.2). The

concentration at the maximum in the looping probability can be found by differentiating above equation with respect to  $[R]$  to obtain,  $[R]_{max} = \sqrt{K_{strong}K_{weak}}$ . Note that the

concentration at which the looping probability is maximized does not depend upon the DNA flexibility as captured in the parameter  $J_{loop}$  factor. However, the maximum looping

probability depends on  $J_{loop}$  factor, as 
$$p_{loop}([R]_{max}) = \frac{J_{loop}/2}{J_{loop}/2 + (\sqrt{K_{strong}} + \sqrt{K_{weak}})^2}$$
. Also,

at high Lacl concentration levels, the looping probability approaches  $J_{loop}/(2[R])$ ,

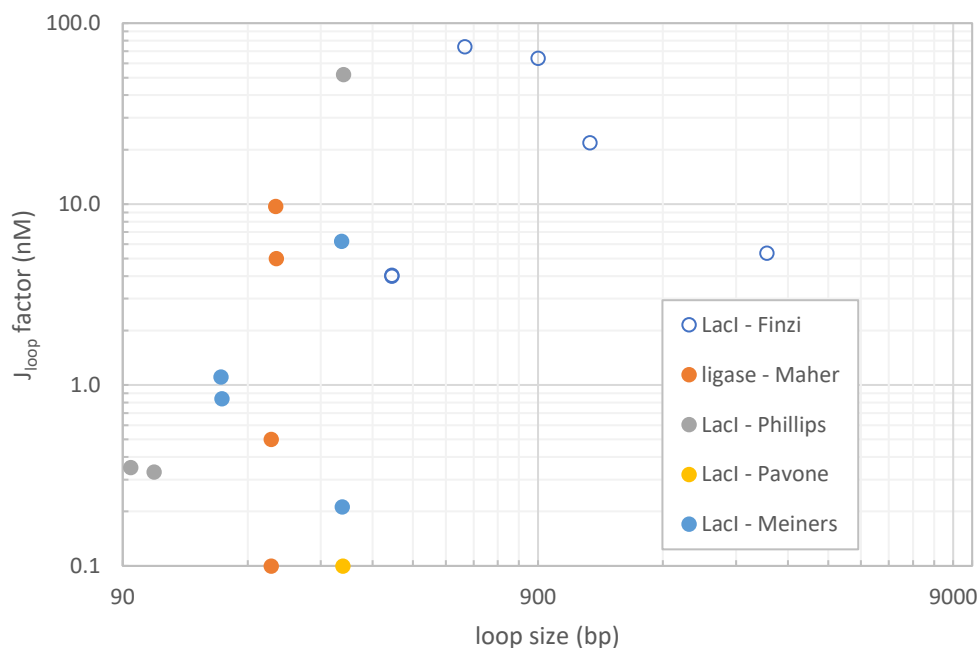
which is independent of the operator strength, but depends on DNA flexibility through the parameter  $J_{loop}$ , explaining why in figure 3.1 (C) the Os-400-O1 and O1-400-O2 curves overlap, in the high Lacl concentration regime, while the Os-900-O1 curve is shifted. Thus, data at low concentrations are essential for determining operator strengths, whereas high concentration data are sufficient for determining  $J_{loop}$  factor.

For the 400 bp loops, I measured a  $J_{loop}$  factor of 4nM, and for the 900 bp loop  $J_{loop} = 64$  nM, which indicates that the 900 bp loop forms more easily than the 400 bp loop. This is inconsistent with the theoretical prediction that the 400 bp loop should have a  $J_{loop}$  factor slightly higher than that of the 900 bp loop (71). The difference may be explained considering that the simulation is based on the assumption that the size of the protein is zero, thus the two ends of the loop segment contact each other. However, the size of Lacl protein is not zero. As we have seen above, Lacl may exist in the “open” and “close” conformation (2) and AFM images from our group show that, when the Lacl is close, size of the protein is about 13 nm, open Lacl conformation, size of the protein is about 18 nm. In ds-B-DNA, one base pair has a length of 0.34 nm and the Kuhn length is  $l = 100$  nm at physiological salt concentration. Thus, the  $J_{loop}$  factor for linear dsDNA

$$J_{loop}(b) = 2.7 \times 10^{-3} \times b^{-3/2} \times \exp\left(\frac{d-2}{1.2 \times 10^{-5} \times b^2 + d}\right) \text{ mol/liter (71)},$$

where  $b$  is the number of base pair of DNA,  $d$  is related to the end-to-end distance of the loop fragment when the loop forms. With this equation, the effective local concentration of the Lacl protein near the weak operator can be calculated. Since the end-to-end distance of the DNA loop,  $r$ , is zero, the value of  $d$  is zero. When  $r = 10$  nm,  $d = 0.13$

(178). Thus, the size of the LacI protein might be the reason for the difference between the theoretical and experimental values of  $J_{loop}$ .



**Fig. 3.3** Comparison of  $J_{loop}$  factor values with previous literature. Finzi group show LacI-mediated loops with loop size larger than a threshold which need to overcome entropy to form a loop. Other data show scattered points with loop size smaller than that threshold which made the formation of loop need to overcome bending affinity.

The  $J_{loop}$  factor for the lac loop has been evaluated by different groups during the past few years (4, 69, 101, 103, 130, 179, 180). Kumar et al. in the Finzi Lab used TPM to measure  $J$  factor values for LacI loops large enough that the loss of entropy hinders loop formation (4). This study reports that  $J_{loop}$  factors decrease as loop size increase past the length ideal for loop formation, with the values of 74.1 nM, 64.0 nM, 21.8 nM, and 5.4 nM measured for 600 bp, 900 bp, 1200 bp and 3200 bp of loop size (Fig. 3.3). However, for loops smaller than that threshold, flexibility and twistability are important for the

formation of loop. Instead of changing linearly with the loop size,  $J_{loop}$  factors from different research groups show scattered points (Fig. 3.3) (69, 101, 103, 130, 179, 180).

I reported that the  $J_{loop}$  factors for two different random 400 bp loops sequence with different operators are 4.0 nM. Thus,  $J_{loop}$  factors are independent of the strength of operators. Revalee et al. from the Meiners group show 303 bp, 304 bp has  $J_{loop}$  value of 6.2 nM and 0.2 nM (103), Vanzi et al. from the Pavone group give  $J_{loop}$  value for 305 bp is 0.1 nM (130), Han et al. from Phillips group show  $J_{loop}$  value for 306 bp is 52 nM (101). Also, Peters et al. from Maher group gave  $J_{loop}$  value is 0.1 or 0.5 nM for 205 bp loop, 9.7 nM and 5.0 nM for 210 and 211 bp loops, respectively (179, 180). Revalee et al. from Meiners group also show that 155 bp and 156 bp-long loops have  $J_{loop}$  values of 1.1 nM and 0.8 nM, respectively (103). Johnson et al. from the Phillips group showed that 94 bp and 107 bp-long loops have a 0.4 and 0.3 nM  $J_{loop}$  value, respectively (69). In most of the situations listed above, a difference in loop size of a single base pair significantly affected the value of the  $J_{loop}$  factor. This may due to the fact that the phase of operators influences the loop formation in the loop size regime where DNA twisting is energetically costly (58, 69, 101, 181). In vivo experiments showed that gene expression levels change periodically as the loop size changes by single base pairs, with a period of 10.4 bp, which is consistent with the helical pitch of B-DNA (58, 69, 101, 181).

# Chapter 4 Quantitation of interactions between two DNA loops demonstrates loop domain insulation in *E. coli* cells

## § 4.1 Introduction

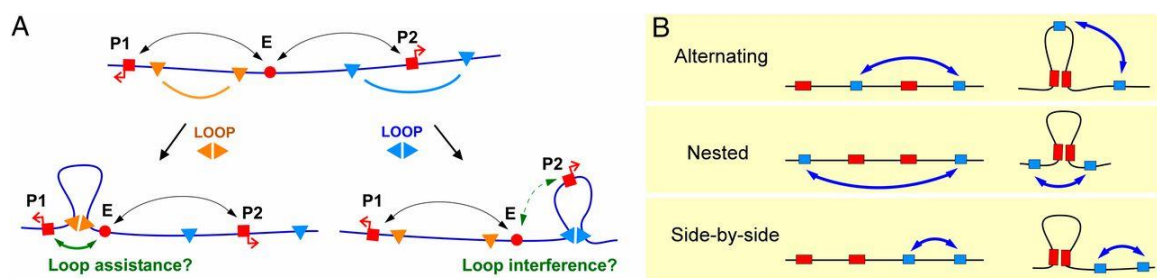
Transcription of genes is regulated by promoter-proximal DNA elements and distal DNA elements (enhancers) that together determine condition-dependent gene expression. In eukaryotic genomes, enhancers can be located many hundreds of kilobases away from the promoter they regulate (182-184), and the intervening DNA can contain other promoters and other enhancers (185-188). How the regulatory influence of distal elements is exerted efficiently and specifically at the correct promoters is poorly understood.

Enhancers are clusters of binding sites for transcription factors and chromatin-modifying enzymes, and activate promoters by directly contacting them via DNA looping (166, 189-192). Enhancer-trap approaches, mapping of transcription factor binding and chromatin modifications have identified tens of thousands of enhancer elements in metazoan genome (188, 193-196). Chromatin capture studies show that enhancers and promoters are connected in highly complex condition-dependent patterns (187, 195, 197). Although core enhancer and promoter elements can provide some specificity (198), enhancers are often able to activate heterologous promoters if they are placed near to each other. Indeed, this lack of specificity is the basis for standard enhancer

assays and screening assays (188, 194, 199). Thus, additional mechanisms are clearly needed to target enhancers to the correct promoters over long distances and to prevent their interaction with the wrong promoters. Dedicated DNA-looping elements that can either assist or interfere with enhancer–promoter looping are thought to play a major role.

In theory, any DNA loop that brings the enhancer and promoter closer together should assist their interaction (Fig. 4.1 A), because the efficiency of contact increases, as the length of the DNA tether between the sites shortens (4, 71, 117, 178, 200).

Promoter-tethering elements in *Drosophila* that allow activation by specific enhancers over long distances are proposed to form DNA loops between sequences near the enhancer and the promoter (198, 201). In the mouse  $\beta$ -globin locus, the Ldb1 protein binds to proteins at the locus control region and at the promoter and appears to form a bridge necessary for efficient enhancer–promoter contact (202). In bacteriophage  $\lambda$ , the



**Fig. 4.1** Interactions between DNA loops. (A) DNA-looping interactions between sites on the DNA, e.g., between an enhancer (E) and promoters (P1 and P2), are proposed to be affected by other DNA loops. Specific interactions between looping elements (triangles) can either assist enhancer–promoter looping by bringing the enhancer and promoter closer together (orange triangles) or are thought to interfere with enhancer–promoter looping by placing them in separate loop domains (blue triangles). (B) The three possible topological arrangements of two pairs of interacting sites on DNA. We tested in each case whether the formation of a loop between one pair of sites affects the propensity of the other sites to interact (and vice versa).

CI protein forms a 2.3-kb DNA loop that brings a distal stimulatory site close to RNA polymerase at the PRM promoter (203). Enhancer–promoter targeting has also been demonstrated on plasmid constructs using heterologous looping proteins—e.g., with  $\lambda$  CI in human cells (204) and the *Drosophila* GAGA protein in human cells and in yeast (205, 206).

DNA looping also seems to be able to inhibit enhancer–promoter contact. Enhancer-blocking insulators are defined by their ability to prevent enhancer activation of the promoter when placed between the enhancer and the promoter. A large body of evidence is consistent with the idea that insulators work by binding proteins that form DNA loops to other insulators (207-211). This mechanism of insulator action is rationalized by the loop domain model (212), which proposes that the formation of a DNA loop creates a separate topological domain that inhibits interaction between any site within the loop and any site outside the loop (Fig. 4.1 A). This model is the only one that can currently explain the requirement that an insulator must be between the enhancer and promoter to block activation, as well as observations that two insulators between the enhancer and promoter sometimes do not block activation (213, 214). The loop domain model is also a potential explanation (211) of the topologically associated domains (TADs) revealed by genome-wide mapping of DNA contacts by chromatin capture methods in genomes from mice to bacteria (215-221). Individual DNA sites between two domain boundaries interact more frequently with each other than they do with individual DNA sites in other TADs—i.e., domain boundaries act like insulators. Consistent with the loop domain model, domain boundaries interact with each other at

high frequency and often contain known DNA-looping elements, being enriched for insulator protein binding sites, as well as for active promoters and enhancers (220).

Despite the fundamental significance of the loop domain model in explaining insulator action and the formation of TADs, unequivocal tests of the model *in vivo* have been hampered by the complexity of eukaryotic genomes and their gene regulatory elements. Evidence from more defined systems has supported the loop domain model. In a plasmid transfection system, a 344-bp DNA loop formed by a Tet repressor derivative around the SV40 enhancer inhibited its activation of a promoter 2 kb away (222). However, the small loop may have affected enhanceosome assembly in these experiments. In an *in vitro* system with *E. coli* proteins and supercoiled plasmids, a 630-bp DNA loop formed by the Lac repressor (LacI) around the NtrC enhancer element inhibited its activation of the *glnA* promoter 2.5 kb away (223). However, in both studies, the lack of information about DNA-looping efficiencies prevents a quantitative analysis of loop interference.

To clearly test the loop domain model *in vitro* by the single molecule technique, tethered particle motion (TPM). Here, we combined LacI and CI DNA loops in each of the three possible topologies (Fig. 4.1 B) to show that alternating DNA loops interfere, nested DNA loops assist, and side-by-side loops do not affect one another's formation. Fitting our data to a general statistical–mechanical model of loop interaction allowed calculation of the strength of looping assistance or interference. Our collaborator in Australia conducted experiments *in vivo* measured interactions between large DNA



loops formed in the *E. coli* chromosome by LacI and bacteriophage  $\lambda$  CI. This chapter include the in vitro work, in vivo part done by our collaborator please reference (176).

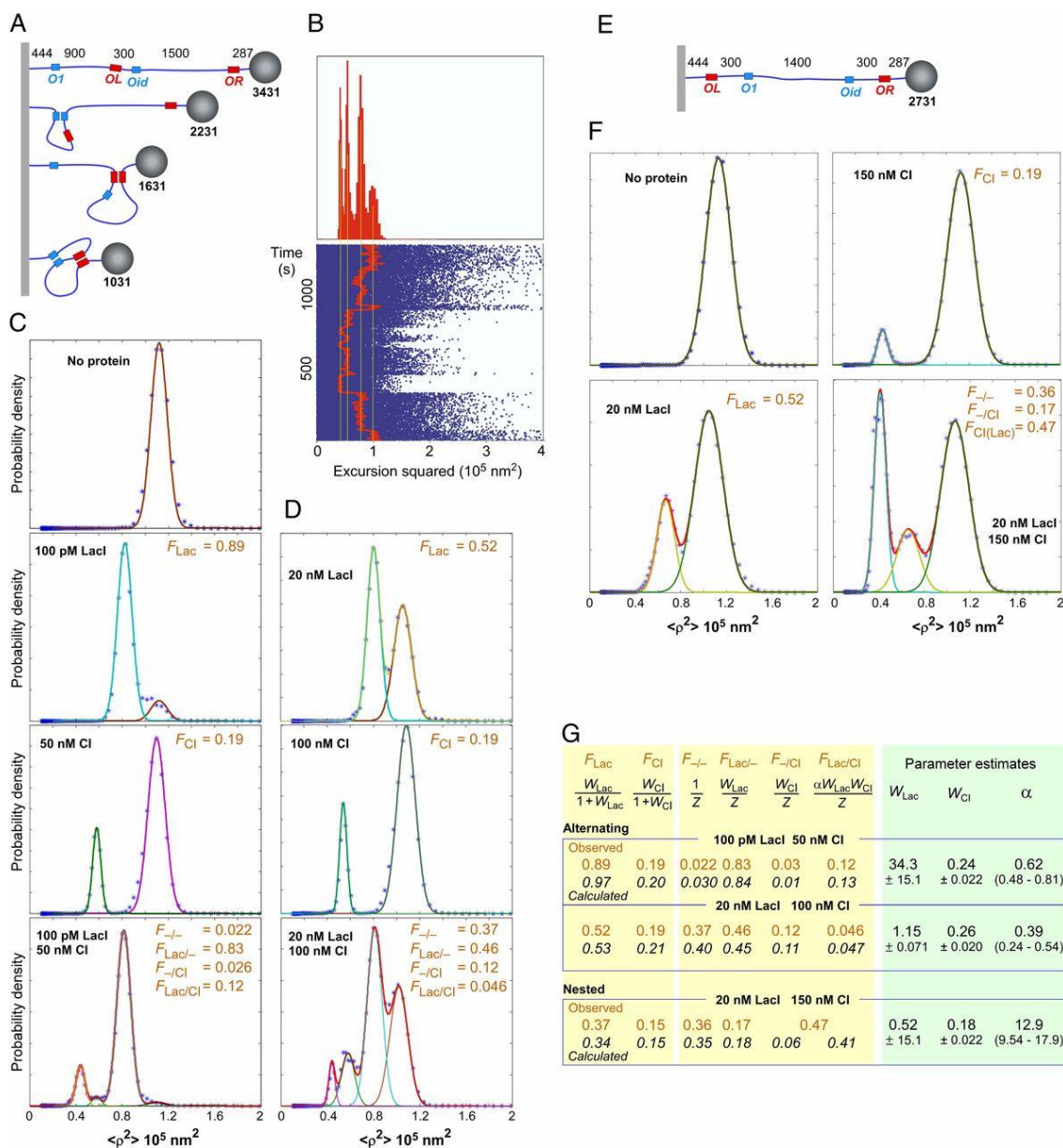
## **§ 4.2 Results**

### **§ 4.2.1 Quantitation of interactions between DNA loops**

There are three topological ways to arrange two pairs of DNA-looping sites: alternating, nested, and side by side (Fig. 4.1 B; note that we ignore the parallel/antiparallel orientation of the strands at the loop clamps). The expectation from the loop domain model is that in the alternating arrangement, the formation of one loop will interfere with the formation of the second loop. For nested loops, we expected that the formation of one loop would assist the formation of the second. If the interior loop forms first, then the linear distance (along the DNA), between the exterior loop sites becomes shorter. When the exterior loop is formed, the distance between the interior loop sites may or may not shorten (depending on the geometry), but they are nevertheless likely to become more spatially constrained because they become linked by a DNA tether on each side. We do not expect side-by-side loops to affect each other.

### **§ 4.2.2 Alternating loops give loop interference**

We used the TPM technique to detect and measure DNA looping in vitro (4, 129, 224, 225). For loop interference, we used a DNA tether with the asymmetrical looping



**Fig. 4.2** TPM analysis of alternating and nested loops. (A and E) Setup of the TPM experiments to measure loop interference or assistance in vitro (not to scale). Distances in base pairs. DNA was attached to the coverslip by digoxigenin and to the bead by streptavidin. (B) An example of switching between tether lengths corresponding to unlooped, Lacl or Cl looped, or both Cl and Lacl looped states for one tether with 20 nM Lacl and 100 nM Cl over a 1,350-s observation interval. (Lower) Observed values of excursion squared (blue) along with an 8-s moving average (red) during an observation interval of 1,350 s. (Upper) Histogram for the entire observation in which four states corresponding to, from left to right, the doubly looped, Cl looped, Lacl looped, and unlooped tether appear. (C and D) Histograms of bead displacement  $\langle \rho^2 \rangle$  for the alternating arrangement.  $F$  values were obtained as the area under each peak by Gaussian fitting. Histograms were compiled from analysis of 12–61 beads

(total 109–1,33x2 min) under each condition. (F) As C and D, but for the nested arrangement. CI-His6 was used instead of CI. Histograms were compiled from analysis of 25-45 beads (total 224-877 min) under each condition. (G) Statistical-mechanical model of loop interaction (Fig. 4.5). The observed  $F$  values (brown) from C, D, and F, and the calculated  $F$  values (italic) using the equations shown and the fitted loop weights and  $\alpha$  for each construct (*Right*).  $W_{\text{Lac}}$ ,  $W_{\text{CI}}$ , and  $\alpha$  were fitted to minimize the difference between the observed and calculated  $F$  values.  $W_{\text{errors}}$  are SDs, and the ranges for the  $\alpha$  estimates are the 2.5–97.5 percentiles (>900 fitting runs).

arrangement (Fig. 4.2, A). This arrangement had been designed so that loops by LacI, CI or both could be distinguished by their effect on the tether length (Fig. 4.2 A), and therefore on the mean displacement of the attached bead. Recordings of the tether lengths vs. time for individual tethers exposed to both proteins exhibited discrete stepping between values expected for the unlooped and all three looped forms (Fig. 4.3 B). By following multiple beads over time in the presence of LacI or CI, histograms of the probability vs. tether length were compiled (Fig. 4.3 C and D; SI). Two combinations of LacI and CI concentrations (100 pM LacI + 50 nM CI and 20 nM LacI + 100 nM CI) were used, and each protein was also used alone at the same concentration. Estimates of the fraction of each looped species were obtained by fitting the histograms to Gaussian curves. Because it was possible to resolve all of the looped species, six  $F$  values were obtained for each condition (Fig. 4.3 B and C).

