Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a 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 in whole or in part in all forms of media, now or hereafter now, 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. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Allison G. Cartee   April 10, 2023

RNA Polymerase Hybrid Passage through Roadblocks

by

Allison G. Cartee

Dr. Laura Finzi
Adviser

Physics

Dr. Laura Finzi

Adviser

Dr. David Dunlap

Committee Member

Dr. Effrosyni Seitaridou

Committee Member

Dr. Fang Chen

Committee Member

2023

RNA Polymerase Hybrid Transit through Roadblocks

By

Allison G. Cartee

Dr. Laura Finzi

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Physics

2023

Abstract

RNA Polymerase Hybrid Transit through Roadblocks

By Allison G. Cartee

The synthesis of messenger ribonucleic acid (mRNA) from deoxynucleic acid (DNA) by RNA

polymerase (RNAP) decodes genetic information. Motor enzyme RNAP translocates at up to 50

base pairs per second *in vivo*, but forces and transcription factors can modulate activity and

pausing, both critical for regulation. Previous studies suggest RNAP may backtrack after intrinsic

pausing, which may prolong the inactive state, upon encountering a physical obstacle bound to

DNA. Yet, the mechanism by which RNAP overcomes these roadblocks to ensure genetic

expression is poorly understood. It is uncertain whether RNAP actively disperses a roadblock, or

passively waits for its dissociation. We collected *E. Coli* RNAP pause times at roadblocks LacI

bound at sites O1, O2, and Os, in decreasing affinity, then EcoR1. Pauses were measured as a

function of forces opposing or assisting RNAP translocation via magnetic tweezers and in the

presence of GreA, a protein that rescues backtracked RNAPs by cleaving nascent mRNA backed

up into RNAP's catalytic site. Regardless of magnitude, forces opposing RNAP at LacI obstacles

increased average pause durations compared to assisting forces without GreA. Including GreA

eliminated this increase in pause time for LacI-O1 and LacI-O2. We speculate opposing forces

may promote backtracking since GreA decreased opposing force average pause times to an

assisting force baseline independent of GreA. Thus, backtracking may not be critical for RNAP to

overcome relatively weaker obstacles. Though LacI-Os demonstrated a similar assisting force

baseline independent of GreA, adding GreA to opposing force conditions lowered average

pause times beneath the assisting force baseline. Repetitive cycles of backtrack and recovery may help RNAP overcome relatively stronger obstacles. To control for any LacI-RNAP co-immunoprecipitation, we performed experiments with an inactive form of EcoR1 since this enzyme is not known to interact with RNAP. We observed similar read-through proportions as with LacI-Os. We propose that RNAP may use two different transit paths to overcome roadblocks of different relative strengths: an active pathway to dislodge stronger proteins and a passive one to wait for the dissociation of weaker proteins. Our biomechanical measurements elucidate how forces on the genome may affect RNAP behavior at roadblocks.

RNA Polymerase Hybrid Transit through Roadblocks

By

Allison G. Cartee

Dr. Laura Finzi

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Physics

2023

Acknowledgements

Foremost, I would like to thank Dr. Laura Finzi and Dr. David Dunlap for the opportunity to contribute to their work and for opening my eyes to the beauty of biophysical research. I would like to thank my graduate student mentor, Jin Qian, for patiently fostering my bench side skills, scientific problem solving, and writing through his time, assistance, and wisdom. I sincerely appreciate the support of lab members Dr. Dylan Collette, Narae Kim, Dr. Wenxuan Xu, Derrica McCalla, Jessie Doan, Anna Eligulashvili, and Joe Piccolo. I deeply appreciate Ms. Susan Cook and Ms. Barbara Connor for supporting and coordinating my activities within the Emory Physics Department.

I would like to thank Dr. Fang Chen who inspired me to share my scientific perspectives with the world. To Dr. Effrosyni Seitaridou, thank you for giving me the foundations to become a biophysical researcher and for inspiring me to pursue excellence wherever I go.

Table of Contents

List of Figures

1 Chapter 1: Introduction

1.1 Research Motivation

All living organisms exhibit the central dogma of biology, the transmission and expression of genetic information (Fig. 1). The motor enzyme ribonucleic acid polymerase (RNAP) initiates gene expression by converting deoxyribonucleic acid (DNA) into messenger RNA (mRNA) in a process called transcription. Subsequent, ribosomal translation of mRNA ensures the faithful expression of proteins fundamental to an organism's functions.
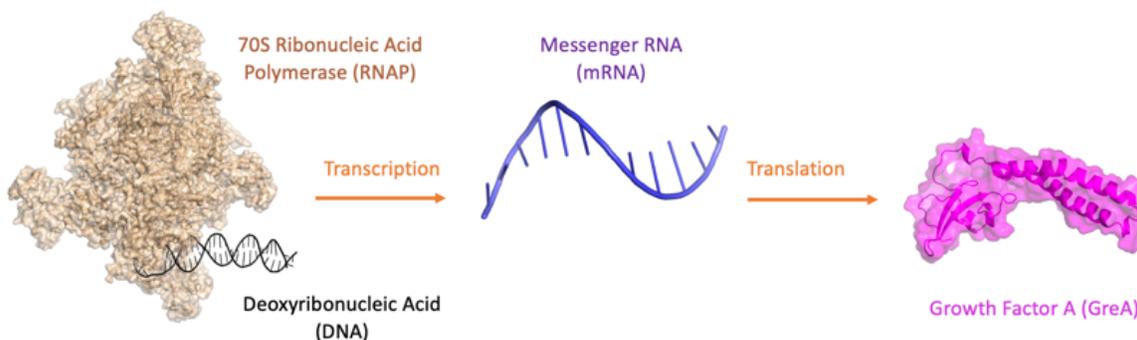
Figure 1. Central Dogma of Biology. RNAP (PDB: 4JKR) initiates the central dogma of biology by transcribing DNA into mRNA (PDB: 1BNA) for subsequent translation by ribosomes into biologically significant protein products, such as growth factor A (GreA) (PDB: 1GRJ). Biological molecules not to scale.

In vivo, transcription is a highly regulated process. RNAP often encounters physical protein roadblocks obstructing its path along DNA. For instance, eukaryotic RNAP often encounters nucleosomes, while prokaryotic RNAP encounters regulatory proteins such as *lac* or *186 Cl* repressors (Fig. 2).[1,2] Previous studies suggest RNAP can either pause and or slightly move backwards, termed backtracking, at these roadblock sites.[3] Though these pauses likely

coordinate mRNA translation into biologically significant protein products, we do not fully understand the mechanism by which RNAP overcomes roadblocks.[4]

As a result, we sought to determine how this motor enzyme overcomes various roadblocks by measuring the effect of transcriptional force for RNAP to transcribe genes. We hypothesized that comparing the transcriptional progress by RNAP through different roadblocks might shed light on these mechanism(s). This study's novel methods modulated RNAP transcriptional force as well as the presence of an endonuclease enzyme known to promote transcription through other roadblocks.[3] Results suggest RNAPs may utilize two different mechanisms depending on the strength of the roadblock obstacle. In an active pathway, RNAPs use repetitive cycles of backtrack and recovery to dislodge tightly bound roadblocks. In a passive pathway, the most efficient means of transit may involve RNAP remaining in an active conformation without backtracking for a weaker roadblock to dissociate before proceeding. Overall, elucidating the behavior of RNAP at roadblocks will improve models of transcription and provide a more holistic understanding of gene expression.
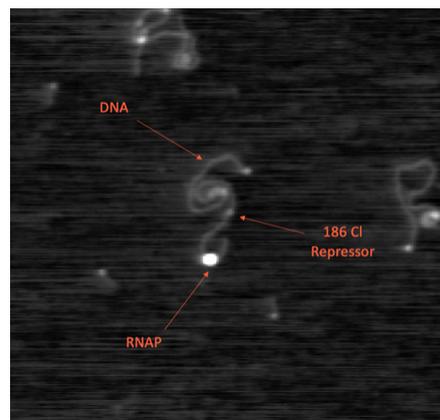
Figure 2. Atomic Force Microscopy (AFM) Scan of Roadblocks. An AFM image of RNAP interrupted before encountering a 186 Cl repressor roadblock (unpublished).