These data were analyzed using the statistical–mechanical loop interaction model to extract loop weights and  $\alpha$  (Fig. 4.3G). An  $\alpha$  of 0.62 (0.48–0.81) indicates ~1.5-fold loop interference for the 100-pM LacI, 50-nM CI data, and an  $\alpha$  of 0.39 (0.24–0.54) indicates ~2.5-fold interference for the 20-nM LacI, 100-nM CI data, with overlapping 95% confidence intervals. We have more confidence in the 2.5-fold interference

estimate because the four-loop species were better balanced in the 20-nM LacI, 100-nM CI condition, and the match between data and prediction was better.

#### **§ 4.2.3 Nested loops give loop assistance.**

Loop assistance was examined by TPM with the Lac-inside nested construct (Fig. 4.2 E). Again, data were collected in the presence of LacI alone, CI alone, or both proteins (Fig. 4.2 F). The fraction of the CI-looped state was substantially increased in the presence of LacI. Note that this state is a mixture of two forms, with the state of the internal Lac loop invisible to TPM.

Fitting to the loop interaction model gave a large  $\alpha$  value of  $\sim 13$  (9.5–18). Thus, in a nested arrangement, formation of the inside loop improved formation of the outside loop (Fig. 4.2 G). Our reporter approach prevented us measuring the fraction of looping for the internal loop.

#### **§ 4.2.4 Side-by-side loops do not interact**

We also made reporters with a 1,500-bp CI loop and a 300-bp LacI loop placed next to each other and separated by 300 bp. We found that the presence of CI did not affect LacI looping, and the presence of LacI did not affect CI looping, confirming our expectation that side-by-side loops do not interact.

#### § 4.2.5 Model to quantitate loop interactions

Our collaborator in Australia develop this loop interaction model to detect the assistance/ interference of the loops, which briefly described as below (176).

If we consider just one pair of operators, then the DNA can exist in either a looped or an unlooped state. This equilibrium is determined by the nature and concentration of the looping protein, the length and nature of the DNA between the operators, and the chemical environment. The propensity of loop formation relative to the unlooped ground state under these fixed conditions can be simply defined by a statistical weight  $W$ . A Monte Carlo fitting procedure was used to find values for  $W_{Lac}$ ,  $W_{CI}$ , and  $\alpha$  that minimized  $\Sigma((\text{observed}-\text{expected})^2/\text{expected})$ . Varied F values were used in repeated fittings to propagate the uncertainty in the F estimates. The fraction looped is a function of this weight  $F = W/(1 + W)$ . Each LacI loop and CI loop has its own weight,  $W_{Lac}$  or  $W_{CI}$ , that determines  $F_{Lac}$  or  $F_{CI}$  in the absence of the other protein. In the case where there are two pairs of operators, there are four loop states: all sites unlooped, only LacI sites looped, only CI sites looped, or both pairs of sites looped. If the loops form independently of each other, i.e., they do not interact, then the statistical weight of the double-looped state is just the product of the individual loop weights  $W_{Lac} \cdot W_{CI}$ . However, if the loops do interact, a loop interaction factor  $\alpha$  can be used to quantitate the direction and strength of the loop interaction, with the weight of the double-looped species represented by  $\alpha \cdot W_{Lac} \cdot W_{CI}$ . Thus, when  $\alpha < 1$ , there is loop interference, the double-looped species forms less frequently than expected; when  $\alpha > 1$ , there is loop assistance, the double looped species forms more frequently than expected.

### § 4.3 Discussion

**Mechanism of loop interference.** My TPM measurements provided data complementary to that obtained *in vivo* by our collaborators to support the loop domain model and show that insulation is not restricted to complex regulatory elements in metazoan genomes, but can occur by loop interference between relatively simple DNA-looping protein-binding sites. Further experiments will be needed to test whether the insulation effects we see for 1.2- and 1.8-kb loops extend to much longer loops.

The asymmetrical alternating construct gave ~four-fold stronger *in vivo* than the interference seen for the same construct *in vitro*, with non-overlapping 95% confidence intervals for the *in vitro* and *in vivo* estimates of  $\alpha$ ; this implies an important role of some *in vivo* factor that affects DNA structure. Our favored explanation is DNA supercoiling, which would be present in our *in vivo* assays but absent in TPM. Brownian dynamics simulations show that DNA supercoiling compacts DNA such that the “search volume” for any two sites on the DNA to find each other is considerably reduced, perhaps by 10-fold to 100-fold (226, 227). Enhancement of protein-mediated DNA looping by supercoiling has been shown *in vitro* (227-229) and also stimulates recombination between distant sites *in vivo* (230). Looping enhancement by supercoiling is also consistent with our measurements of a 5- to 10-fold increase in the efficiency of long-range LacI looping of DNA *in vivo* compared with relaxed DNA *in vitro* (4). Much of the enhancement of search volume by DNA supercoiling is likely to be lost

when the sites are in separate topological domains, such as formed by a protein-mediated loop (231). Indeed, LacI loop inhibition of NtrC–promoter contact was dependent on DNA supercoiling of the plasmid template (223). Further experiments are needed to confirm the involvement of DNA supercoiling in the loop interference we measured in vivo. Of particular interest is whether the prevalent DNA supercoiling in eukaryotic genomes (232) plays a role in the efficiency and specificity of enhancer–promoter contact, or whether more complex mechanisms such as nucleosome–nucleosome interactions are also involved (233, 234).

We did not expect to see any loop interference in vitro with relaxed DNA, so the ~2.5-fold loop interference effect is intriguing. Whether this supercoiling-independent effect is due to specifics of the TPM setup (e.g., the inability of the bead to pass through a loop) or contributes to loop interference in vivo is not clear. It has been proposed that entropic effects can drive DNA circles apart when they are in a confined volume (235).

**Mechanism of loop assistance.** The loop assistance in vitro,  $\alpha = 13$ , was ~two-fold greater than we expected on the basis of distance shortening. In previous TPM experiments with LacI, we found that  $J_{loop}$  decreased with distance to the power 1.5 (4). Because the internal loop brings the CI sites 3.3-fold closer together (2,000/600 bp), we expected that the distance-shortening effect alone would give  $\alpha = 3.3^{1.5} = 6$ . The higher assistance seen may be due to specific angles imposed on the two 300-bp DNA “arms”

as they exit the internal LacI tetramer (236), which, combined with the relative stiffness of DNA in vitro (persistence length  $\sim 150$  bp) (71), may tend to juxtapose the CI sites.

**Achieving enhancer-promoter specificity by loop interactions.** Our analysis suggests that large changes in enhancer–promoter specificity could be caused by a single DNA loop that combines loop assistance and loop interference. We imagined an enhancer that is able to interact with either one of two promoters, with this contact regulated by a controlling loop that simultaneously assists the enhancer to loop to one promoter and interferes with looping to the other. Interestingly, the strengths of the enhancer–promoter loops are not important in determining the specificity change; the critical parameters are the  $\alpha$  values for the assisting and interfering effects of the controlling loop on the enhancer–promoter loops, ( $\alpha_{1c}$  and  $\alpha_{2c}$ ), as well as the strength of the controlling loop ( $W_c$ ). The maximal specificity change obtainable is given by the ratio of the assisting and interfering  $\alpha$  values ( $\alpha_{1c}/\alpha_{2c}$ ), which is approached as the controlling loop gets stronger. Thus, interactions between DNA loops provide a potentially powerful mechanism for regulating enhancer–promoter specificity.

Note: This chapter was originally published as part of the work "Quantitation of interactions between two DNA loops demonstrates loop domain insulation in *E.coli* cells" PNAS, 2014. DOI: 10.1073/pnas.1410764111



## Chapter 5 Protein-mediated looping of DNA under slight tension requires supercoiling

### § 5.1 Introduction

Protein-mediated DNA looping is a ubiquitous feature of gene regulation and other DNA transactions (237-240). The paradigmatic lac repressor (LacI)-mediated loop inhibits the expression of the lac operon in *E. coli* (241), and the looping probability is set by LacI tetramer concentration, binding site affinities, and the loop size (4, 242). Indeed, loops secured by a single LacI tetramer bridging two binding sites (operators) are suppressed when separate LacI tetramers occupy both operators, or when either the stiffness of short, or entropy of long, intervening DNA overwhelms the free energy of protein binding (71, 243). Thus, approximately 500 bp-long DNA segments most easily form LacI-mediated loops (71) at intermediate LacI concentrations (4, 101).

In vivo, bacterial DNA is extensively decorated and configured by nucleoid associated proteins (NAPs), including the abundant histone-like U (HU) protein (244), which compacts DNA into a flexible HU-DNA filament in 150 - 200 mM salt (245). HU potently changes the bacterial transcriptome (244), and binding correlates with negative supercoiling in stationary phase *E. coli* (246). 500 nM or more HU also drives LacI-mediated looping of roughly 140 bp DNA to maxima without altering the binding of LacI (247).

Supercoiling has been also shown to dynamically coordinate the transcription of bacterial (248, 249) as well as eukaryotic genes (250, 251), and the overall level of supercoiling is an indicator of cell health (252, 253). Supercoiling is widely and inhomogeneously distributed in the bacterial chromosome, being generated by the binding of nucleoid-associated proteins, such as HU (254, 255), and by various DNA transactions (246, 256). Negative supercoiling stabilizes regulatory DNA loops, for example, those that repress the lytic genes of bacteriophage  $\lambda$  (142, 257) or the *gal* or *lac* operons (138, 258). Such loops act as barriers to supercoiling diffusion and confine supercoiling density (142, 259) as part of biochemical interplay in which supercoiling promotes looping and loops constrain stabilizing superhelical density.

Small and large loops have different energy barriers to formation, but supercoiling and nucleoid-associated proteins may enhance looping in both cases. They can be distinguished by comparing the loop segment to the persistence length of DNA, about 150 bp (~50 nm). For small loops on the order of a persistence length or less, energy is required to kink or bend the short, stiff segment of DNA and juxtapose DNA segments for looping, as well as to twist the intervening DNA segment to orient both DNA binding sites toward the protein that joins them. Supercoiling can modify the twist of DNA to minimize the difference in rotational orientations of two binding sites about the helical axis (86). In this way a small, bidentate protein can simultaneously bind two sites on the same face of the helix to produce a loop. Several reports in the literature have shown that the probability of forming loops oscillates as a function of the rotational offset determined by the contour length between the binding sites (260-262).

This may be important for loops between the O1 and O3 operators of the *lac* operon in *E. coli*, which are separated by a torsionally rigid 92 bp segment, and may be relevant to observations of the supercoiling-driven modulation of *in vivo* repression of a reporter gene by protein-mediated DNA loops 70-85 bp-long (263). As shown previously, supercoiling-induced and HU-stabilized kinks likely bent the segment and optimized the orientation of the binding sites enhanced the formation of the 113 bp loop mediated by the *gal* repressor (258). Thus, there is clear evidence that HU and supercoiling may enhance the formation of short loops, whether or not the mechanism involves plectonemes.

For long, flexible loops, the foremost barrier is the separation between the binding sites (264). This is especially critical in DNA under tension which extends the molecule and favors twist over writhe, which opposes juxtaposition of the protein binding sites for looping. Proteins like RNA polymerase intermittently translocate DNA, creating as much as 20 pN of tension before stalling (265). This intermittent activity will modulate the twist/writhe equilibrium and thereby the opportunity for proteins like the *lac* repressor to secure loops. As will be shown below, supercoiling-induced plectonemes draw binding sites together to promote looping. The O2 and O1 operators of the *lac* operon in *E. coli* are separated by a 401 bp. The intrinsic flexibility of such a long span relieves flexural and torsional strain and randomizes the orientations of juxtaposed binding sites. Nevertheless, *in vitro* experiments from thirty years ago on similarly long segments showed that supercoiling enhanced looping (266). Those experiments which did not include protein factors, like HU, that might modify looping *in*

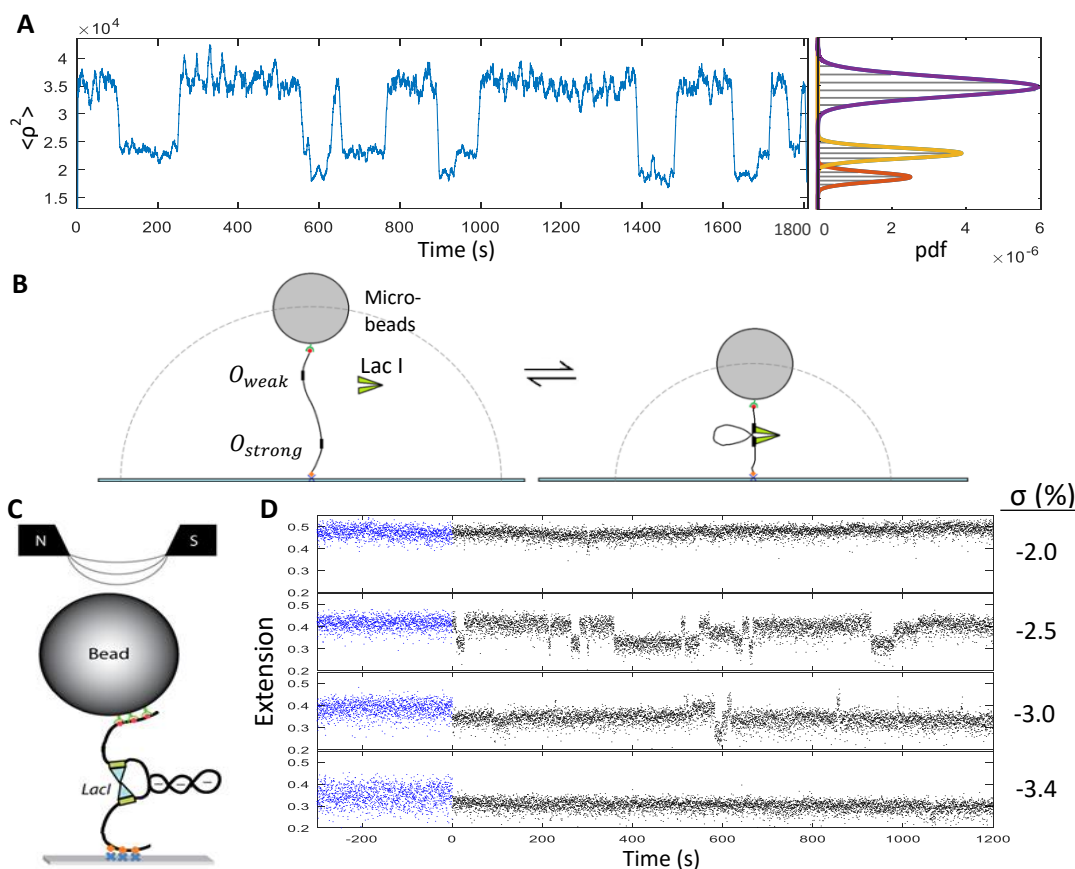
*vivo*, did not lead to a mechanical view of how supercoiling enhances the looping of long segments and whether accessory proteins can produce similar effects.

More recently, *in vivo* levels of LacI-mediated looping were found to be higher than those measured *in vitro* for a range of large loop sizes in identical DNA templates (4). To determine whether NAPs and/or supercoiling were likely to have caused the differences, the probabilities of LacI-mediated DNA loops in single DNA molecules were measured as a function of either the superhelicity of the tether, or the concentration of HU for DNA molecules under little or no tension. Both parameters progressively decreased the fractional extension of DNA under 0.25 pN of tension and enhanced looping, but only supercoiling drove the formation of stable loops in DNA under higher tension.

## § 5.2 Results

### § 5.2.1 HU decreased tether lengths and promoted LacI-mediated looping

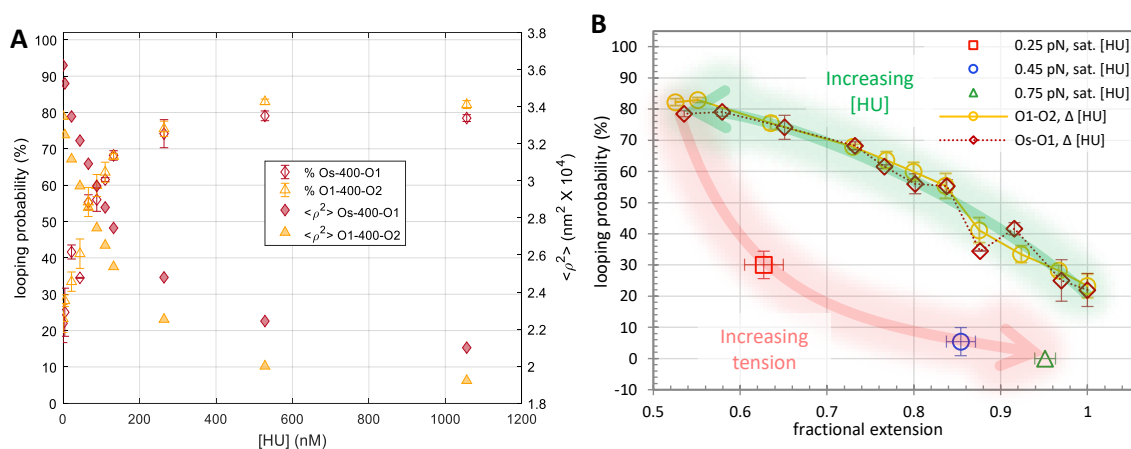
After establishing a concentration of LacI to set a baseline level of looping, increasing amounts of the HU protein were added to determine the effect on looping. In previously published magnetic tweezer experiments, the salt concentration affected the binding of HU to DNA (245). At monovalent salt concentrations in the range between 150 mM and 200 mM, increasing the HU concentration produced compaction and increased the flexibility of extremely long, 48 kbp, DNA tethers (245). At least 500 nM, HU also contracted 445 bp DNA tethers by 7% and drove LacI-mediated looping of short,



**Fig. 5.1** Experimental schematics and representative data. (A) A representative recording of mean squared excursion versus time shows intermittent changes in the average excursion of the bead as LacI-induced looping changes the overall tether length in a 909 bp Os-400-O1 DNA tether. At right, a histogram of the excursion values for the recordings exhibits three peaks which represent the two looped states and the unlooped state fitted with Gaussian curves, shown in red, yellow, and purple respectively. Similar data were recorded for 831-bp-long O1-400-O2 tethers (not shown). (B) A schematic representation of tether length changes due to LacI-induced DNA looping during a TPM experiment (not to scale). (C) A schematic representation of writhe created using a magnetic tweezer and trapped within a LacI-mediated loop (not to scale). Red/green dots indicate biotin/streptavidin linkages of DNA to beads. Orange diamonds/blue crosses indicate digoxigenin/antidigoxigenin linkages of DNA to glass. O1-400-O2 DNA looped by LacI is shown to trap three negative (-) supercoils. The north and south poles of magnets above the microscope stage create a magnetic field to attract and rotate the bead to allow stretching and twisting of the DNA. (D) Recordings of the extension of a 2115 bp-long O1-400-O2 DNA tether under 0.45 pN of tension at different levels of negative supercoiling in the absence (blue) and in the presence (black) of LacI. Extended (unlooped) states dominate at low negative linking number values but become intermittent and finally disappear altogether as negative supercoiling increases and favors the looped state.

roughly 140 bp, DNA segments to maximum levels without altering the binding of LacI (247). To test whether or not HU similarly affects looping of DNA segments longer than a persistence length, the looping probability was measured as HU protein was titrated from 0 to approximately 1  $\mu\text{M}$ , a high concentration yet still below maximum estimates for exponentially proliferating *E. coli* ( $\sim 50 \mu\text{M}$ ) (17). For Os-400-O1 or O1-400-O2 tethers in the presence of LacI, excursions of beads,  $\langle \rho^2 \rangle$ , tethered by both the looped and unlooped DNA decreased as the HU concentration increased (Fig. 5.2 A, Fig. S4 in (267)). To establish the associated looping probabilities, combinations of three Gaussian curves were fit to histograms of the excursion values aggregated from recordings in identical conditions (Fig. S4 in (267)). As [HU] increased, the DNA tethers contracted to reduce excursions of the attached beads. The fractional extension of the unlooped tether contracted more than the looped tethers (Fig. S5 in (267)). Since HU should bind randomly along the tether, this is most likely due to the greater contraction of centrally located, HU-induced bends compared to those in the flanking ends. This effect has been used previously in circular permutation assays to discover the location and extent of protein-induced bends in DNA segments (268). The looping probability was calculated as the area of the Gaussians for the looped states divided by the area of the combination of three Gaussians representing all the states. As HU gradually contracted the DNA tether and decreased excursions, the looping probability increased from 25 to a plateau near 80% (Fig. 5.2 A). Apparent contour lengths of the DNA tethers were determined from measurements of the mean square excursions of beads (Table S3 in (267)), which in  $\lambda$  buffer are related to the contour length of the DNA tether by:  $\langle \rho^2 \rangle = 100.89$

$\times L+3445$  (146). Looping probabilities were plotted as a function of the fractional extension at a given HU concentration to that without HU,  $L_f = L_{[HU]}/L_0$ . Fig. 5.2 B shows that HU shifted the looped/unlooped equilibrium to 80% looped when the fractional extension of the DNA dropped by 50%.