Transcription, the process by which organisms initiate then finely tune genetic expression, often involves interplay between multiple factors. At the most fundamental level, the DNA sequence itself can modulate RNAP transcription through two documented methods. First, regulatory sequences called enhancers may increase transcriptional levels.[5] Or, a native DNA sequence may help transcription past obstacles.[6] If roadblock that singly-binds DNA is added, transcription through this linear obstacle can depend on how tightly roadblock binds DNA (k_d) within the Hill function.[7] Indeed, roadblocks bound to stronger sites impede transcription more than roadblocks bound to weaker sites.[1] Even in this relatively simple model, bacteria utilize over 400 of these roadblock systems to regulate genetic expression.[8] Furthermore, roadblocks can bind DNA at multiple sites and induce topologies such as loops or wraps that influence transit more than just individual k_d's alone.[1,2,6]

In addition to DNA sequences or roadblocks intrinsic to genome organization, transcribing RNAPs *in vivo* may encounter other transcribing RNAPs along the same DNA strand. If translocating in opposite directions, the collision between RNAPs can terminate transcription.[9] However, if collisions occur in the same direction, the additional force exerted by trailing RNAPs colliding with the leading RNAP may help the passage of the leading RNAP through roadblock obstacles.[3] Though such studies posit an effect of force on transcription, individual collisions may slightly vary, and their kinetics cannot be quantified. Salt may also affect transcription, though such studies only explore the open conformation, before RNAP begins synthesizing mRNA, or the closed conformation,

where RNAP terminates transcription fully.[10] Consequently, the effects of both force and salt on RNAPs transcribing against roadblocks have yet to be fully investigated.

Currently, the process by which RNAPs overcome protein obstacles is poorly understood. Previous studies document RNAP transcription through this study's roadblocks of interest, but these studies' results only capture snapshots of RNAPs' position through roadblocks rather than the kinetics of transit.[3,11] This thesis builds upon past work by delivering pico-Newton (pN) forces to alter RNAP's intrinsic transcriptional force while continually measuring its location as they approach near, pause at, then read through roadblocks in real time. In particular, this study demonstrates increasing RNAPs' intrinsic transcriptional force by even an additional 0.2 pN can effectively prevent backtracking entirely while decreasing it by the same amount can promote backtracking. Even up to 5 pN, this force-biased finding depends solely on the direction of force, rather than its magnitude. The sensitivity of results with respect to direction but independence with respect to magnitude compel us to elucidate the dynamics and a formal mechanism of RNAP transcriptional pausing at roadblock proteins.

1.2   Nucleic Acid Structure and Mechanics

The polymer DNA contains genetic information across all life within the sequence of its nitrogenous bases: adenosine (A), cytosine (C), guanine (G), and thymine (T). Nucleotide monomers of DNA consist of a 2-deoxyribose sugar with a nitrogenous base attached to the 1' carbon, a phosphate group attached to the 3' carbon, and a hydroxyl attached to the 5' carbon.