**Fig. 5.2** HU promotes looping. (A) Increasing HU concentration progressively decreased the mean squared extension values of Os-400-O1 or O1-400-O2 tethers from 100 to 50% (right hand Y-axis) simultaneously driving the looping probability from 25 to 80% (left hand Y-axis). (B) Plotting the looping probability as a function of the HU-modified extension of the DNA reveals an inverse relationship for DNA without tension. Application of slight tension reversed the HU-induced contraction and lowered the looping probability. The lower looping probability of tethers under tension in a magnetic tweezer may reflect a lower probability of the tether to adopt conformations with writhe when the bead was torsionally constrained.

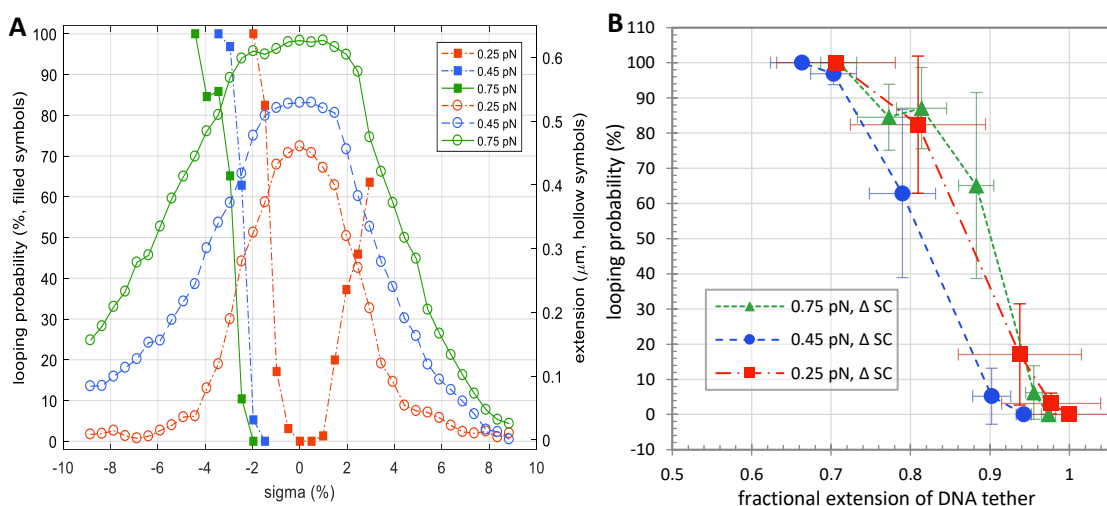
Applying tension opposed the HU-driven contraction restoring the extended state and reducing the looping probability. At 1  $\mu$ M HU concentration, which produced 80% looping in DNA without tension, 0.25, 0.45, or 0.75 pN of tension reduced the probability of looping to 30, 5, or 0% (Fig. 5.2 B). Although the contraction that was induced by HU (Fig. 5.2 B, green arrow) was essentially reversed by tension (Fig. 5.2 B, red arrow), for a given extension, the percentage of looping depended on the concentration of HU and the tension. Thus, the induction of looping by HU can be

reversed with tension and appears to be related to the contraction of DNA. A reduction of the overall extension of the DNA does not indicate that the DNA is shorter, because the contour length is unchanged. It indicates random coiling. We have compared the extension of rotationally unconstrained DNA to the looping probability, because coiling creates juxtaposed binding sites that can be linked by the lac repressor to form loops. The HU protein stabilizes large bends in DNA and may induce coiling in rotationally free DNA tethers like those in the TPM experiments described in the manuscript. However, even slight tension opposes coiling and reverses the enhancement of looping by HU. The looping probability drops sharply with slight tension, so that a DNA tether that is partially saturated with HU exhibits a much higher probability of looping without tension than a similarly extended, HU-saturated DNA tether under 0.25 pN of tension. The difference is that HU binding induces kinks in a DNA tether under no tension that reduce the overall extension by 45% and promote random coiling. Slight tension extends the kinked tether which reduces the frequency of juxtaposition of binding sites required for lac repressor to secure a loop.

### **§ 5.2.2 Supercoiling decreased tether lengths and induced LacI-mediated looping**

Supercoiling can also contract DNA when the torsional strain exceeds the buckling threshold and the molecule forms plectonemes. Supercoiling was recently shown to enhance  $\lambda$  repressor (CI)-mediated looping in both linear and circular DNA (142, 257). It influences gene expression (269, 270) and is a ubiquitous feature of genomes. To quantitatively characterize the impact of supercoiling on LacI-mediated looping, the





**Fig. 5.3** Supercoiling promotes looping. (A) No looping was observed in tethers under slight tension in a magnetic tweezer. However, over- or underwinding O1-400-O2 DNA tethers under 0.25 pN of tension produced plectonemes that reduced extension (open red circles) and increased the looping probability (filled red squares). Higher tensions (green and blue data) extended the tether further and required more negative supercoiling to produce comparable looping. For each tension, supercoiling decreased the extension and drove the looping probability from 0 to 100%! (B) Although achieving equivalent looping probabilities required more negative supercoiling in tethers under higher tension, the looping probabilities as a function of fractional extension were similar for the three different levels of tension tested. When the supercoil-induced plectoneme contracted the tether to 90-95% of the extended length, a loop intermittently formed and became the dominant conformation as further unwinding reduced the extension to 70%.

probability of the looped state was characterized while gently stretching and twisting 2115 bp DNA tethers using a magnetic tweezer (see schematic in Fig. 5.1 C).

Experiments on tethers with a centrally located O1-400-O2 loop sequence (Fig. S1 in (267)) were carried out at 0.25, 0.45, and 0.75 pN, in the presence of 1 nM Lacl. No looping occurred without Lacl (Fig. 5.1 D, blue recordings) or with Lacl acting on torsionally relaxed tethers under tension (Fig. 5.3 A), because such tension precludes juxtaposition of operators separated by 400 bp, roughly 130 nm, along the tether (271, 272). However, superhelical densities between -1.9 and -3.4 % induced switching

between looped and unlooped states (Fig. 5.1 D, black traces). Under low tension, DNA tethers buckle randomly to form plectonemes that rapidly slither or hop to different locations along the length (273). LacI has the opportunity to secure a loop when the binding sites slither into juxtaposition. In this conformation plectonemic supercoils are constrained inside the loop. Dwell times in the looped and unlooped states were calculated as described above to determine the looping probabilities shown in Fig. 5.3 A. Loops persisted several tens of seconds as reported previously (89, 164) and significantly longer than those of the extended and plectonemic states observed as a twisted DNA tether begins to buckle (274).

Supercoil-driven, LacI-mediated looping exhibited several noteworthy features. First, at each tension, looping probability increased from 0 to 100 % as superhelical density became more negative (Fig. 5.3 A). The superhelical density of the tethers is  $\sigma = (Lk - Lk_0)/Lk_0 = n/Lk_0$ , where  $Lk$  and  $Lk_0$  are the linking numbers of twisted and relaxed DNA respectively, and  $n$  is the number of mechanically introduced twists.  $Lk_0$  is given by the total length in base pairs divided by 10.4 base pairs per turn. Additional superhelicity lowered the free energy,  $\Delta G$ , for looping at each tension (Fig. S6 in (267)) and negative superhelicity was more effective than positive. Second, the free energy change for LacI-induced looping as a function of supercoiling was compared with data for the formation of a 393 bp DNA loop by the  $\lambda$  repressor protein (CI) (142), which secures looping via cooperative interactions amongst multiple dimers. Negative supercoiling decreased the free energy for LacI- more than for CI-mediated looping (Fig. S6 in (267)). These free energy measurements support theoretical investigations indicating that the free energy

change for 400 bp LacI-mediated looping *in vivo* is -10 kcal/mol (-4.4 *kT*) or less (275). Third, overlay of the looping probability versus twist measurements with the DNA extension versus twist curves (Fig. 5.3 A), revealed that the onset of loop formation induced by negative supercoiling coincided with the onset of plectoneme formation. Indeed, in DNA tethers under tension, plectoneme formation was essential for looping, and more extreme (negative) supercoiling densities were required at higher tensions to achieve similar looping probabilities.

The coincidence of the initial increase in looping probability and plectoneme formation suggested that contraction juxtaposes binding sites for LacI looping. Graphically comparing the contraction of a DNA tether and the associated LacI-mediated looping probability shows that the percentage of looping resembles the percentage of contractions below a threshold that is one loop length shorter than the full-length, unlooped tether, such as that produced when LacI secures a loop. After introducing -2% negative supercoiling during an extension-vs.-twist experiment at 0.45 pN, momentary tether lengths (Fig. S7 grey points in (267)) occasionally dip below a level that is one loop segment shorter (Fig. S7 red dashed line in (267)) and the looping probability begins to rise. With additional negative supercoiling, more measurements fall below that level and the looping probability increases, finally reaching 100%. This is also reflected in extension versus time records which display increases in the frequency and duration of intermittent looping as a DNA tether is negatively supercoiled (Fig. 5.1 D), until the looped state becomes perpetual.

If contraction of the DNA were indeed the mechanism by which supercoiling induced looping, then it would be expected to dictate the looping probability independently of the tension. To quantitatively examine this relationship, the fractional extension was defined as the tether extension at a given supercoiling density divided by that of torsionally relaxed DNA,  $L_f = L_\sigma/L_0$ . In Figure 5.3 B, plots of the looping probability associated with different levels of negative supercoiling at three levels of tensions are shown. The looping probability began to rise after supercoiling induced contraction to 0.9-0.95 of the torsionally relaxed extension. Thereafter, the looping probability rapidly increased, and tethers contracted to 0.7 of the torsionally relaxed extension were 100% looped for all three levels of tension tested.

## § 5.3 Discussion and conclusion

### § 5.3.1 Contraction, especially with writhe, enhances loop formation

The conformation of DNA stretched and twisted to the buckling point under different tensions rapidly fluctuates between plectonemic and extended conformations (274). Comparison of the onset of loop formation in Os-400-O1 DNA constructs under different tensions (Fig. 5.3 B), shows that LacI-mediated loops begin to form just beyond the buckling threshold at which plectonemes form. The contraction of DNA caused by the initial plectonemic gyre has been studied as a function of tension using an optical tweezer with high temporal resolution. This gyre of DNA is about 130 nm-long under 0.45 pN of tension (274). The LacI-induced loop is similarly sized, and this coincidence

indicates that lac repressor captures and stabilizes looped configurations produced by writhe. Indeed, Figure S7 in (267) shows that the looping probability correlates with the probability that a tether exhibits an instantaneous extension with at least one loop length of slack. This is the only way in which LacI, which is not a DNA translocase, can establish DNA loops.

HU also seems to promote looping by reducing the fractional extension. It is noteworthy that saturating HU promoted looping more effectively in DNA tethers without tension. Although equivalent levels of contraction were produced in tethers exposed to various concentrations of HU without tension and other tethers exposed to saturating levels of HU and varying levels of tension, the looping probabilities observed for tethers under tension were much lower (Fig. 5.2 B). This is likely due to the fact that tension opposes juxtaposition of the protein binding sites and looping even in a DNA tether with a saturating level of HU-induced kinks.

Tension is likely to be physiologically relevant for at least three reasons: First, models for the hierarchical packaging of DNA into chromatin include loops ranging from 15 kb segments in *Caulobacter crescentus* (276) to hundreds of kbps in mammals (277), and supercoiling appears to be critical for these models. The Meiners laboratory has shown that sub-piconewton tensions can prevent the formation of 305 bp loops (264), and even femtonewton-scale tensions would significantly interfere with loop formation in much larger segments and interfere with hierarchical packaging. Indeed, novobiocin treatment of *Caulobacter crescentus* cells, which inhibits DNA gyrase and lessens negative supercoiling, altered interactions between segments separated by 20 to 800

kbp (276). Our finding that supercoiling rescues loop formation in DNA under tension, suggests that DNA *in vivo* is under tension and that supercoiling is required to condense chromatin through plectoneme formation and looping. Second, *in vitro* experiments have shown that the translocation velocity of RNA polymerase is unaffected by up to 20 pN of opposing tension on the DNA (265), and that the nucleic acid translocation enzyme that packages DNA into the p29 bacteriophage exerts as much as 50 pN of tension on the DNA (278). It is unlikely that these enzymes would have evolved such capabilities unless they operate against mechanical tension. Third, the Belmont and Wang labs have recently demonstrated that the tension exerted by forces applied to microscopic beads attached to the surface of a cell can be transmitted to the nucleus where it draws segments of chromatin apart and simultaneously enhances transcription from genes in the intervening DNA (279). This demonstration of enhanced transcription by stretching chromatin in a live cell suggests that cells in mechanically distorted tissues may respond altering their transcription profile to the stress.

### **§ 5.3.2 Supercoiling lowers the free energy of DNA looping in a protein-dependent manner**

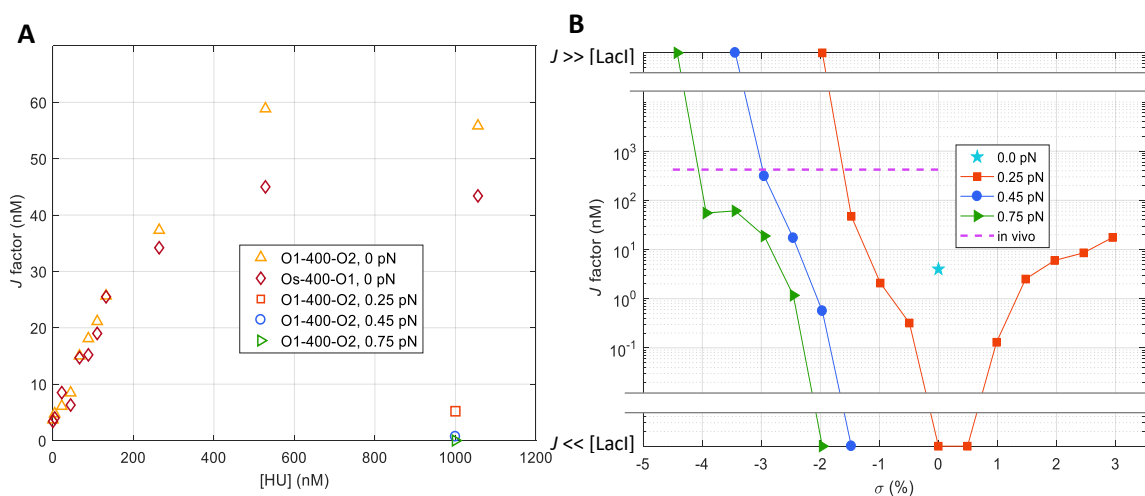
Once the loop closes the DNA is partitioned into different topological domains. Previous work from our lab demonstrates that the lambda CI loop serves as a barrier for supercoiling diffusion (142), and work in this paper shows that LacI behaves similarly. Key features of this interplay are that supercoiling can become sequestered within a

loop, but also that supercoiling modulates the looping probability (138). Figure S6 in (267) plots the energy change between looped and unlooped states under different supercoiling densities. Both positive and negative supercoiling lowered the free energy of LacI-mediated loop formation, but negative supercoiling was more effective. In addition, the response of different loop-mediating proteins, LacI and  $\lambda$  CI negative was different. The free energy dropped more rapidly for LacI-mediated looping as negative supercoiling increased. This may reflect different binding interfaces with DNA, or their structures which may have different tolerances for changes in binding site phasing.

### § 5.3.3 Comparing the effects of HU and supercoiling on DNA looping via $J_{loop}$ factor

The  $J_{loop}$  factor is a succinct parameter for comparing the effects of HU and supercoiling on DNA looping. It represents the effective concentration driving the binding of a protein at one site along a DNA molecule while simultaneously bound to another site along the same DNA tether. This factor accounts for the energy changes involved in forming a DNA loop independently of the affinity of protein linkages. The  $J$  factor can be calculated (101) from the ratio of the unlooped and looped state probabilities,  $\frac{p_{loop}}{p_{unloop}}$ , the dissociation constant from the secondary operator,  $K_d$ , and the concentration of *lac* repressor,  $[R]$ , using the equation  $J = 2 \frac{p_{loop}}{p_{unloop}} ([LacI] + K_d)$  (101). Increasing the HU concentration surrounding torsionally relaxed DNA tethers drove  $J$  factors to plateau at 11-15 times the initial values, and slight tension reversed the gains (Fig. 5.4 A). In contrast, supercoiling, especially negatively, increased  $J$  factors by several orders of

magnitude from far below the free LacI concentration (1 nM), for which there was no looping under tension. Negative supercoiling drove the concentration of lac repressor binding sites to extremely high levels corresponding to  $J$  factors far greater (effectively infinite) than those reported for a 400 bp loop *in vivo* (Fig. 5.4 B; purple dashed line). This produced perpetual loops which were not observed with HU protein alone even in tethers under no tension.



**Fig. 5.4**  $J_{loop}$  factors decrease as DNA tethers shorten especially when the DNA becomes negative supercoiled. (A) Increasing the HU concentration surrounding Os-400-O1 (diamonds) or O1-400-O2 (upright triangles) tethers drives the  $J_{loop}$  factors to saturation at multiples of 11-15. Slight tension lowered the HU-improved  $J$  factors back toward their original values (left and right-pointing triangles). (B) The  $J_{loop}$  factor measured for O1-400-O2 in TPM experiments (with negligible tension) was 4 nM (cyan star). Slight tension precluded looping ( $J_{loop} \ll [\text{LacI}] = 1$  nM) until the DNA became sufficiently supercoiled (Fig. 5.1B). Indeed, supercoiling, especially negative, increased  $J$  factors by several orders of magnitude to levels ( $J_{loop} \gg [\text{LacI}] = 1$  nM) beyond that reported for a 400 bp loop *in vivo* (purple dashed line).

## § 5.4 Conclusion

Although HU binding promoted the formation of 400 bp, LacI-mediated loops in DNA molecules without tension, supercoiling was essential to drive significant looping of



molecules under slight tension. While direct measurements of tension in genomic DNA have not been reported, the mechanical fact that DNA translocases can operate against as much as tens of pN of tension and the tension sensitivity of looping described above suggest that supercoiling is critical for looping *in vivo*. Simultaneously, a cellular perspective of the prominent role that protein-mediated loops play in chromatin condensation and transcriptional regulation, and the ubiquitous genetic requirement for topoisomerases that regulate supercoiling, reveal its importance of supercoiling in DNA biochemistry. Indeed, supercoiling directly reflects the current state of cellular metabolism and is a fundamental genomic parameter. Future studies are likely to show that, similarly to proteins that by virtue of recognition sites act at very specific sites, supercoiling can be constrained within loops and plectonemes to control dispersion and locally regulate DNA transactions.

Note: This chapter was originally published on Nucleic Acids Research, 2018.

DOI:10.1093/nar/gky021

# Chapter 6 Protein-mediated loops in supercoiled DNA create large topological domains

## § 6.1 Introduction

Supercoiling is an inherent and dynamic feature of DNA that sensitively regulates genome-based reactions (263, 280-282), altering base-pairing or the form of the helix (283, 284) to modify protein binding. It affects the thermodynamic stability of regulatory, protein-mediated loops (142, 285), and in addition may become trapped within repressor mediated loops (138, 142, 259). The interplay between supercoiling and looping appears to be fundamental to eukaryotic gene regulation, since topologically associating domains (TADs) with loops catalyzed by CTCF (277) are circumscribed by topoisomerase (TOP2B), an enzyme that modifies supercoiling, in mouse liver cells (286). Such positioning may effectively isolate supercoiling in adjacent TADs. In addition, psoralen incorporation has shown that transcriptionally active TADs are negatively supercoiled (251), and models of transcription-induced supercoiling operating in 600 kbp regions generates supercoiled domains with interaction patterns similar to those observed experimentally in *S. pombe* (287). Supercoiling also affects the hierarchical organization and function of chromatin in prokaryotes like *Streptococcus pneumoniae* which has genes organized in topology-sensitive clusters (288).

This evidence suggests that supercoiling plays a role in establishing domains in chromatin. It is straightforward to imagine the conformational differences between a linkage between two sites in a gently curved DNA molecule to form a simple loop versus a linkage between juxtaposed sites in a solenoidally or plectonemically coiled molecule that traps supercoiling and may even produce knots. The entanglement that accompanies supercoiling creates topology that, when stabilized by a protein-mediated junction, may have regulatory consequences due to the compaction of the DNA, the difference in supercoiling within and outside the loop, and the partitioning that isolates segments inside a loop from those outside the loop (147).

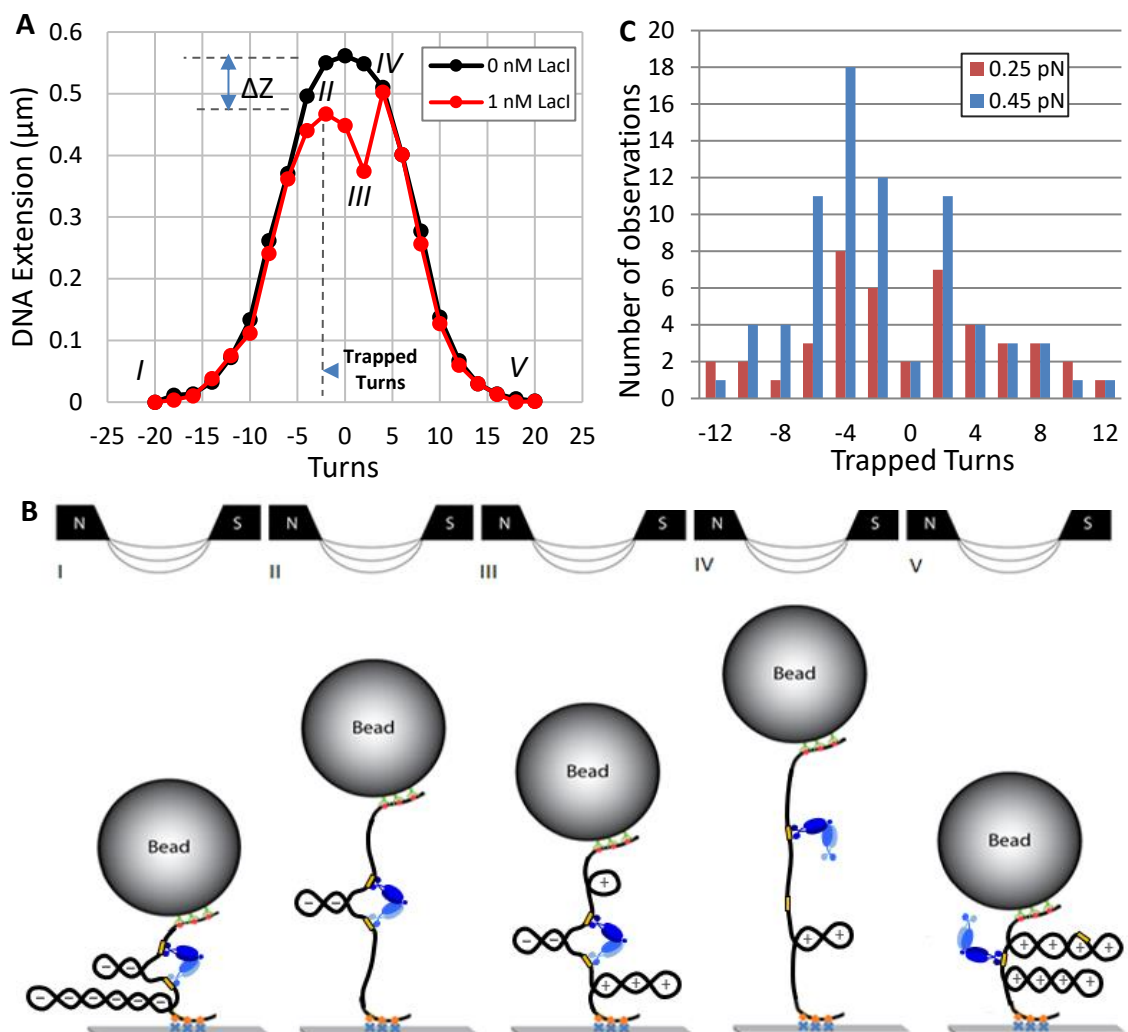
Although the supercoiling and topology of DNA segments are dynamic and so complex *in vivo* that measurements would be difficult, single molecule manipulation *in vitro* allows arbitrary twisting of DNA while monitoring the overall extension of the molecule to detect loop formation. Therefore, to investigate the interplay between supercoiling and looping, superparamagnetic beads were tethered to the surface of a glass microchamber by DNA molecules containing two *lac* repressor (LacI) binding sites separated by 400 bases, and LacI was added to the solution. LacI protein is a paradigmatic DNA looping protein (289) that has two DNA-binding domains connected by a hinge, and can bind simultaneously two, non-adjacent binding sites, producing a loop in the intervening DNA. We used magnetic tweezers (MT) (290) to impart controlled amounts of supercoiling to the DNA tethers, and monitored intermittent shortening of the molecules due to LacI-mediated looping and entanglement, as well as the amount of writhe (291) trapped by loops. LacI-mediated loops often sequestered

supercoiling and topologically constrained domains larger than the loop. Thus, protein-mediated loops in supercoiled DNA affect not only the circumscribed DNA but also topologically extend to flanking sequences. This newly discovered way of creating extended topological domains, may enable the coordinated regulation of serially arranged elements of the genome.

## **§ 6.2 Results**

### **§ 6.2.1 Loop-securing LacI resists torsional stress in dynamically supercoiled DNA**

To investigate supercoiled loops, para-magnetic microspheres were tethered to a cover glass by single 2121 bp DNA molecules that were gently stretched and twisted using magnetic tweezers. Twisting in the absence of LacI produced symmetric extension-vs-turns curves in which progressively larger plectonemes reduced the extension, as the amount of mechanically introduced twist increased (Fig. 6.1 A black, S4 black in (1)). In the presence of LacI, extension-vs-turns curves often exhibited shifted and/or reduced peaks (Figures 6.1 A red and S4 red or blue in (1)) with respect to controls without LacI. In the representative example depicted in Figure 6.1 A (red), initially the DNA was extensively negatively supercoiled (-20 turns) with an extended plectoneme (Figure 6.1 B, state I). Successive cartoons in Figure 6.1 B represent conformations that the molecule likely visited during winding from -20 to +20 turns (Figure 6.1 A, red curve). Winding progressively relaxed plectonemes in the flanking segments, but plectonemes within the



**Fig. 6.1** Lacl-mediated loops trap supercoiling and resist torsion. The formation of plectonemes in an O1-400-O2 construct under 0.25 pN of tension without Lacl produced a typical extension-vs-turns curve (black). When Lacl was added to the buffer, the maxima of subsequent extension-vs-turns curves were often reduced by  $\Delta Z$  due to loop formation by Lacl and shifted according to the number of trapped supercoils. (A) In this example, after unwinding by -20 turns, an extended plectoneme reduced the extension of the tether (state I) with or without Lacl. The relative shift of the maxima of the extension-vs-turns curves acquired with (red) and without (black) Lacl while winding from -20 to -2 turns indicates that two negative supercoils were trapped within the loop (state II). After further winding to +2 turns to reach state III, the loop ruptured, and the extension became identical to that without Lacl (state IV). Further winding to +20 produced extensions equal to those observed without Lacl. (B) Schematic representations of supercoiled configurations compatible with the red extension-vs-turns curve in A. At -20 turns a large plectoneme is formed and Lacl can secure a loop between juxtaposed operators (state I). Winding the tether from -20 to -2 turns, relaxed flanking plectonemic gyres but not those within

the LacI-mediated loop, which shifted the maximum of the extension-vs-turns curve from 0 to -2 turns (state II). Further winding and unwinding cycles, negative supercoils are more often trapped in loops formed at 0.45 pN. Note also that few loops formed without trapping any supercoiling. to +2 turns produced (+) plectonemes in the flanking DNA, while the plectoneme within the loop remained negatively (-) supercoiled (state III). In this state, the LacI junction sustains maximal torsional stress between oppositely supercoiled DNA segments. When the loop spontaneously ruptured, negative supercoiling within immediately relaxed an equal amount of positive supercoiling in the flanking DNA giving a more extended tether with a smaller plectoneme (state IV). Thereafter, further winding extended the plectoneme and the extension-vs-turns curve superimposed with that of an unlooped tether as a large plectoneme with positive supercoils formed (state V). LacI might secure loops between juxtaposed operators in this configuration that could be detected by a shift in the extension-vs-turns maximum during unwinding. (C) Repeatedly winding and unwinding DNA tethers in the presence of LacI and noting the shifts in the extension-vs-turns curves produced distributions of the number of supercoils trapped by the LacI-mediated loops in tethers under 0.25 (red) or 0.45 (blue) pN of tension. While the distribution of trapped supercoils is symmetric about zero for tethers under 0.25 pN of tension for equal numbers of winding and unwinding cycles, negative supercoils are more often trapped in loops formed at 0.45 pN. Note also that few loops formed without trapping any supercoiling.

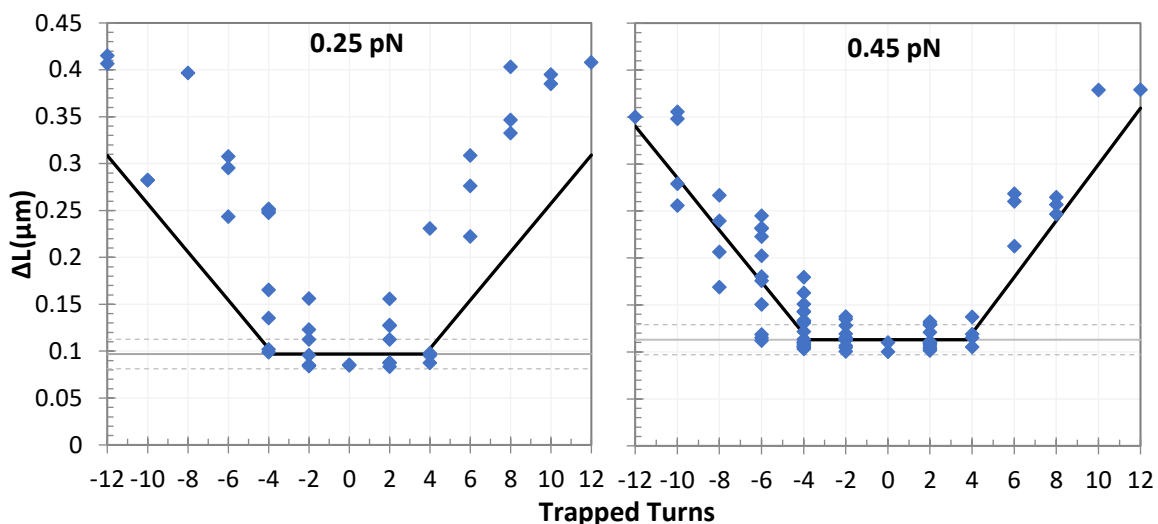
LacI-mediated loop were protected, so that the extension-vs-turns curve displayed a reduced maximum at -2 turns (state II). Further winding produced positive plectonemes in the flanking DNA (+), while the plectoneme within the loop remained negatively (-) supercoiled (state III). This condition, with plectonemes of opposite handedness inside and outside the loop, exerts the maximal possible torsional stress on the LacI tetramer securing the loop. When the loop ruptured, negative supercoiling within the loop immediately relaxed an equal amount of positive supercoiling in the flanking DNA giving a new torsional equilibrium (state IV). Thereafter, the curve was identical to that of an unlooped tether.

### § 6.2.2 LacI-mediated loops preferentially trap negative supercoiling at higher tension

Loops that trapped superhelicity were observed under two levels of tension (Figure 6.1 C). Under 0.25 pN of tension, loops trapped similar distributions of negative and positive supercoils and torsion-free loops rarely formed. Under 0.45 pN of tension, torsion-free loops were rare as well, but loops that trapped supercoiling readily formed and more frequently trapped negative than positive supercoils. At either tension, some of the LacI-mediated loops trapped as many as  $\pm 12$  turns. Several studies of DNA with moderately sized plectonemes have established the contour length per writhe at about 50 nm under approximately half a piconewton of tension (274, 292-295). However, twelve plectonemic gyres 50 nm in length do not fit into a 400 bp segment. Note also that at 0.45, but not 0.25 pN, LacI more often trapped negative supercoiling. Since unwinding DNA under 0.45 pN of tension produces a partial transition to L-DNA (283, 284), which has a left-handed helical repeat of 15 bp, it would be possible to absorb twelve negative supercoils as L-DNA in a 400 bp loop under 0.45 pN of tension. However, the phase change to L-DNA does not occur at 0.25 pN for which high numbers of trapped supercoils were also observed. In addition, positive supercoiling under either tension does not induce a helical phase change that might allow +12 turns to become trapped in the loop. All together these observations indicate that reductions of the extension to the maxima observed in the shifted peaks involve a more general mechanism of trapping supercoils.

### § 6.2.3 LacI-mediated loops dynamically trap large topological domains in supercoiled DNA

A naïve expectation is that the boundaries of topological domains induced by looping proteins should correspond to the contour length between the delimiting protein binding



**Fig. 6.2** Loops that trapped high levels of supercoiling greatly reduced the extension of tethers under tensions of 0.25 (left) and 0.45 (right) pN. Thin grey horizontal lines in both panels show  $\Delta Z$  values corresponding to the loop size (400 bp). The grey dashed lines represent  $\pm 2$  standard error of  $\Delta Z$  measurements for the unlooped tether (Table S1 in (1)). The black lines represent estimates of  $Z_{loop}$  values derived from Equation 6.1 using the average  $Z_{gyre}$  values reported in Table S1 in (1). They are minima, because the loop was assumed to have absorbed its' full capacity of trapped supercoiling before any additional supercoiling was attributed to flanking plectonemes entangled with the loop.

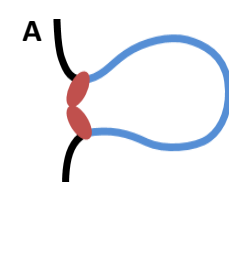
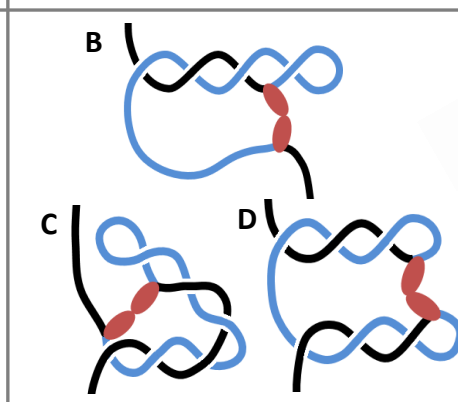
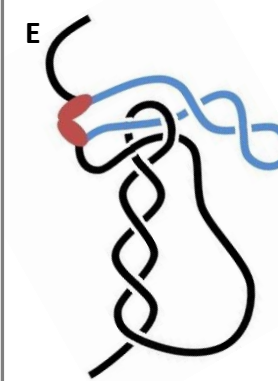
sites. In this study, such a topological domain would reduce the maxima of extension-vs-turns curves by a value,  $Z_{loop}$ , proportional to the length of the loop,  $L_{loop}$ . Figure 6.2 reports the reduction of the extension-vs-turns maxima for LacI-mediated loops relative to unlooped DNA tethers,  $\Delta Z$ , as a function of the number of turns trapped by the loop.

Expected  $Z_{loop}$  values were estimated using the proportionality constants described in figure S3 in (1). This produced  $\frac{400 \text{ bp}}{4124 \text{ bp}/\mu\text{m}} = 0.096 \mu\text{m}$  or  $\frac{400 \text{ bp}}{3540 \text{ bp}/\mu\text{m}} =$



0.113  $\mu\text{m}$  for 0.25 pN or 0.45 pN, respectively.  $Z_{loop}$  values were expected to lie along the grey, horizontal lines. 0.41  $\mu\text{m}$ , the largest  $\Delta Z$  observed at 0.25 pN, is 4.3-fold larger than  $Z_{loop}$ . Under 0.45 pN of tension,  $\Delta Z$  values as large as 0.38  $\mu\text{m}$ , 3.4 times  $Z_{loop}$  were observed. Thus, LacI-mediated loops in supercoiled DNA frequently established topological domains much larger than the loop length, and the size of the increase was inversely correlated with tension. In Figure S5 in (1), a cumulative histogram of  $\Delta Z$  values for DNA tethers shows that at 0.45 pN 50 percent of the topological domains were greater than  $Z_{loop}$  by as much 3.4-fold. At 0.25 pN 65 percent of  $\Delta Z$  values exceeded  $Z_{loop}$  reaching a maximum of 4.3. This difference may be due to tension which opposes coiling as well as the juxtaposition of coils that might become entangled when a loop forms. It is remarkable that these topological domains constrained 64 to 80 % of the tether while the loop segment was only 19% of the total length.

$\Delta Z$  values equal to  $Z_{loop}$  only occur when all trapped supercoiling,  $n_{trapped}$ , is entirely within the loop (Figure 6.2, grey horizontal line). To determine the amount of supercoiling trapped within and peripheral to the loop, one can estimate the maximum number of supercoils that might be contained in the looped segment,  $n_{loop}$ , by dividing  $Z_{loop}$  by the extension corresponding to a plectonemic gyre,  $Z_{gyre}$ . Using the  $Z_{gyre}$  estimates in Table S1 in (1) indicates that 3.8 gyres could fit within a loop segment. In Figure 6.2, it is noteworthy that values equal to  $Z_{loop}$  only occurred for  $|n_{trapped}|$  values below 3.8, unless there was sufficient tension (0.45) to induce the formation of some easily trapped left-handed helicity in unwound tethers. Other larger changes in  $\Delta Z$  must have included not

$\Delta Z$	Loop $\Delta Z = Z_{loop}$	Loop + plectonemic entanglement $2L_{loop} \geq \Delta L > L_{loop}$	Loop + knot $\Delta L > L_{loop}$
Conformation			

**Fig. 6.3** Different loop topologies may trap flanking DNA. Lacl might secure a loop without entangling any flanking DNA and shorten the tether by a value proportional to the loop size,  $Z_{loop}$  (A). Alternatively, a blue looped segment may entangle with flanking black segments to further reduce the extension. For example, Lacl may connect operators in a plectoneme and flanking DNA (B, C), in two different plectonemes (D), or in knotted DNA (E). Entanglements between the loop and flanking segments can reduce the extension by up to twice  $Z_{loop}$  (B-D). Knotting between the loop and flanking segments can reduce the extension by more than  $Z_{loop}$  (E).

only the loop segment, but also plectonemically entangled flanking sequences, that became topologically constrained. Since high values of writhe must be partitioned between the loop and flanking segments,  $n_{trapped} - n_{loop}$  represents the topologically constrained writhe in the flanking DNA. The length of this topologically constrained flanking DNA will be  $Z_{gyre} (n_{trapped} - n_{loop})$ . Thus, a minimum estimate for  $\Delta Z$  would include  $Z_{loop}$  plus flanking plectonemic entanglements:

$$\Delta Z = Z_{loop} + Z_{gyre} (n_{trapped} - n_{loop}) \quad [\text{Eqn.6.1}]$$

Estimation of the minimum  $\Delta Z$  was obtained by substituting 3.8 for  $n_{loop}$  in Eqn. 6.1 (Figure 6.2, black lines). In addition to a simple loop (Figure 6.3 A) that would produce  $\Delta Z$  values

equal to  $Z_{loop}$ , Figure 6.3 depicts hypothetical, topological configurations that would exhibit larger values. When  $2Z_{loop} \geq \Delta Z > Z_{loop}$ , flanking DNA must be entwined with the loop segment (Figures 6.3 B-D) or knotted (Figure 6.3 E). Any larger  $\Delta Z$  values, most likely involve Lacl-mediated knotting (Figure 6.3 E).

Since supercoil partitioning may change the entwining/knotting when the loop stochastically breaks and reforms, the structure and size of the loop-trapped domain may vary as loops break and re-form (Figures 6.2, 6.3, S4 in (1)). This suggests that the size of the topological domains may be dynamic. Indeed, records of the extension of supercoiled DNA with and without Lacl, clearly showed rapid, Lacl-mediated switching between multiple states of different length (Figure S6 in (1)).