The primary structure of DNA is defined by covalent bonds linking nucleotides through 5' to 3' phosphodiester bonds (Fig. 3). Secondary structure involves Watson-Crick base pairing between two single stranded DNA (ssDNA) molecules to form a doubled strand DNA (dsDNA). This base pairing consists of hydrogen-bonding between nucleotides A T and C G to connect two ssDNA in an anti-parallel manner as a single dsDNA (Fig. 3). DNA's primary and secondary structures contribute to the molecule's stability and characteristic, double-helical tertiary structure.
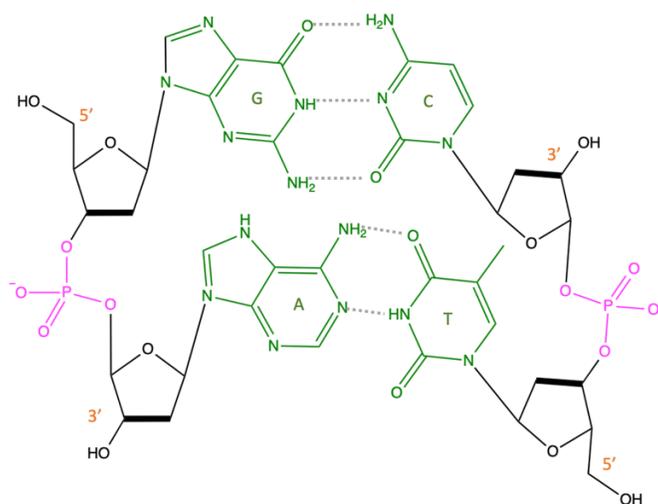
Figure 3. DNA Structure and Base Pairing. Covalently bonded primary structure and hydrogen bonded secondary structure between complementary nitrogenous bases. The primary structure of DNA consists of phosphodiester bonds linking 2-deoxyribose sugars in the 5' to 3' direction. Deoxyribose sugars are shown in black, labeled nitrogenous bases in green, and phosphate groups in pink.

The DNA double helix shields base pairs internally and positions the negatively charged phosphodiester backbones outwards. Inside the helix, the unique pairing of double-ringed purines, A & G, with single-ringed pyrimidines, T & C, permit a constant 2 nanometer (nm)

diameter (Fig. 4). *In vivo* B-DNA forms a right-handed helix with a pitch of 10.4 base pairs (bp) or 3.4 nm per turn (Fig. 4). However, DNA's pitch may deviate from this commonly observed, standard value in genetically silenced DNA or when RNAP exerts transcriptional forces.[12] Although DNA can accommodate some torsional stress, twisting beyond a certain threshold induces plectonemes, flat or planar loops that can compact DNA.[13] Plectonemic formation, however, only occurs for molecules 10 kilo base pair (kbp) or longer, outside the consideration of this study's 4 kbp molecules.[14] These structural properties of DNA influence its function within organisms. For example, bacterial DNA is naturally negatively supercoiled–twisted or coiled in the clockwise direction of the double-helix looking down the double helix depicted in figure 4–to promote genome compaction and organization.[14]
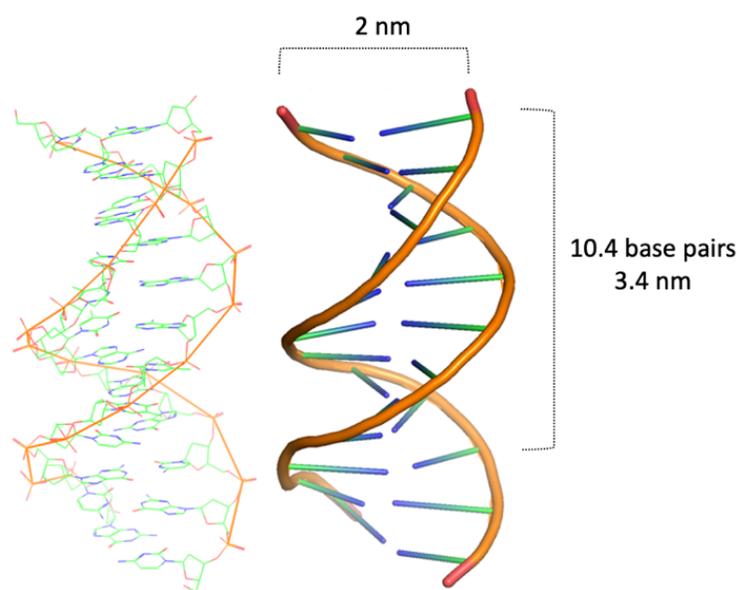
Figure 4. DNA Tertiary Structure. The tertiary structure of B-DNA (PDM: 1BNA) is a right-handed double helix shown as a wire and cartoon representation for the same sequence.

In addition to DNA, cells utilize RNA, another nucleic acid with a few key differences. Unlike

helical dsDNA, RNA is typically single stranded. Also, whereas DNA includes a 2-deoxyribose

sugar as shown in figure 3, RNA possesses a more reactive ribose sugar with a 2' hydroxyl group

(Fig. 5). Finally, instead of thymine, RNAP incorporates another pyrimidine, uracil (U), a

nitrogenous base that lacks thymine's 5' methyl group (Fig. 5). The genetic code, however, is

preserved despite this slight structural modification since uracil also base pairs with adenosine.