#### § 6.2.4 Exclusion of artifacts

Non-specific interactions could in principle decrease the length of the DNA tether and can be of three different types: 1) bead to surface, 2) DNA to surface or bead, 3) DNA-bound protein to surface or bead. We have ruled out the possibility that any of these could be the cause of the DNA length shortening we observed in this study. First, non-specific sticking of a DNA-tethered bead to the surface of the flow chamber would produce tether length measurements equal to zero. Such tethers were discarded from the analysis. Second, several observations indicated that transient, non-specific interactions between DNA and either the surface of the flow-chamber or the bead did not occur in the analysed data: i) control measurements in the absence of Lacl (looping protein) showed only one

state, and the extension-vs-turns curves recorded in this condition had reproducible maxima; ii) decreases in the extension of the tethers were greater than or equal to the expected length of the loop segment,  $Z_{loop}$ , under the applied tensions. Random, nonspecific sticking would have also produced smaller decreases, which were not observed; iii) The clear correlation between the tether length decrease and number of turns trapped in the LacI-mediated loop would not result from non-specific interactions that caused temporary, sticking of random segments of the DNA to either the surface of the flow-chamber, or the bead. Third, no interaction has ever been reported between LacI and streptavidin, or anti-digoxigenin, and observations that allow us to exclude artifacts caused by DNA sticking to the bead or glass surfaces (ii and iii) allow us to exclude protein-mediated DNA sticking as well.

Finally, non-specific binding of LacI to DNA is negligible (if not non-existent) in the buffer condition and LacI concentration used, as seen by atomic force imaging (137, 162) and tethered particle microscopy assays where, in the absence of supercoiling, the extension of a DNA tether without operator sequences in the presence of LacI remained constant in time (89, 164, 173, 296).

### **§ 6.3 Discussion**

These data show that LacI can: i) separate domains of oppositely supercoiled DNA withstanding considerable torsional stress, ii) mediate loops in DNA under higher tension (0.45 pN), that likely contain L-DNA and preferentially trap negative supercoiling (Figure

6.1 C, 6.2), iii), topologically constrain segments larger than the distance between LacI operators (Figure 6.2) by entwining, or knotting, flanking and loop segments of DNA (Figure 6.3), and iv) catalyze variations of the size of these topological domains (Figure S6 in (1)). The latter agrees with the recent suggestion, from microarray assays and Monte Carlo simulations, that the sizes of topological domains in the chromosome change dynamically (118). Such dynamics might moderate extensive genomic restructuring or compaction due to the cooperative binding of architectural proteins, such as HU or HNS in prokaryotes, or histone octamers in eukaryotes.

Topological domains like those described here were observed also with  $\lambda$  CI repressor-mediated loops of different sizes (Figure S7 in (1)), indicating that constraining topological domains larger than the distance between binding sites in supercoiled genomes may be a general feature of protein-mediated DNA looping. These findings have intriguing implications for DNA/chromatin packaging, especially for coordinated gene expression. For example, it is straightforward to envision that a supercoiled loop might encompass nearby genes and influence the transcription from a nearby promoter, but coarse-grained models show that the interaction of two distant sites along contiguous DNA is most sensitive to supercoiling (297). Plectonemes in distant DNA segments topologically pinned by protein-mediated loops may explain this observation.

There might be significant advantages to topologically entangling more DNA than just the segment between the binding sites of a loop-mediating protein. First, a larger topological domain protects larger segments of DNA from damage following single-stranded nicks (Figure S8 in (1)). Second, larger topological domains may catalyze DNA

damage repair by keeping the two ends to be re-joined closer to each other, since they must slither to untangle from each other instead of being free to separate along multiple three-dimensional trajectories (297, 298). Such confinement would be particularly important when double-strand breaks must be repaired through non-homologous end joining. Third, larger domains will produce greater compaction of the chromosome. Fourth, these larger topological domains exhibit different configurational isomers of DNA, which may increase the opportunities for regulation. Fifth, larger topological entanglements facilitate bridging by proteins that maintain structural integrity of the chromosome, such as HNS in prokaryotes (298). Lastly, protein-mediated knotting involving a distant plectoneme radically alters topology and perhaps DNA transactions quite distant from the loop (Figure 6.3 E).

This work suggests that large, dynamic domains, which characterize genomes across biological kingdoms, may result from protein-mediated loops that topologically isolate segments of supercoiled DNA larger than the actual contour length between the protein binding sites. Although relatively small DNA loops, in DNA molecules only five times longer, were examined in this study, the possibility of the formation of extensive plectonemic structures between large loops and flanking DNA is predicted to significantly impact chromosome structure, dynamics and function, including coordinated gene expression.

Note: This chapter was originally published on *Nucleic Acids Research*, 2018. DOI: 10.1093/nar/gky153

# Chapter 7 RNA polymerase pauses at *lac* repressor obstacles

## § 7.1 Introduction

Proteins bound to DNA may act as roadblocks that hinder elongation by RNA polymerase during transcription. The mechanisms by which RNA polymerases might surpass such roadblocks are poorly understood. Previous single-molecule studies have focused on RNA polymerases disrupting nucleosomes (299-302). However, nucleosomes are only found in eukaryotes, interact with DNA non-specifically, and are substrates for post-translational modifications that regulate chromatin remodeling and the recruitment of accessory factors that regulate transcription (303). In contrast, many transcription factors from organisms spanning all kingdoms recognize specific sites on DNA to shape the genome and regulate various genomic functions, and may or may not undergo chemical modifications regulated by complex pathways. Very often they recognize multiple specific sequences to which they bind with different affinities and cooperatively (304, 305). These tunable, cooperative interactions determine ubiquitous architectural DNA modifications such as repressor bound to operator, the dynamic of which has not been directly investigated in earlier studies on transcription roadblocks either *in vivo* or *in vitro*.

While an accessory factor protein may regulate initiation from the promoter near one of the bridged binding sites, its presence at the other site may be an obstacle

for RNA polymerases transcribing a different gene. To observe the road blocking effect of accessory protein bind operator alters the extent to which a protein halts transcription, the lac repressor protein (LacI) was used. LacI can bind to a high affinity site adjacent to the promoter (O1), a ~10-fold lower affinity, secondary site ~400 bp downstream (O2), and an ~100-fold lower affinity, tertiary site ~90 bp upstream (O3) (306-308). A strong, symmetric operator, O<sub>s</sub>, which LacI binds with 10-fold greater affinity than O1, has also been engineered (309). The auxiliary sites are thought to serve as reservoirs to elevate the concentration of LacI in proximity to the primary, promoter-blocking site by forming loops to deliver LacI to the primary site (310). In addition, LacI tetramers exhibit higher affinity for operators in looped as compared to unlooped DNA (266). LacI bound to an operator can block transcription initiation with up to 99.5% efficiency *in vivo* (311, 312) that depends on the promoter firing rate and whether multiple RNA polymerases act cooperatively (313). Lac repressor can also obstruct eukaryotic RNA polymerase II (314) to a greater or lesser degree depending on accessory factors (315).

The extent to which a DNA loop accentuates the repression of transcription initiation at the lac operon has been amply characterized; however, its strength as a roadblock to an advancing transcription elongation complex has never been characterized. Therefore, in the study, the progress of RNA polymerase through LacI obstacles bound to O1 operator on torsionally relaxed, unlooped DNA templates was monitored using magnetic tweezers (MT). When the RNA polymerase encountered the LacI, it paused for several minutes. The average paused lifetime was comparable to the



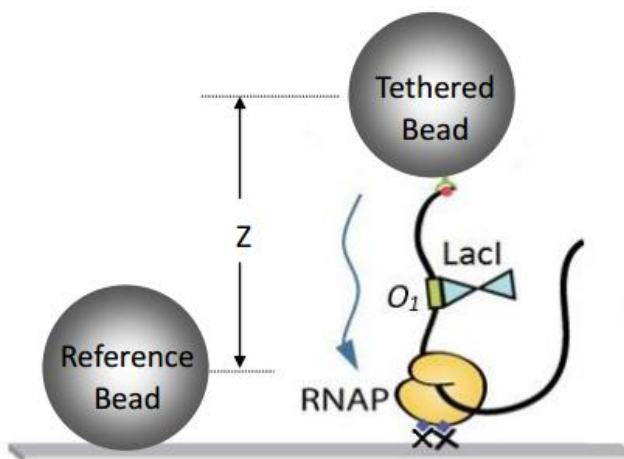
lifetime of LacI binding measures on templates with no active transcription, which suggests that instead of dislodging the obstacle, RNA polymerase waits for *lac* repressor to dissociate before proceeding. Occasionally, either immediately or after a pause, the tether length abruptly to the initial length observed when RNA polymerase was posed at the promoter. This suggests that RNA polymerase may somehow shuttle back towards the promoter.

## § 7.2 Results

### § 7.2.1 RNA polymerase overcomes pausing in magnetic tweezer experiments

#### Analysis of transcription pausing using magnetic tweezers

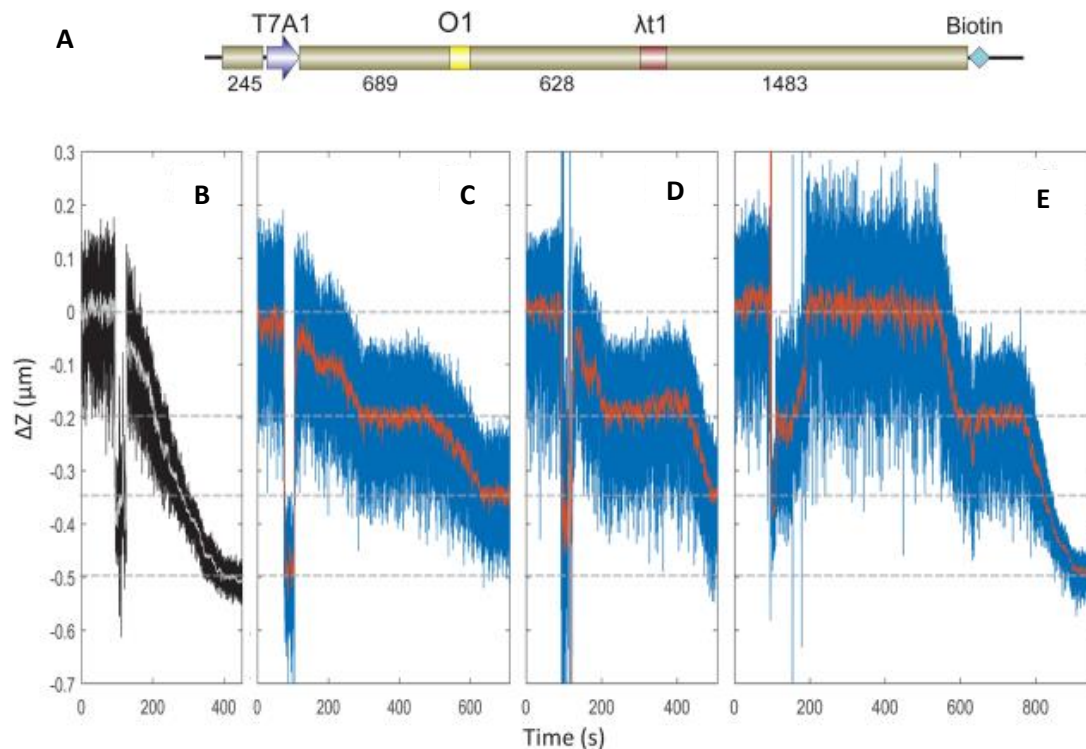
To understand what happens to an RNA polymerase that encounters a LacI obstacle, the real-time progress was tracked using magnetic tweezers (Fig. 7.1). A DNA tether containing only the O1 operator was used (Fig. 7.2 A), and the change in the length of the tether was monitored after a missing nucleotide was added to



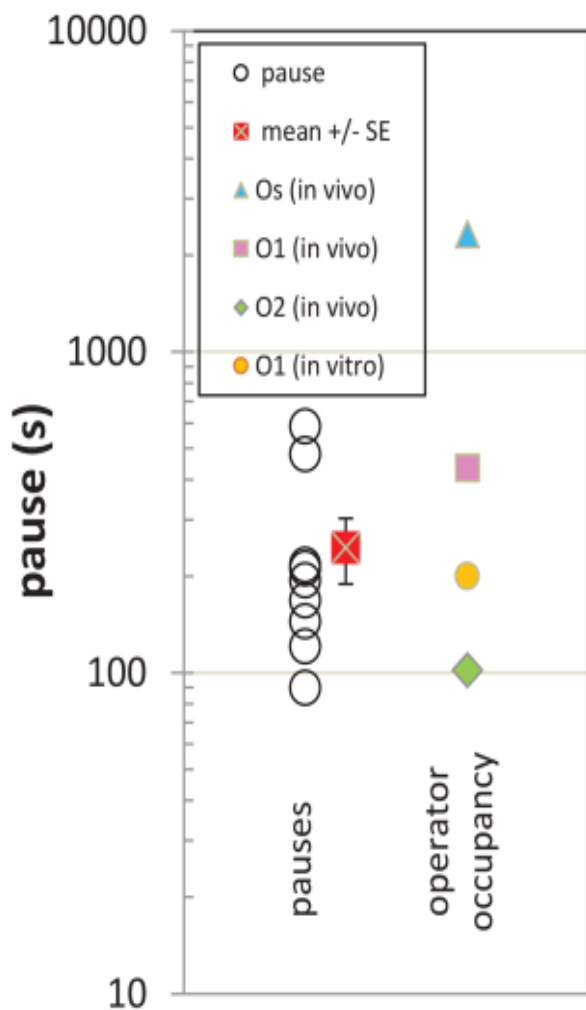
**Fig. 7.1** A schematic diagram of the magnetic tweezer transcription assay. The DNA tether is attached to the bead through a streptavidin-biotin linkage (green-red), and RNAP (yellow) is attached to the surface through linkages between two HA tags on the enzyme and the anti-HA coated glass surface of the microchamber. The tether includes an O1 operator to which lac repressor may bind. Transcriptional activity will shorten the DNA tether as RNA polymerase draws the DNA template through itself.

allow stalled RNA polymerases to resume transcription. Figure 7.2 B–E shows representative traces of the average change in the extension of DNA tethers as a function of time. The DNA extension is maximal at the beginning of the experiment when TECs were stalled at position +22 for lack of CTP (Fig. 7.2 B–E, top dashed line at 0). When the missing nucleotide was added, elongation resumed after random delays and was detected as a progressive decrease in tether length after turbulence subsided.

The leftmost record (Fig. 7.2 B) was in the absence of LacI and transcription proceeded smoothly consuming the entire tether, so that the bead was drawn down to the glass surface. Figure 7.2 C–E show transcription progress in the presence of 1 or 10 nM LacI. Pauses in tether shortening (transcription elongation) were observed at the extension corresponding to O1 (dashed line at  $-0.2 \mu\text{m}$ ). Pauses at the roadblock were measured and are reported in Figure 7.3 together with the mean lifetimes of LacI-operator complexes estimated as the inverse of the measured off-rates (313). Notice that the mean pause time is similar to the average dwell time of LacI on O1 *in vitro*. Furthermore, the average pause is longer than that measured previously for TECs with no obstacles,  $90 \pm 14 \text{ s}$ . (316). Sometimes transcription ceased at the terminator (Fig. 7.2 C, D, dashed line at  $-0.35 \mu\text{m}$ ), but alternatively it continued until the bead was drawn down to the micro-chamber surface (Fig. 7.2 B, E, dashed line at  $-0.5 \mu\text{m}$ ).



**Fig. 7.2** LacI bound to an O1 operator pauses transcription. (A) A schematic representation of the DNA template used in magnetic tweezer transcription assays. The template contained a T7A1 promoter close to the upstream end, a stall site at position +22, an O1 operator, the lambda t1 terminator ( $\lambda t1$ ) and a biotin label at the downstream end. A streptavidin-labeled paramagnetic bead was coupled to the biotin label to for micromanipulation in the magnetic tweezer. Four examples of transcriptional elongation recorded using the magnetic tweezers are displayed. In (B) no LacI was included and transcription shortened the DNA tether progressively without interruption. When LacI was included (C-E), transcription shortened the tether by about 0.2  $\mu\text{m}$  before pausing for about 200 s and then resuming. Transcription finally ceased after the tether shortened by either 0.35  $\mu\text{m}$  (C and D), a distance corresponding to the location of a terminator sequence, or 0.5  $\mu\text{m}$  (B and E), a distance corresponding to the end of the template.

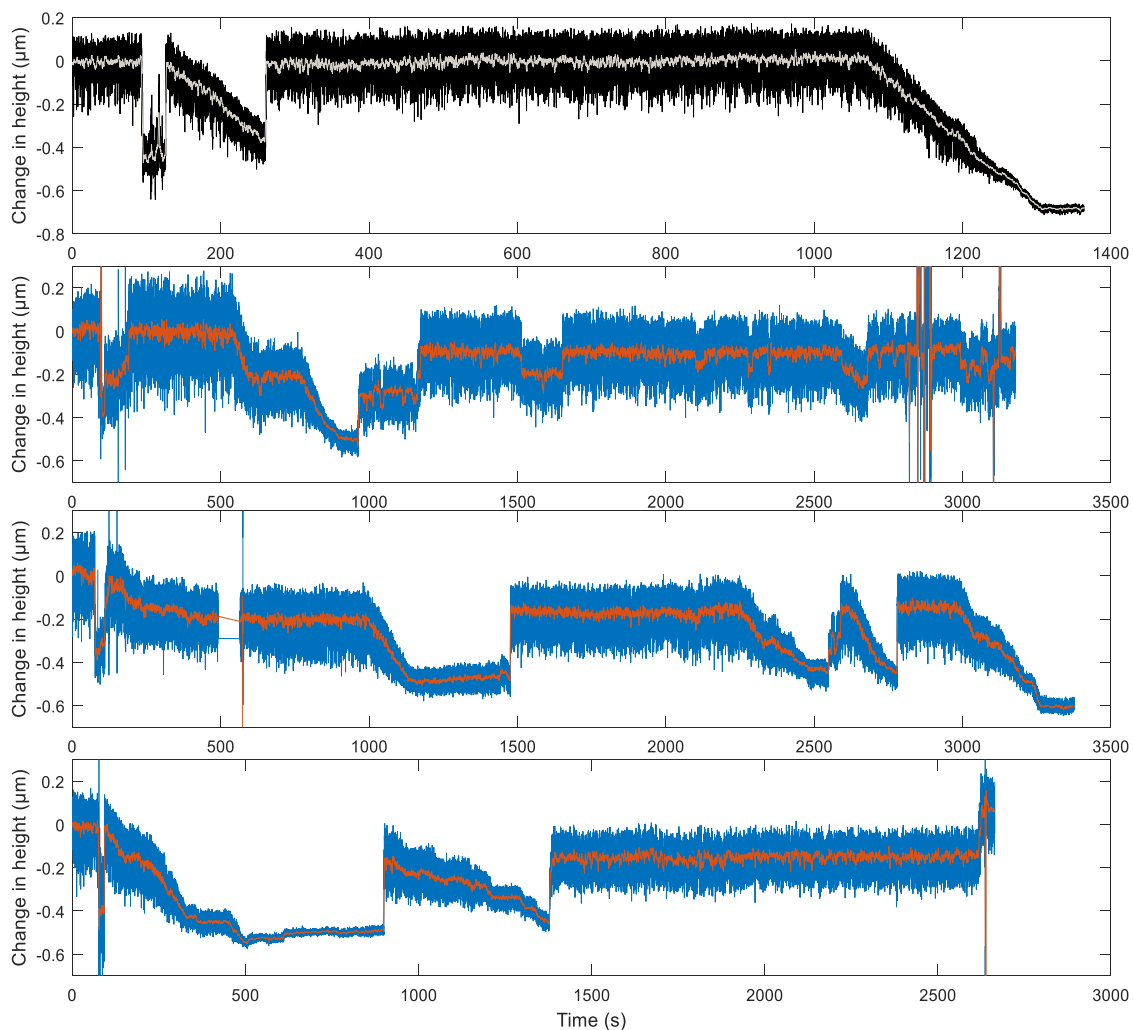


**Fig. 7.3** TEC pauses and expected operator occupancy. Sections of time series like those displayed in Fig. 7.4 were fitted as described in Fig. S6 in (2) to determine pause intervals at the position corresponding to the O1 operator site (o). The mean value is also plotted with standard error indicated (red x: 253 s). This mean is considerably shorter than expected the mean interval of occupancy of the operator in vivo calculated as the inverse of the reported dissociation rate of LacI from O1 (pink □: 434 s). Also shown are values for the expected mean intervals of in vivo occupancy of operators O2 (green ◇: 102 s) and Os (blue Δ: 2326 s) and in vitro occupancy of O1 (orange circle: 200 s).

### § 7.2.2 RNA polymerase may shuttling backward once encounter obstacle

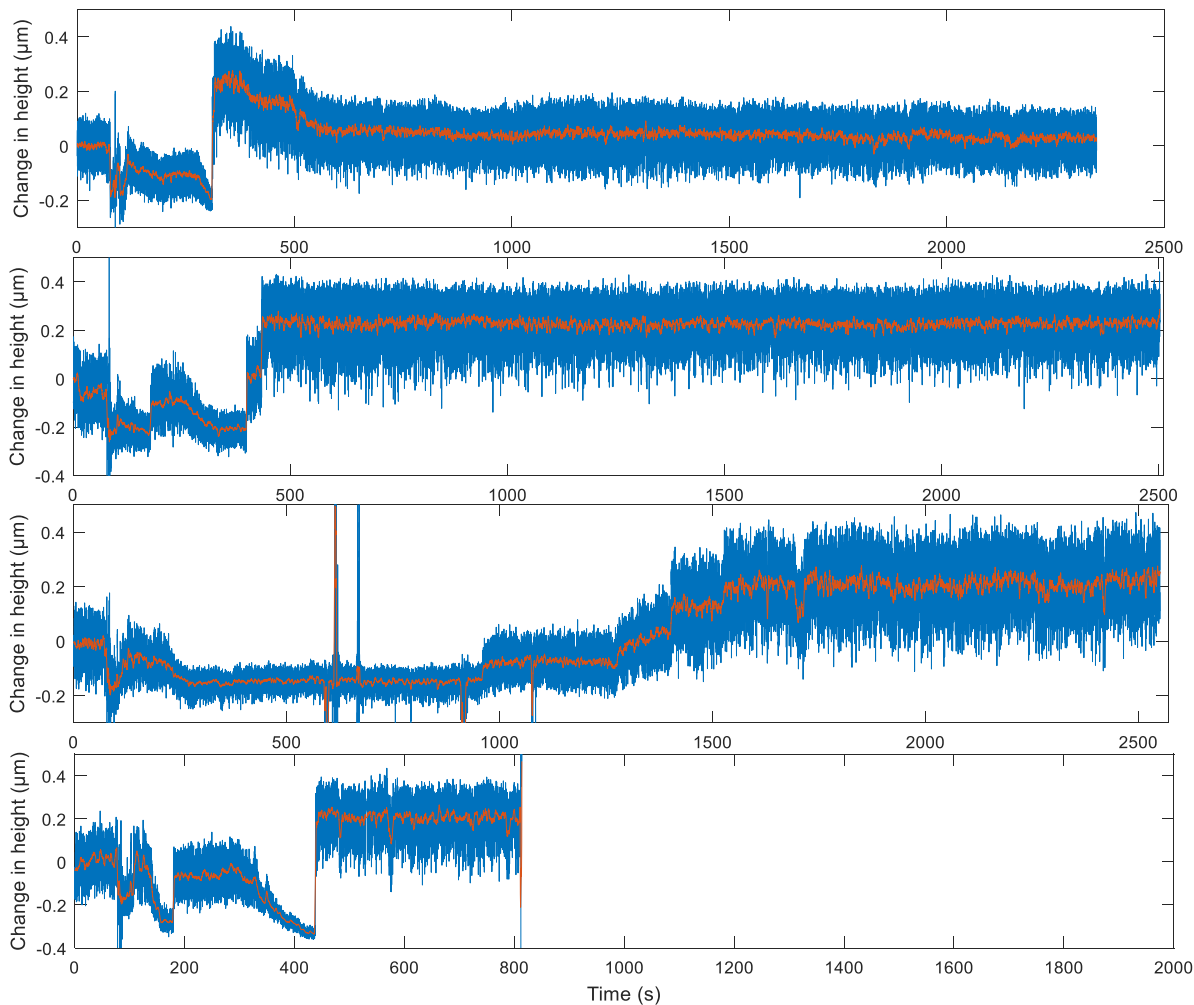
Other than pausing when RNA polymerase encounter obstacle, RNA polymerase may shuttling back toward promoter or back to the LacI obstacle (Fig. 7.4). Using the same magnetic tweezers set up as above (Fig. 7.2 A), the change in the length of the tether in transcription was monitored after a missing nucleotide was added to allow stalled RNA polymerases to resume transcription.

The top record (Fig. 7.4, black) was in the absence of LacI and transcription proceeded smoothly consuming the entire tether, so that the bead was drawn down



**Fig. 7.4** RNA polymerase (RNAP) show transcription forward and shuttling back in the absence of Lacl (black) and in the presence of Lacl (blue). In the absence of Lacl, RNAP transcribe to terminator then immediately shuttling back to promoter and re-transcribe after a long pause. In the presence of Lacl, RNAP transcribe among promoter, operator, terminator and the end of DNA template.

smoothly. When RNA polymerase encounters the obstacle ( $\lambda$ t1 terminator in this example), instead of dissociates from DNA tether so that the bead dis-attach, RNAP sliding back to promoter. Then RNAP transcribe smoothly to glass surface after a long pause. In the presence of 1 or 10 nM Lacl (Fig. 7.4 blue), transcription cycles also appear among promoter, operator, terminator and the end of DNA template.



**Fig. 7.5** RNA polymerase (RNAP) show transcription forward and shuttling back to an extension exceed the upstream end of the DNA template.

### § 7.2.3 DNA extension can reach an extension beyond the end of DNA template

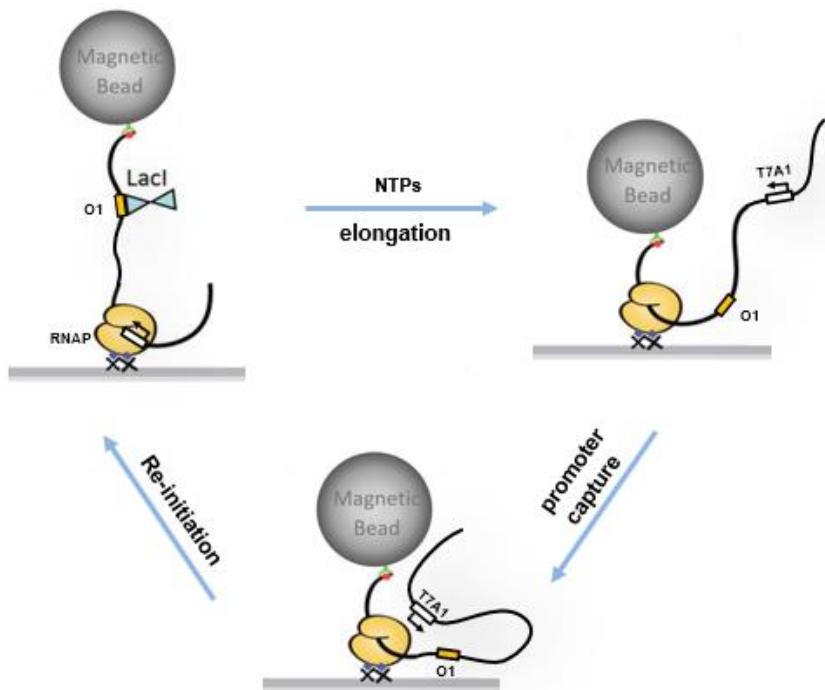
When RNA polymerase sliding back, it may reach DNA extension beyond the predicted upstream end of DNA as it shown in Figure 7.5. In the figure, after the DNA tether transcribe to the predicted loci operator or terminator, it jumps to an extension of approximately 0.2  $\mu\text{m}$  which correspond to 600 bp upstream of the +22 stalled site (0 extension in the figure). However, based on the DNA template (Fig. 7.2 A), there are less

than 300 bp DNA sequence upstream of the +22 stalled site, which is unlikely that RNA polymerase sliding to an upstream location and induce a longer extension. What's more, since the extension of single stranded DNA is smaller than double stranded DNA under the tension applied in this experiment, 0.3 pN (153, 317), it is unlikely that double stranded DNA unzipping to two single stranded DNA during transcription.

### § 7.3 Discussion and Conclusion

#### Does RNAP wait for LacI to dissociate or does it eject it?

Our results demonstrate that RNA polymerase halts but eventually bypasses LacI bound operator sites. The dissociation rates for LacI bound to DNA have been determined



**Fig. 7.6** Possible mechanism for RNAP shuttling. Transcription elongate with the existence of NTPs, and able to recapture promoter after dissociate from the DNA template and reinitiate the transcriptional elongation.

previously and suggest that in vivo LacI will dissociate after 2326, 434, or 102 s from Os, O1, and O2 respectively (308). For in vitro conditions similar to those described here, the O1 dissociation rate is reported to be 0.3/min (163); the reciprocal suggests that LacI occupies an O1 operator for an average of 200 s. In the magnetic tweezer experiments pauses at the O1 site lasted an average of 253 s. This seems to indicate that in vitro, without accessory factors, RNA polymerase waits for LacI to dissociate.

### **Possible mechanism for shuttling by RNAP**

In our MT observation, RNAP shuttling back toward promoter or obstacles immediately upon encounter the obstacles, one possible mechanism can be illustrated as figure 7.6. RNAP elongates on DNA template in the existence of NTPs. When RNAP dissociate from DNA template, instead of diffusing away, it may recapture the promoter and reinitiate the transcription (Fig. 7.4, black).

Note: Part of this chapter was originally published as part of the work "Proteins mediating DNA loops effectively block transcription", Protein Science, 2017. DOI: 10.1002/pro.3156



## References

1. Yan Y, Ding Y, Leng F, Dunlap D, & Finzi L (2018) Protein-mediated loops in supercoiled DNA create large topological domains. *Nucleic Acids Research*:gky153-gky153.
2. Vörös Z, Yan Y, Kovari DT, Finzi L, & Dunlap D (2017) Proteins mediating DNA loops effectively block transcription. *Protein Science* 26(7):1427-1438.
3. Lodish H, *et al.* (1995) *Molecular cell biology* (Scientific American Books New York).
4. Priest DG, *et al.* (2014) Quantitation of the DNA tethering effect in long-range DNA looping in vivo and in vitro using the Lac and  $\lambda$  repressors. *Proceedings of the National Academy of Sciences* 111(1):349-354.
5. Gruber S (2014) Multilayer chromosome organization through DNA bending, bridging and extrusion. *Current opinion in microbiology* 22:102-110.
6. Watson JD & Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171(4356):737-738.
7. Schiessel H (2003) The physics of chromatin. *Journal of Physics: Condensed Matter* 15(19):R699.
8. Alberts B, *et al.* (Molecular biology of the cell. 4th edn. New York: Garland Science; 2002. (ISBN 0-8153-3218-1).
9. Dickerson RE, *et al.* (1982) The anatomy of A-, B-, and Z-DNA. *Science (New York, N.Y.)* 216(4545):475-485.
10. Marko JF & Neukirch S (2013) Global force-torque phase diagram for the DNA double helix: structural transitions, triple points, and collapsed plectonemes. *Physical review. E, Statistical, nonlinear, and soft matter physics* 88(6):062722.
11. Bar A, Kabakcoglu A, & Mukamel D (2011) Denaturation of circular DNA: supercoil mechanism. *Physical review. E, Statistical, nonlinear, and soft matter physics* 84(4 Pt 1):041935.
12. Sipski ML & Wagner TE (1977) Probing DNA quaternary ordering with circular dichroism spectroscopy: Studies of equine sperm chromosomal fibers. *Biopolymers* 16(3):573-582.
13. Drlica K & Rouviere-Yaniv J (1987) Histone-like proteins of bacteria. *Microbiological reviews* 51(3):301-319.
14. Kar S, Edgar R, & Adhya S (2005) Nucleoid remodeling by an altered HU protein: reorganization of the transcription program. *Proceedings of the National Academy of Sciences of the United States of America* 102(45):16397-16402.
15. Hodges-Garcia Y, Hagerman PJ, & Pettijohn DE (1989) DNA ring closure mediated by protein HU. *The Journal of biological chemistry* 264(25):14621-14623.
16. Rouviere-Yaniv J & Kjeldgaard NO (1979) Native Escherichia coli HU protein is a heterotypic dimer. *FEBS letters* 106(2):297-300.
17. Ali Azam T, Iwata A, Nishimura A, Ueda S, & Ishihama A (1999) Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid. *Journal of Bacteriology* 181(20):6361-6370.
18. Dixon NE & Kornberg A (1984) Protein HU in the enzymatic replication of the chromosomal origin of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* 81(2):424-428.
19. Wong JTY, New DC, Wong JCW, & Hung VKL (2003) Histone-Like Proteins of the Dinoflagellate *Cryptothecodinium cohnii* Have Homologies to Bacterial DNA-Binding Proteins. *Eukaryotic Cell* 2(3):646-650.

20. Broyles SS & Pettijohn DE (1986) Interaction of the Escherichia coli HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch. *Journal of molecular biology* 187(1):47-60.
21. Skarstad K, Baker TA, & Kornberg A (1990) Strand separation required for initiation of replication at the chromosomal origin of E.coli is facilitated by a distant RNA-DNA hybrid. *The EMBO Journal* 9(7):2341-2348.
22. Swinger KK & Rice PA (2007) Structure-based analysis of HU-DNA binding. *Journal of molecular biology* 365(4):1005-1016.
23. Kamashev D, Balandina A, & Rouviere-Yaniv J (1999) The binding motif recognized by HU on both nicked and cruciform DNA. *The EMBO Journal* 18(19):5434-5444.
24. Pinson V, Takahashi M, & Rouviere-Yaniv J (1999) Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA. *Journal of molecular biology* 287(3):485-497.
25. Castaing B, Zelwer C, Laval J, & Boiteux S (1995) HU protein of Escherichia coli binds specifically to DNA that contains single-strand breaks or gaps. *The Journal of biological chemistry* 270(17):10291-10296.
26. Kamashev D & Rouviere-Yaniv J (2000) The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *Embo j* 19(23):6527-6535.
27. Pontiggia A, Negri A, Beltrame M, & Bianchi ME (1993) Protein HU binds specifically to kinked DNA. *Molecular microbiology* 7(3):343-350.
28. Xiao B, Johnson RC, & Marko JF (2010) Modulation of HU-DNA interactions by salt concentration and applied force. *Nucleic Acids Res* 38(18):6176-6185.
29. van Noort J, Verbrugge S, Goosen N, Dekker C, & Dame RT (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proceedings of the National Academy of Sciences of the United States of America* 101(18):6969-6974.
30. Skoko D, Wong B, Johnson RC, & Marko JF (2004) Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochemistry* 43(43):13867-13874.
31. Lia G, et al. (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(20):11373-11377.
32. Sagi D, Friedman N, Vorgias C, Oppenheim AB, & Stavans J (2004) Modulation of DNA conformations through the formation of alternative high-order HU-DNA complexes. *Journal of molecular biology* 341(2):419-428.
33. Lal A, et al. (2016) Genome scale patterns of supercoiling in a bacterial chromosome. *Nat Commun* 7:11055.
34. Crick FH (1958) On protein synthesis. *Symposia of the Society for Experimental Biology* 12:138-163.
35. Crick F (1970) Central dogma of molecular biology. *Nature* 227(5258):561-563.
36. Finn RD, Orlova EV, Gowen B, Buck M, & van Heel M (2000) Escherichia coli RNA polymerase core and holoenzyme structures. *The EMBO Journal* 19(24):6833-6844.
37. Darst SA, Kubalek EW, & Kornberg RD (1989) Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography. *Nature* 340(6236):730-732.
38. Burgess RR, Travers AA, Dunn JJ, & Bautz EK (1969) Factor stimulating transcription by RNA polymerase. *Nature* 221(5175):43-46.
39. Zhang G & Darst SA (1998) Structure of the Escherichia coli RNA polymerase alpha subunit amino-terminal domain. *Science (New York, N.Y.)* 281(5374):262-266.

40. Gunnelius L, *et al.* (2014) The omega subunit of the RNA polymerase core directs transcription efficiency in cyanobacteria. *Nucleic acids research* 42(7):4606-4614.
41. Murakami KS & Darst SA (2003) Bacterial RNA polymerases: the whole story. *Current opinion in structural biology* 13(1):31-39.
42. Young BA, Gruber TM, & Gross CA (2002) Views of transcription initiation. *Cell* 109(4):417-420.
43. Kapanidis AN, *et al.* (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science (New York, N.Y.)* 314(5802):1144-1147.
44. Revyakin A, Liu C, Ebright RH, & Strick TR (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science (New York, N.Y.)* 314(5802):1139-1143.
45. Hsu LM (2002) Promoter clearance and escape in prokaryotes. *Biochimica et biophysica acta* 1577(2):191-207.
46. Luse DS & Jacob GA (1987) Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter. *The Journal of biological chemistry* 262(31):14990-14997.
47. Carpousis AJ & Gralla JD (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry* 19(14):3245-3253.
48. Ellinger T, Behnke D, Bujard H, & Gralla JD (1994) Stalling of Escherichia coli RNA polymerase in the +6 to +12 region in vivo is associated with tight binding to consensus promoter elements. *Journal of molecular biology* 239(4):455-465.
49. Hsu LM (2009) Monitoring Abortive Initiation. *Methods (San Diego, Calif.)* 47(1):25-36.
50. Abbondanzieri EA, Shaevitz JW, & Block SM (2005) Picocalorimetry of transcription by RNA polymerase. *Biophysical journal* 89(6):L61-63.
51. Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, & Block SM (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* 438(7067):460-465.
52. Adelman K, *et al.* (2002) Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proceedings of the National Academy of Sciences of the United States of America* 99(21):13538-13543.
53. Wang MD, *et al.* (1998) Force and velocity measured for single molecules of RNA polymerase. *Science (New York, N.Y.)* 282(5390):902-907.
54. Liu LF & Wang JC (1987) Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences* 84(20):7024-7027.
55. Epshtein V, *et al.* (2002) Swing-gate model of nucleotide entry into the RNA polymerase active center. *Molecular cell* 10(3):623-634.
56. Richardson JP (1991) Preventing the synthesis of unused transcripts by Rho factor. *Cell* 64(6):1047-1049.
57. Whitaker JW, Chen Z, & Wang W (2015) Predicting the Human Epigenome from DNA Motifs. *Nature methods* 12(3):265-272.
58. Muller J, Oehler S, & Muller-Hill B (1996) Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *Journal of molecular biology* 257(1):21-29.
59. Majors J (1975) Initiation of in vitro mRNA synthesis from the wild-type lac promoter. *Proceedings of the National Academy of Sciences of the United States of America* 72(11):4394-4398.
60. Nick H & Gilbert W (1985) Detection in vivo of protein-DNA interactions within the lac operon of Escherichia coli. *Nature* 313(6005):795-798.

61. Friedman AM, Fischmann TO, & Steitz TA (1995) Crystal structure of lac repressor core tetramer and its implications for DNA looping. *Science (New York, N.Y.)* 268(5218):1721-1727.
62. CHEN B, *et al.* (1971) On the mechanism of action of lac repressor. *Nature* 233(37):67-70.
63. Straney SB & Crothers DM (1987) Lac repressor is a transient gene-activating protein. *Cell* 51(5):699-707.
64. Bahl CP, Wu R, Stawinsky J, & Narang SA (1977) Minimal length of the lactose operator sequence for the specific recognition by the lactose repressor. *Proceedings of the National Academy of Sciences of the United States of America* 74(3):966-970.
65. Mehta RA & Kahn JD (1999) Designed hyperstable Lac repressor. DNA loop topologies suggest alternative loop geometries. *Journal of molecular biology* 294(1):67-77.
66. Schleif R (1992) DNA looping. *Annual review of biochemistry* 61:199-223.
67. Zhang Y, McEwen AE, Crothers DM, & Levene SD (2006) Statistical-Mechanical Theory of DNA Looping. *Biophysical journal* 90(6):1903-1912.
68. Hao N, Sneppen K, Shearwin KE, & Dodd IB (2017) Efficient chromosomal-scale DNA looping in Escherichia coli using multiple DNA-looping elements. *Nucleic Acids Res* 45(9):5074-5085.
69. Johnson S, Linden M, & Phillips R (2012) Sequence dependence of transcription factor-mediated DNA looping. *Nucleic Acids Res* 40(16):7728-7738.
70. Mulligan Peter J, Chen Y-J, Phillips R, & Spakowitz Andrew J (Interplay of Protein Binding Interactions, DNA Mechanics, and Entropy in DNA Looping Kinetics. *Biophysical journal* 109(3):618-629.
71. Rippe K (2001) Making contacts on a nucleic acid polymer. *Trends in Biochemical Sciences* 26(12):733-740.
72. Dunn TM, Hahn S, Ogden S, & Schleif RF (1984) An operator at -280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proceedings of the National Academy of Sciences* 81(16):5017-5020.
73. Lee DH & Schleif RF (1989) In vivo DNA loops in araCBAD: size limits and helical repeat. *Proceedings of the National Academy of Sciences of the United States of America* 86(2):476-480.
74. Becker NA, Kahn JD, & Maher LJ, 3rd (2005) Bacterial repression loops require enhanced DNA flexibility. *Journal of molecular biology* 349(4):716-730.
75. Camsund D, Heidorn T, & Lindblad P (2014) Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *Journal of biological engineering* 8(1):4.
76. Bintu L, *et al.* (2005) Transcriptional regulation by the numbers: models. *Current opinion in genetics & development* 15(2):116-124.
77. Buchler NE, Gerland U, & Hwa T (2003) On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences* 100(9):5136-5141.
78. Law SM, Bellomy GR, Schlax PJ, & Record MT, Jr. (1993) In vivo thermodynamic analysis of repression with and without looping in lac constructs. Estimates of free and local lac repressor concentrations and of physical properties of a region of supercoiled plasmid DNA in vivo. *Journal of molecular biology* 230(1):161-173.
79. Vilar JM & Saiz L (2005) DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Current opinion in genetics & development* 15(2):136-144.