Although mRNA is rather transient genetic information that does not serve for long-term

storage and transmission, it is vital to cellular function. Furthermore, there are other essential

types of RNA such as transfer RNA which positions amino acids in translation, and micro-RNA

which controls post-translational protein regulation.

Thymine Nucleotide   Uracil Nucleotide
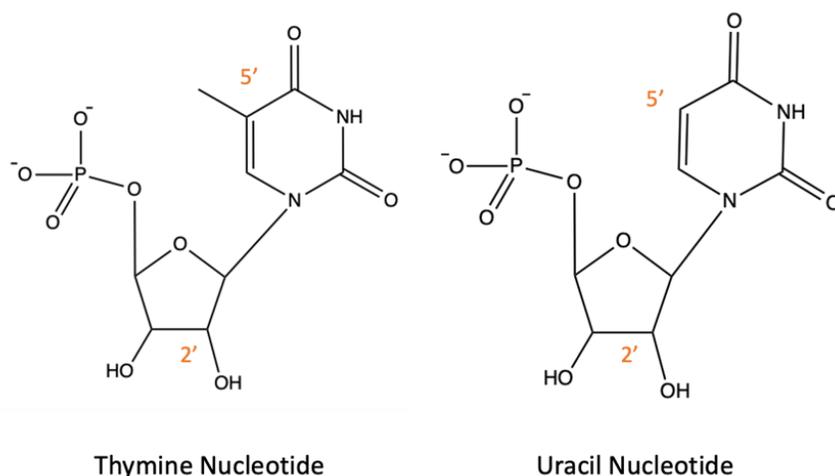
Figure 5. Thymine versus Uracil Nucleotides. The uracil ribonucleotide lacks a 5' methyl found in
thymine and has a 2' ribose hydroxyl group atypical of deoxyribonucleotides.

Though all organisms share DNA and RNA, eukaryotes generally have larger genomes. As a

result, eukaryotes store linear DNA negatively supercoiled around positively charged histone

proteins within a nucleus.[2] Prokaryotes such as our model organism *Escherichia coli* (*E. Coli*)

store DNA in circular plasmids containing approximately 4.6 million base pairs under less ambient torsional stress than their eukaryotic counterparts.[14] Moreover, prokaryotic DNA involves fewer regulatory checkpoints, smaller polymerase proteins, and less molecular machinery. This study investigates prokaryotic transcription through roadblocks with specially prepared *E. Coli* RNAPs, roadblocks, and linear segments from plasmid chromosomes.