80. Saiz L, Rubi JM, & Vilar JMG (2005) Inferring the in vivo looping properties of DNA. *Proceedings of the National Academy of Sciences of the United States of America* 102(49):17642-17645.
81. Allemand JF, Cocco S, Douarche N, & Lia G (2006) Loops in DNA: an overview of experimental and theoretical approaches. *The European physical journal. E, Soft matter* 19(3):293-302.
82. Vologodskii A & Cozzarelli NR (1996) Effect of supercoiling on the juxtaposition and relative orientation of DNA sites. *Biophysical journal* 70(6):2548-2556.
83. Krämer H, *et al.* (1987) lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. *The EMBO Journal* 6(5):1481-1491.
84. Hsieh W-T, Whitson PA, Matthews K, & Wells RD (1987) Influence of sequence and distance between two operators on interaction with the lac repressor. *Journal of Biological Chemistry* 262(30):14583-14591.
85. Whitson PA, Hsieh WT, Wells RD, & Matthews KS (1987) Influence of supercoiling and sequence context on operator DNA binding with lac repressor. *The Journal of biological chemistry* 262(30):14592-14599.
86. Krämer H, Amouyal M, Nordheim A, & Müller-Hill B (1988) DNA supercoiling changes the spacing requirement of two lac operators for DNA loop formation with lac repressor. *The EMBO Journal* 7(2):547-556.
87. Schafer DA, Gelles J, Sheetz MP, & Landick R (1991) Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature* 352(6334):444-448.
88. Kovari DT, Yan Y, Finzi L, & Dunlap D (2018) Tethered Particle Motion: An Easy Technique for Probing DNA Topology and Interactions with Transcription Factors. *Single Molecule Analysis: Methods and Protocols*, ed Peterman EJG (Springer New York, New York, NY), pp 317-340.
89. Finzi L & Gelles J (1995) Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules. *Science (New York, N.Y.)* 267(5196):378-380.
90. Yin H, Landick R, & Gelles J (1994) Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophysical journal* 67(6):2468-2478.
91. Tolić-Nørrelykke SF, Engh AM, Landick R, & Gelles J (2004) Diversity in the rates of transcript elongation by single RNA polymerase molecules. *Journal of Biological Chemistry* 279(5):3292-3299.
92. Blumberg S, Gajraj A, Pennington MW, & Meiners J-C (2005) Three-Dimensional Characterization of Tethered Microspheres by Total Internal Reflection Fluorescence Microscopy. *Biophysical journal* 89(2):1272-1281.
93. Pouget N, *et al.* (2004) Single-particle tracking for DNA tether length monitoring. *Nucleic Acids Res* 32(9):e73.
94. Segall DE, Nelson PC, & Phillips R (2006) Volume-exclusion effects in tethered-particle experiments: bead size matters. *Physical review letters* 96(8):088306.
95. Nelson PC, *et al.* (2006) Tethered particle motion as a diagnostic of DNA tether length. *The journal of physical chemistry. B* 110(34):17260-17267.
96. Zurla C, *et al.* (2006) Novel tethered particle motion analysis of CI protein-mediated DNA looping in the regulation of bacteriophage lambda. *Journal of Physics: Condensed Matter* 18(14):S225.
97. Zurla C, *et al.* (2007) Integration host factor alters LacI-induced DNA looping. *Biophysical chemistry* 128(2-3):245-252.

98. van den Broek B, Vanzi F, Normanno D, Pavone FS, & Wuite GJL (2006) Real-time observation of DNA looping dynamics of Type IIE restriction enzymes NaeI and NarI. *Nucleic Acids Research* 34(1):167-174.
99. Wong OK, Guthold M, Erie DA, & Gelles J (2008) Interconvertible lac repressor-DNA loops revealed by single-molecule experiments. *PLoS Biol* 6(9):e232.
100. Rutkauskas D, Zhan H, Matthews KS, Pavone FS, & Vanzi F (2009) Tetramer opening in LacI-mediated DNA looping. *Proceedings of the National Academy of Sciences* 106(39):16627-16632.
101. Han L, *et al.* (2009) Concentration and length dependence of DNA looping in transcriptional regulation. *PLoS One* 4(5):e5621.
102. Johnson S, Chen Y-J, & Phillips R (2013) Poly (dA: dT)-rich DNAs are highly flexible in the context of DNA looping. *PLoS One* 8(10):e75799.
103. Revalee Joel D, Blab Gerhard A, Wilson Henry D, Kahn Jason D, & Meiners J-C (2014) Tethered Particle Motion Reveals that LacI-DNA Loops Coexist with a Competitor-Resistant but Apparently Unlooped Conformation. *Biophysical journal* 106(3):705-715.
104. Edelman LM, Cheong R, & Kahn JD (2003) Fluorescence resonance energy transfer over approximately 130 basepairs in hyperstable lac repressor-DNA loops. *Biophysical journal* 84(2 Pt 1):1131-1145.
105. Morgan MA, Okamoto K, Kahn JD, & English DS (2005) Single-Molecule Spectroscopic Determination of Lac Repressor-DNA Loop Conformation. *Biophysical journal* 89(4):2588-2596.
106. Ruben GC & Roos TB (1997) Conformation of Lac repressor tetramer in solution, bound and unbound to operator DNA. *Microscopy research and technique* 36(5):400-416.
107. Johnson S, van de Meent J-W, Phillips R, Wiggins CH, & Lindén M (2014) Multiple LacI-mediated loops revealed by Bayesian statistics and tethered particle motion. *Nucleic Acids Research* 42(16):10265-10277.
108. Towles KB, Beausang JF, Garcia HG, Phillips R, & Nelson PC (2009) First-principles calculation of DNA looping in tethered particle experiments. *Physical biology* 6(2):025001.
109. Swigon D, Coleman BD, & Olson WK (2006) Modeling the Lac repressor-operator assembly: the influence of DNA looping on Lac repressor conformation. *Proceedings of the National Academy of Sciences of the United States of America* 103(26):9879-9884.
110. Zhang Y, McEwen AE, Crothers DM, & Levene SD (2006) Analysis of in-vivo LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS One* 1(1):e136.
111. Mirkin SM (2001) DNA topology: fundamentals. *eLS*.
112. Worcel A & Burgi E (1972) On the structure of the folded chromosome of Escherichia coli. *Journal of molecular biology* 71(2):127-147.
113. Sinden RR & Pettijohn DE (1981) Chromosomes in living Escherichia coli cells are segregated into domains of supercoiling. *Proceedings of the National Academy of Sciences of the United States of America* 78(1):224-228.
114. Kavenoff R & Bowen BC (1976) Electron microscopy of membrane-free folded chromosomes from Escherichia coli. *Chromosoma* 59(2):89-101.
115. Kavenoff R & Ryder OA (1976) Electron microscopy of membrane-associated folded chromosomes of Escherichia coli. *Chromosoma* 55(1):13-25.
116. Delius H & Worcel A (1974) Electron microscopic studies on the folded chromosome of Escherichia coli. *Cold Spring Harbor symposia on quantitative biology* 38:53-58.

117. Higgins NP, Yang X, Fu Q, & Roth JR (1996) Surveying a supercoil domain by using the gamma delta resolution system in *Salmonella typhimurium*. *Journal of Bacteriology* 178(10):2825-2835.
118. Postow L, Hardy CD, Arsuaga J, & Cozzarelli NR (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes & Development* 18(14):1766-1779.
119. Trun NJ & Marko JF (1998) Architecture of a bacterial chromosome. *Asm News* 64(5):276-283.
120. Gellert M, O'Dea MH, Itoh T, & Tomizawa J (1976) Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proceedings of the National Academy of Sciences of the United States of America* 73(12):4474-4478.
121. Gellert M, Mizuuchi K, O'Dea MH, Itoh T, & Tomizawa JI (1977) Nalidixic acid resistance: a second genetic character involved in DNA gyrase activity. *Proceedings of the National Academy of Sciences of the United States of America* 74(11):4772-4776.
122. Ullsperger CJ, Vologodskii AV, & Cozzarelli NR (1995) Unlinking of DNA by Topoisomerases During DNA Replication. *Nucleic Acids and Molecular Biology*, eds Eckstein F & Lilley DMJ (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 115-142.
123. Schwartzman JB & Stasiak A (2004) A topological view of the replicon. *EMBO reports* 5(3):256-261.
124. Espeli O, Levine C, Hassing H, & Marians KJ (2003) Temporal regulation of topoisomerase IV activity in *E. coli*. *Molecular cell* 11(1):189-201.
125. Vologodskii AV & Cozzarelli NR (1993) Monte Carlo analysis of the conformation of DNA catenanes. *Journal of molecular biology* 232(4):1130-1140.
126. Beausang JF, *et al.* (2007) DNA looping kinetics analyzed using diffusive hidden Markov model. *Biophysical journal* 92(8):L64-66.
127. Singh-Zocchi M, Dixit S, Ivanov V, & Zocchi G (2003) Single-molecule detection of DNA hybridization. *Proceedings of the National Academy of Sciences* 100(13):7605-7610.
128. Tolic-Norrelykke SF, Rasmussen MB, Pavone FS, Berg-Sorensen K, & Oddershede LB (2006) Stepwise bending of DNA by a single TATA-box binding protein. *Biophysical journal* 90(10):3694-3703.
129. Kumar S, *et al.* (2014) Enhanced tethered-particle motion analysis reveals viscous effects. *Biophysical journal* 106(2):399-409.
130. Vanzi F, Broggio C, Sacconi L, & Pavone FS (2006) Lac repressor hinge flexibility and DNA looping: single molecule kinetics by tethered particle motion. *Nucleic Acids Res* 34(12):3409-3420.
131. Smith SB, Finzi L, & Bustamante C (1992) Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science (New York, N.Y.)* 258(5085):1122-1126.
132. Strick T, Allemand J, Bensimon D, Bensimon A, & Croquette V (1996) The elasticity of a single supercoiled DNA molecule. *Science (New York, N.Y.)* 271(5257):1835.
133. Marko JF & Siggia ED (1995) Stretching DNA. *Macromolecules* 28(26):8759-8770.
134. Bednar J, *et al.* (1995) Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *Journal of molecular biology* 254(4):579-594.
135. Strick TR, Croquette V, & Bensimon D (1998) Homologous pairing in stretched supercoiled DNA. *Proceedings of the National Academy of Sciences of the United States of America* 95(18):10579-10583.

136. Mosconi F, Allemand JF, Bensimon D, & Croquette V (2009) Measurement of the Torque on a Single Stretched and Twisted DNA Using Magnetic Tweezers. *Physical review letters* 102(7):078301.
137. Voros Z, Yan Y, Kovari DT, Finzi L, & Dunlap D (2017) Proteins mediating DNA loops effectively block transcription. *Protein Science* 26(7):1427-1438.
138. Fulcrand G, *et al.* (2016) DNA supercoiling, a critical signal regulating the basal expression of the lac operon in Escherichia coli. *Scientific reports* 6:19243.
139. Yan Y, Leng F, Finzi L, & Dunlap D (Supercoiling drives DNA looping *submitted*).
140. Fulcrand G, *et al.* (2016) DNA supercoiling, a critical signal regulating the basal expression of the lac operon in Escherichia coli. *Sci Rep* 6:19243.
141. Fernández-Sierra M, Shao Q, Fountain C, Finzi L, & Dunlap D (2015) E. coli Gyrase Fails to Negatively Supercoil Diaminopurine-Substituted DNA. *Journal of Molecular Biology* 427(13):2305-2318.
142. Ding Y, *et al.* (2014) DNA supercoiling: A regulatory signal for the  $\lambda$  repressor. *Proceedings of the National Academy of Sciences* 111(43):15402-15407.
143. Ucuncuoglu S, Schneider DA, Weeks ER, Dunlap D, & Finzi L (2017) Multiplexed, Tethered Particle Microscopy for Studies of DNA-Enzyme Dynamics. *Methods in Enzymology*, eds Spies M & Chemla RY (Academic Press), Vol Volume 582, pp 415-435.
144. Kovari DT, Yan Y, Finzi L, & Dunlap D (2018) Tethered Particle Motion: An Easy Technique for Probing DNA Topology and Interactions with Transcription Factors. *Methods in Molecular Biology* 1665:317-340.
145. Han L, *et al.* (2009) Calibration of Tethered Particle Motion Experiments. *Mathematics of DNA Structure, Function and Interactions*, IMA Volumes in Mathematics and its Applications, eds Benham CJ, Harvey S, Olson WK, Sumners DWL, & Swigon D (Springer, New York), Vol 150, pp 123-138.
146. Kumar S, *et al.* (2014) Enhanced Tethered-Particle Motion Analysis Reveals Viscous Effects. *Biophysical Journal* 106(2):399-409.
147. Priest DG, *et al.* (2014) Quantitation of interactions between two DNA loops demonstrates loop domain insulation in E. coli cells. *Proceedings of the National Academy of Sciences* 111(42):E4449-4457.
148. Strick TR, Allemand JF, Bensimon D, Bensimon A, & Croquette V (1996) The elasticity of a single supercoiled DNA molecule. *Science* 271(5257):1835-1837.
149. Kriegel F, Ermann N, & Lipfert J (2017) Probing the mechanical properties, conformational changes, and interactions of nucleic acids with magnetic tweezers. *Journal of Structural Biology* 197(1):26-36.
150. Manzo C & Finzi L (2010) Quantitative Analysis of DNA-Looping Kinetics from Tethered Particle Motion Experiments. *Methods In Enzymology*, ed Nils GW (Academic Press), Vol Volume 475, pp 199-220.
151. Watkins LP & Yang H (2005) Detection of Intensity Change Points in Time-Resolved Single-Molecule Measurements. *J. Phys. Chem. B* 109(1):617-628.
152. Blumberg S, Pennington MW, & Meiners JC (2006) Do femtonewton forces affect genetic function? A review. *J Biol Phys* 32(2):73-95.
153. Wuite GJ, Smith SB, Young M, Keller D, & Bustamante C (2000) Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* 404(6773):103-106.
154. Shao Q, Goyal S, Finzi L, & Dunlap D (2012) Physiological Levels of Salt and Polyamines Favor Writhe and Limit Twist in DNA. *Macromolecules* 45(7):3188-3196.



155. Strick TR, Allemand JF, Bensimon D, & Croquette V (1998) Behavior of Supercoiled DNA. *Biophysical journal* 74(4):2016-2028.
156. Ribeck N & Saleh OA (2008) Multiplexed single-molecule measurements with magnetic tweezers. *The Review of scientific instruments* 79(9):094301.
157. Seol Y & Neuman KC (2011) Magnetic tweezers for single-molecule manipulation. *Methods in molecular biology (Clifton, N.J.)* 783:265-293.
158. Gosse C & Croquette V (2002) Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophysical journal* 82(6):3314-3329.
159. Parthasarathy R (2012) Rapid, accurate particle tracking by calculation of radial symmetry centers. *Nat Methods* 9(7):724-726.
160. Azam TA & Ishihama A (1999) Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *Journal of Biology Chemistry* 274(46):33105-33113.
161. Voros Z, Yan Y, Kovari DT, Finzi L, & Dunlap D (2017) Proteins mediating DNA loops effectively block transcription. *Protein Sci* 26(7):1427-1438.
162. Zurla C, et al. (2007) Integration host factor alters LacI-induced DNA looping. *Biophysical Chemistry* 128(2-3):245-252.
163. Johnson S, van de Meent JW, Phillips R, Wiggins CH, & Linden M (2014) Multiple LacI-mediated loops revealed by Bayesian statistics and tethered particle motion. *Nucleic Acids Res* 42(16):10265-10277.
164. Vanzi F, Broggio C, Sacconi L, & Pavone FS (2006) Lac repressor hinge flexibility and DNA looping: single molecule kinetics by tethered particle motion. *Nucleic Acids Research* 34(12):3409-3420.
165. Ding Y, et al. (2014) DNA supercoiling: a regulatory signal for the lambda repressor. *Proceedings of the National Academy of Sciences of the United States of America* 111(43):15402-15407.
166. Ptashne M (1986) Gene regulation by proteins acting nearby and at a distance. *Nature* 322(6081):697-701.
167. Matthews KS (1992) DNA looping. *Microbiological reviews* 56(1):123-136.
168. Bulger M & Groudine M (2010) Enhancers: the abundance and function of regulatory sequences beyond promoters. *Developmental biology* 339(2):250-257.
169. Gibcus JH & Dekker J (2013) The hierarchy of the 3D genome. *Molecular cell* 49(5):773-782.
170. Vilar JM & Leibler S (2003) DNA looping and physical constraints on transcription regulation. *Journal of molecular biology* 331(5):981-989.
171. Dodd IB, et al. (2004) Cooperativity in long-range gene regulation by the lambda CI repressor. *Genes Dev* 18(3):344-354.
172. Garcia HG & Phillips R (2011) Quantitative dissection of the simple repression input-output function. *Proceedings of the National Academy of Sciences* 108(29):12173-12178.
173. Johnson S, van de Meent JW, Phillips R, Wiggins CH, & Linden M (2014) Multiple LacI-mediated loops revealed by Bayesian statistics and tethered particle motion. *Nucleic Acids Research* 42(16):10265-10277.
174. Wong OK, Guthold M, Erie DA, & Gelles J (2008) Interconvertible lac repressor-DNA loops revealed by single-molecule experiments. *PLoS Biology* 6(9):e232.
175. Normanno D, Vanzi F, & Pavone FS (2008) Single-molecule manipulation reveals supercoiling-dependent modulation of lac repressor-mediated DNA looping. *Nucleic Acids Res* 36(8):2505-2513.

176. Priest DG, *et al.* (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(42):E4449-4457.
177. Tsodikov OV, *et al.* (1999) Wrapping of flanking non-operator DNA in lac repressor-operator complexes: implications for DNA looping1. *Journal of molecular biology* 294(3):639-655.
178. Ringrose L, Chabanis S, Angrand PO, Woodroffe C, & Stewart AF (1999) Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *The EMBO Journal* 18(23):6630-6641.
179. Peters Justin P, Mogil Lauren S, McCauley Micah J, Williams Mark C, & Maher L J (2014) Mechanical Properties of Base-Modified DNA Are Not Strictly Determined by Base Stacking or Electrostatic Interactions. *Biophysical journal* 107(2):448-459.
180. Peters JP, Yelgaonkar SP, Srivatsan SG, Tor Y, & James Maher L, 3rd (2013) Mechanical properties of DNA-like polymers. *Nucleic Acids Res* 41(22):10593-10604.
181. Bond LM, Peters JP, Becker NA, Kahn JD, & Maher LJ (2010) Gene repression by minimal lac loops in vivo. *Nucleic Acids Research* 38(22):8072-8082.
182. Lettice LA, *et al.* (2002) Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proceedings of the National Academy of Sciences of the United States of America* 99(11):7548-7553.
183. Nobrega MA, Ovcharenko I, Afzal V, & Rubin EM (2003) Scanning human gene deserts for long-range enhancers. *Science (New York, N.Y.)* 302(5644):413.
184. Jin F, *et al.* (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503(7475):290-294.
185. Maeda RK & Karch F (2003) Ensuring enhancer fidelity. *Nature genetics* 34(4):360-361.
186. Li G, *et al.* (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148(1-2):84-98.
187. Kwon K-RK, *et al.* (2013) Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* 155(7):1507-1520.
188. Marinic M, Aktas T, Ruf S, & Spitz F (2013) An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Developmental cell* 24(5):530-542.
189. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, & de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Molecular cell* 10(6):1453-1465.
190. Carter D, Chakalova L, Osborne CS, Dai YF, & Fraser P (2002) Long-range chromatin regulatory interactions in vivo. *Nature genetics* 32(4):623-626.
191. Bulger M & Groudine M (Functional and Mechanistic Diversity of Distal Transcription Enhancers. *Cell* 144(3):327-339.
192. Krivega I & Dean A (2012) Enhancer and promoter interactions-long distance calls. *Current opinion in genetics & development* 22(2):79-85.
193. Bellen HJ, *et al.* (1989) P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes Dev* 3(9):1288-1300.
194. Ruf S, *et al.* (2011) Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nature genetics* 43(4):379-386.
195. Shen Y, *et al.* (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature* 488(7409):116-120.
196. Anonymous (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57-74.

197. Nord AS, *et al.* (2013) Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* 155(7):1521-1531.
198. Kwon D, *et al.* (2009) Enhancer-promoter communication at the *Drosophila* engrailed locus. *Development (Cambridge, England)* 136(18):3067-3075.
199. Visel A, *et al.* (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457(7231):854-858.
200. Lieberman-Aiden E, *et al.* (2009) Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)* 326(5950):289-293.
201. Calhoun VC, Stathopoulos A, & Levine M (2002) Promoter-proximal tethering elements regulate enhancer-promoter specificity in the *Drosophila* Antennapedia complex. *Proceedings of the National Academy of Sciences of the United States of America* 99(14):9243-9247.
202. Deng W, *et al.* (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149(6):1233-1244.
203. Cui L, Murchland I, Shearwin KE, & Dodd IB (2013) Enhancer-like long-range transcriptional activation by lambda CI-mediated DNA looping. *Proceedings of the National Academy of Sciences of the United States of America* 110(8):2922-2927.
204. Nolis IK, *et al.* (2009) Transcription factors mediate long-range enhancer-promoter interactions. *Proceedings of the National Academy of Sciences of the United States of America* 106(48):20222-20227.
205. Mahmoudi T, Katsani KR, & Verrijzer C (2002) GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *The EMBO Journal* 21(7):1775-1781.
206. Petrascheck M, *et al.* (2005) DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res* 33(12):3743-3750.
207. Bell AC, West AG, & Felsenfeld G (2001) Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science (New York, N.Y.)* 291(5503):447-450.
208. Bushey AM, Dorman ER, & Corces VG (Chromatin Insulators: Regulatory Mechanisms and Epigenetic Inheritance. *Molecular cell* 32(1):1-9.
209. Hou C, Zhao H, Tanimoto K, & Dean A (2008) CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proceedings of the National Academy of Sciences of the United States of America* 105(51):20398-20403.
210. Gohl D, *et al.* (2011) Mechanism of chromosomal boundary action: roadblock, sink, or loop? *Genetics* 187(3):731-748.
211. Chetverina D, Aoki T, Erokhin M, Georgiev P, & Schedl P (2014) Making connections: insulators organize eukaryotic chromosomes into independent cis-regulatory networks. *BioEssays : news and reviews in molecular, cellular and developmental biology* 36(2):163-172.
212. Corces VG (1995) Chromatin insulators. Keeping enhancers under control. *Nature* 376(6540):462-463.
213. Cai HN & Shen P (2001) Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science (New York, N.Y.)* 291(5503):493-495.
214. Muravyova E, *et al.* (2001) Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science (New York, N.Y.)* 291(5503):495-498.
215. Dixon JR, *et al.* (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485(7398):376-380.
216. Nora EP, *et al.* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485(7398):381-385.

217. Sexton T, *et al.* (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* 148(3):458-472.
218. Hou C, Li L, Qin ZS, & Corces VG (2012) Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. *Molecular cell* 48(3):471-484.
219. Sofueva S, *et al.* (2013) Cohesin-mediated interactions organize chromosomal domain architecture. *The EMBO Journal* 32(24):3119-3129.
220. Phillips-Cremins JE & Corces VG (2013) Chromatin Insulators: Linking genome organization to cellular function. *Molecular cell* 50(4):461-474.
221. Akhtar W, *et al.* (2013) Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell* 154(4):914-927.
222. Ameres SL, *et al.* (2005) Inducible DNA-loop formation blocks transcriptional activation by an SV40 enhancer. *The EMBO Journal* 24(2):358-367.
223. Bondarenko VA, Jiang YI, & Studitsky VM (2003) Rationally designed insulator-like elements can block enhancer action in vitro. *Embo j* 22(18):4728-4737.
224. Nelson PC (2007) Colloidal particle motion as a diagnostic of DNA conformational transitions. *Current Opinion in Colloid & Interface Science* 12(6):307-313.
225. Biton YY, Kumar S, Dunlap D, & Swigon D (2014) Lac repressor mediated DNA looping: Monte Carlo simulation of constrained DNA molecules complemented with current experimental results. *PLoS One* 9(5):e92475.
226. Jian H, Schlick T, & Vologodskii A (1998) Internal motion of supercoiled DNA: brownian dynamics simulations of site juxtaposition. *Journal of molecular biology* 284(2):287-296.
227. Polikanov YS, *et al.* (2007) Probability of the site juxtaposition determines the rate of protein-mediated DNA looping. *Biophysical journal* 93(8):2726-2731.
228. Liu Y, Bondarenko V, Ninfa A, & Studitsky VM (2001) DNA supercoiling allows enhancer action over a large distance. *Proceedings of the National Academy of Sciences of the United States of America* 98(26):14883-14888.
229. Norregaard K, *et al.* (2013) DNA supercoiling enhances cooperativity and efficiency of an epigenetic switch. *Proceedings of the National Academy of Sciences of the United States of America* 110(43):17386-17391.
230. Rovinskiy N, Agbleke AA, Chesnokova O, Pang Z, & Higgins NP (2012) Rates of gyrase supercoiling and transcription elongation control supercoil density in a bacterial chromosome. *PLoS genetics* 8(8):e1002845.
231. Leng F, Chen B, & Dunlap DD (2011) Dividing a supercoiled DNA molecule into two independent topological domains. *Proceedings of the National Academy of Sciences of the United States of America* 108(50):19973-19978.
232. Naughton C, *et al.* (2013) Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature structural & molecular biology* 20(3):387-395.
233. Kulaeva OI, *et al.* (2012) Internucleosomal Interactions Mediated by Histone Tails Allow Distant Communication in Chromatin. *The Journal of biological chemistry* 287(24):20248-20257.
234. Mukhopadhyay S, Schedl P, Studitsky VM, & Sengupta AM (2011) Theoretical analysis of the role of chromatin interactions in long-range action of enhancers and insulators. *Proceedings of the National Academy of Sciences of the United States of America* 108(50):19919-19924.

235. Jun S & Mulder B (2006) Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proceedings of the National Academy of Sciences of the United States of America* 103(33):12388-12393.
236. Villa E, Balaeff A, & Schulten K (2005) Structural dynamics of the lac repressor–DNA complex revealed by a multiscale simulation. *Proceedings of the National Academy of Sciences of the United States of America* 102(19):6783-6788.
237. Brennan LD, Forties RA, Patel SS, & Wang MD (2016) DNA looping mediates nucleosome transfer. *Nature Communications* 7:13337.
238. Hnisz D, Day DS, & Young RA (2016) Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell* 167(5):1188-1200.
239. Duderstadt KE, *et al.* (2016) Simultaneous Real-Time Imaging of Leading and Lagging Strand Synthesis Reveals the Coordination Dynamics of Single Replisomes. *Molecular Cell* 64(6):1035-1047.
240. Qian Z, *et al.* (2016) Genome-Wide Transcriptional Regulation and Chromosome Structural Arrangement by GalR in *E. coli*. *Frontiers in Molecular Biosciences* 3(74).
241. Becker NA, Peters JP, Lionberger TA, & Maher LJ (2013) Mechanism of promoter repression by Lac repressor–DNA loops. *Nucleic Acids Research* 41(1):156-166.
242. Hao N, Sneppen K, Shearwin KE, & Dodd IB (2017) Efficient chromosomal-scale DNA looping in *Escherichia coli* using multiple DNA-looping elements. *Nucleic Acids Research*.
243. Mulligan Peter J, Chen Y-J, Phillips R, & Spakowitz Andrew J (2015) Interplay of Protein Binding Interactions, DNA Mechanics, and Entropy in DNA Looping Kinetics. *Biophysical Journal* 109(3):618-629.
244. Kar S, Edgar R, & Adhya S (2005) Nucleoid remodeling by an altered HU protein: Reorganization of the transcription program. *Proceedings of the National Academy of Sciences* 102(45):16397-16402.
245. Xiao B, Johnson RC, & Marko JF (2010) Modulation of HU–DNA interactions by salt concentration and applied force. *Nucleic Acids Research* 38(18):6176-6185.
246. Lal A, *et al.* (2016) Genome scale patterns of supercoiling in a bacterial chromosome. *Nature Communications* 7:11055.
247. Boedicker JQ, Garcia HG, Johnson S, & Phillips R (2013) DNA sequence-dependent mechanics and protein-assisted bending in repressor-mediated loop formation. *Physical Biology* 10(6):066005.
248. Dorman CJ (2013) Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. *Biochemical Society Transactions* 41(2):542-547.
249. Ferrándiz M-J, *et al.* (2016) An increase in negative supercoiling in bacteria reveals topology-reacting gene clusters and a homeostatic response mediated by the DNA topoisomerase I gene. *Nucleic Acids Research* 44(15):7292-7303.
250. Kouzine F, *et al.* (2013) Transcription-dependent dynamic supercoiling is a short-range genomic force. *Nature Structural & Molecular Biology* 20(3):396-403.
251. Naughton C, *et al.* (2013) Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature Structural & Molecular Biology* 20(3):387-395.
252. Hsieh LS, Rouviere-Yaniv J, & Drlica K (1991) Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. *Journal of Bacteriology* 173(12):3914-3917.
253. van Workum M, *et al.* (1996) DNA supercoiling depends on the phosphorylation potential in *Escherichia coli*. *Molecular Microbiology* 20(2):351-360.

254. Broyles SS & Pettijohn DE (1986) Interaction of the Escherichia coli HU protein with DNA: Evidence for formation of nucleosome-like structures with altered DNA helical pitch. *Journal of Molecular Biology* 187(1):47-60.
255. Rouvière-Yaniv J, Yaniv M, & Germond J-E (1979) E. coli DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA. *Cell* 17(2):265-274.
256. Rovinskiy N, Agbleke AA, Chesnokova O, Pang Z, & Higgins NP (2012) Rates of Gyrase Supercoiling and Transcription Elongation Control Supercoil Density in a Bacterial Chromosome. *PLoS Genet* 8(8):e1002845.
257. Norregaard K, *et al.* (2013) DNA supercoiling enhances cooperativity and efficiency of an epigenetic switch. *Proceedings of the National Academy of Sciences*.
258. Lia G, *et al.* (2003) Supercoiling and denaturation in Gal repressor/heat unstable nucleoid protein (HU)-mediated DNA looping. *Proceedings of the National Academy of Sciences* 100(20):11373-11377.
259. Leng F, Chen B, & Dunlap DD (2011) Dividing a supercoiled DNA molecule into two independent topological domains. *Proceedings of the National Academy of Sciences* 108(50):19973-19978.
260. Du Q, Smith C, Shiffeldrim N, Vologodskaia M, & Vologodskii A (2005) Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proceedings of the National Academy of Sciences* 102(15):5397-5402.
261. Becker NA, Kahn JD, & Maher LJ (2007) Effects of nucleoid proteins on DNA repression loop formation in Escherichia coli. *Nucleic Acids Research* 35(12):3988-4000.
262. Johnson S, Lindén M, & Phillips R (2012) Sequence dependence of transcription factor-mediated DNA looping. *Nucleic Acids Research* 40(16):7728-7738.
263. Mogil LS, Becker NA, & Maher LJ, III (2016) Supercoiling Effects on Short-Range DNA Looping in E. coli. *PLOS ONE* 11(10):e0165306.
264. Chen YF, Milstein JN, & Meiners JC (2010) Femtonewton Entropic Forces Can Control the Formation of Protein-Mediated DNA Loops. *Physical Review Letters* 104(4).
265. Forde NR, Izhaky D, Woodcock GR, Wuite GJ, & Bustamante C (2002) Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by Escherichia coli RNA polymerase. *Proc. Natl. Acad. Sci. USA* 99(18):11682-11687.
266. Whitson PA, Hsieh WT, Wells RD, & Matthews KS (1987) Influence Of Supercoiling And Sequence Context On Operator Dna-Binding With Lac Repressor. *Journal of Biological Chemistry* 262(30):14592-14599.
267. Yan Y, Leng F, Finzi L, & Dunlap D (2018) Protein-mediated looping of DNA under tension requires supercoiling. *Nucleic Acids Research* 46(5):2370-2379.
268. Wu HM & Crothers DM (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature* 308(5959):509-513.
269. Matsumoto K & Hirose S (2004) Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of *Drosophila*. *Journal of Cell Science* 117(17):3797-3805.
270. Travers A & Muskhelishvili G (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth? *Nature Reviews Microbiology* 3(2):157-169.
271. Blumberg S, Pennington MW, & Meiners JC (2006) Do femtonewton forces affect genetic function? A review. *Journal of biological physics* 32(2):73-95.
272. Chen Y-F, Milstein JN, & Meiners J-C (2010) Protein-Mediated DNA Loop Formation and Breakdown in a Fluctuating Environment. *Physical Review Letters* 104(25):258103.
273. van Loenhout MTJ, de Grunt MV, & Dekker C (2012) Dynamics of DNA Supercoils. *Science* 338(6103):94-97.

274. Forth S, *et al.* (2008) Abrupt Buckling Transition Observed during the Plectoneme Formation of Individual DNA Molecules. *Physical Review Letters* 100(14):148301.
275. Saiz L & Vilar JMG (2006) DNA looping: the consequences and its control. *Current Opinion in Structural Biology* 16(3):344-350.
276. Le TBK, Imakaev MV, Mirny LA, & Laub MT (2013) High-Resolution Mapping of the Spatial Organization of a Bacterial Chromosome. *Science*.
277. Dixon Jesse R, Gorkin David U, & Ren B (2016) Chromatin Domains: The Unit of Chromosome Organization. *Molecular Cell* 62(5):668-680.
278. Smith DE, *et al.* (2001) The bacteriophage phi 29 portal motor can package DNA against a large internal force. *Nature* 413(6857):748-752.
279. Tajik A, *et al.* (2016) Transcription upregulation via force-induced direct stretching of chromatin. *Nat. Mater.* 15(12):1287-1296.
280. Badrinarayanan A, Le TBK, & Laub MT (2015) Bacterial Chromosome Organization and Segregation. *Annual Review of Cell and Developmental Biology* 31(1):171-199.
281. Higgins NP (2016) Species-specific supercoil dynamics of the bacterial nucleoid. *Biophysical reviews* 8(Suppl 1):113-121.
282. Sobetzko P (2016) Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. *Nucleic Acids Research* 44(4):1514-1524.
283. Efremov AK, Winardhi RS, & Yan J (2016) Transfer-matrix calculations of DNA polymer micromechanics under tension and torque constraints. *Physical Review E* 94(3):032404.
284. Marko JF & Neukirch S (2013) Global force-torque phase diagram for the DNA double helix: Structural transitions, triple points, and collapsed plectonemes. *Physical Review E* 88(6):062722.
285. Whitson PA, Hsieh WT, Wells RD, & Matthews KS (1987) Supercoiling Facilitates Lac Operator-Repressor-Pseudooperator Interactions. *Journal Of Biological Chemistry* 262(11):4943-4946.
286. Uuskula-Reimand L, *et al.* (2016) Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biology* 17(1):182.
287. Benedetti F, Racko D, Dorier J, Burnier Y, & Stasiak A (2017) Transcription-induced supercoiling explains formation of self-interacting chromatin domains in *S. pombe*. *Nucleic Acids Research* 45(17):9850-9859.
288. Ferrandiz MJ, Martin-Galiano AJ, Schvartzman JB, & de la Campa AG (2010) The genome of *Streptococcus pneumoniae* is organized in topology-reacting gene clusters. *Nucleic Acids Research* 38(11):3570-3581.
289. Adhya S (1989) Multipartite Genetic-control Elements - Communication by DNA Loop. *Annual Review of Genetics* 23:227-250.
290. Finzi L & Dunlap DD (2010) Single-molecule approaches to probe the structure, kinetics, and thermodynamics of nucleoprotein complexes that regulate transcription. *Journal of Biological Chemistry* 285(25):18973-18978.
291. Mirkin SM (2001) DNA Topology: Fundamentals. in *Encyclopedia of Life Sciences* (John Wiley & Sons, Ltd).
292. Charvin G, Allemand JF, Strick TR, Bensimon D, & Croquette V (2004) Twisting DNA: single molecule studies. *Contemporary Physics* 45(5):383-403.
293. Lipfert J, *et al.* (2014) Double-stranded RNA under force and torque: Similarities to and striking differences from double-stranded DNA. *Proceedings of the National Academy of Sciences* 111(43):15408-15413.
294. Neukirch S & Marko JF (2011) Analytical description of extension, torque, and supercoiling radius of a stretched twisted DNA. *Physical Review Letters* 106(13):138104.

295. Seol Y & Neuman KC (2011) Magnetic tweezers for single-molecule manipulation. *Methods in Molecular Biology* 783:265-293.
296. Yin H, Landick R, & Gelles J (1994) Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophysical Journal* 67(6):2468-2478.
297. Benedetti F, Dorier J, & Stasiak A (2014) Effects of supercoiling on enhancer–promoter contacts. *Nucleic Acids Research* 42(16):10425-10432.
298. Muskhelishvili G & Travers A (2016) The regulatory role of DNA supercoiling in nucleoprotein complex assembly and genetic activity. *Biophys Rev* 8(Suppl 1):5-22.
299. Hodges C, Bintu L, Lubkowska L, Kashlev M, & Bustamante C (2009) Nucleosomal Fluctuations Govern the Transcription Dynamics of RNA Polymerase II. *Science* 325(5940):626-628.
300. Bintu L, *et al.* (2011) The elongation rate of RNA polymerase determines the fate of transcribed nucleosomes. *Nat Struct Mol Biol* 18(12):1394-1399.
301. Bintu L, *et al.* (2012) Nucleosomal Elements that Control the Topography of the Barrier to Transcription. *Cell* 151(4):738-749.
302. Jin J, *et al.* (2010) Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. *Nat Struct Mol Biol* 17(6):745-752.
303. Zhang T, Cooper S, & Brockdorff N (2015) The interplay of histone modifications – writers that read. *EMBO reports* 16(11):1467-1481.
304. Long Hannah K, Prescott Sara L, & Wysocka J (2016) Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* 167(5):1170-1187.
305. Mendes MA, *et al.* (2013) MADS Domain Transcription Factors Mediate Short-Range DNA Looping That Is Essential for Target Gene Expression in Arabidopsis. *The Plant Cell* 25(7):2560-2572.
306. Hsieh WT, Whitson PA, Matthews KS, & Wells RD (1987) Influence of sequence and distance between two operators on interaction with the lac repressor. *Journal of Biological Chemistry* 262(30):14583-14591.
307. Pfahl M, Gulde V, & Bourgeois S (1979) “Second” and “third operator” of the lac operon: An investigation of their role in the regulatory mechanism. *Journal of Molecular Biology* 127(3):339-344.
308. Garcia HG & Phillips R (2011) Quantitative dissection of the simple repression input-output function. *Proc Natl Acad Sci U S A* 108(29):12173-12178.
309. Sadler JR, Sasmor H, & Betz JL (1983) A perfectly symmetric lac operator binds the lac repressor very tightly. *Proc Natl Acad Sci U S A* 80(22):6785-6789.
310. Oehler S & Müller-Hill B (2010) High Local Concentration: A Fundamental Strategy of Life. *Journal of Molecular Biology* 395(2):242-253.
311. Oehler S, Amouyal M, Kolkhof P, von Wilcken-Bergmann B, & Muller-Hill B (1994) Quality and position of the three lac operators of E. coli define efficiency of repression. *Embo J* 13(14):3348-3355.
312. Deuschle U, Gentz R, & Bujard H (1986) lac Repressor blocks transcribing RNA polymerase and terminates transcription. *Proc Natl Acad Sci U S A* 83(12):4134-4137.
313. Hao N, *et al.* (2014) Road rules for traffic on DNA-systematic analysis of transcriptional roadblocking in vivo. *Nucleic Acids Res* 42(14):8861-8872.
314. Deuschle U, Hipskind RA, & Bujard H (1990) RNA polymerase II transcription blocked by Escherichia coli lac repressor. *Science* 248(4954):480-483.



315. Reines D & Mote J, Jr. (1993) Elongation factor SII-dependent transcription by RNA polymerase II through a sequence-specific DNA-binding protein. *Proc Natl Acad Sci U S A* 90(5):1917-1921.
316. Forde NR, Izhaky D, Woodcock GR, Wuite GJ, & Bustamante C (2002) Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by *Escherichia coli* RNA polymerase. *Proc Natl Acad Sci U S A* 99(18):11682-11687.
317. Lee JB, *et al.* (2006) DNA primase acts as a molecular brake in DNA replication. *Nature* 439(7076):621-624.