1.3 RNA Polymerase and Transcription

The transcription of DNA into mRNA is the first step of genetic expression and consists of three generalizable steps: initiation, elongation, and termination. First, an RNAP attaches to DNA at a promoter sequence (Fig. 6A). RNAP may recruit various transcriptional factors or other proteins to initiate or enhance transcription. After assembling the necessary components at the promoter, RNAP separates 10-15 base pairs of dsDNA into two ssDNA strands.[15] Then, it transcribes approximately 12 bps before switching to elongation and moving along a template strand of DNA in the 3' to 5' direction.[15] It simultaneously synthesizes a strand of mRNA, 5' to 3', complementary to the DNA template using free ribonucleotide triphosphates in the cellular environment. During elongation, the nascent mRNA polymerizes at the active site and emerges from the RNA exit channel of the enzyme into the cellular environment (Fig. 6B). The transcriptional process terminates when RNAP either reaches an intrinsic terminator with many uracil residues or a protein factor assists in termination (Fig. 6C).[15] The RNAP and mRNA dissociate from the DNA, and the translation of mRNA by ribosomes into biologically significant protein products completes genetic expression.
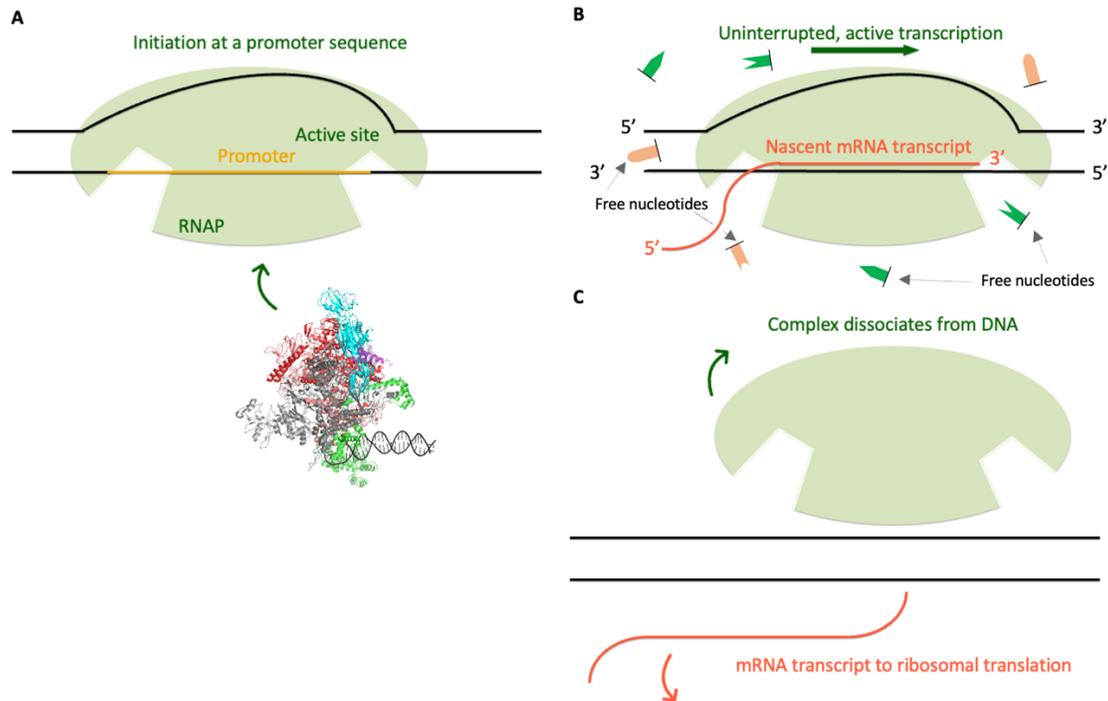
Figure 6. Transcription Initiation, Elongation, and Termination. Transcription involves three steps: initiation (A), elongation (B), and termination (C). (A) RNAP (PDB: 4JKR) searches for and binds to a promoter sequence in the DNA. (B) RNAP translocates forward in the 3' to 5' direction, synthesizing a strand of mRNA in the 5' to 3' direction with free nucleotides in solution. (C) RNAP and mRNA dissociate from the DNA for other processes to take over genetic expression. Molecules not depicted to scale.

Because the structures of eukaryotic and prokaryotic genomes differ, their RNAPs and various transcriptional factors are also different. This study's model organism *E. Coli* has a core polymerase enzyme that consists of 5 core protein subunits and requires a σ-factor coenzyme (Fig. 6B).[15,16] Together, these components form the holoenzyme, which can search for promoters. Previous studies estimate that *E. Coli* RNAP translocates with 20-25 pN of force

exerted by the motor enzyme on the DNA and at approximately 50 bp per second *in vivo*, but such velocity is sequence dependent.[15,17]

1.4   LacI and EcoR1 (Gln111) Roadblock Structure and Mechanics

In vivo transcription is a highly regulated part of gene expression. During elongation, RNAP may encounter many regulatory proteins obstructing its forward translocation along a DNA strand.[18] One widely studied prokaryotic example is the *lac* repressor protein (LacI), which regulates genes for sugar metabolism (Fig. 7A, B). This homo-tetramer with two DNA-binding sites can bind to either one or two high affinity sites along a DNA molecule, including Os (K_d = 0.01 nM), O1 (K_d = 0.13 nM), and O2 (K_d = 0.6 nM) (Fig. 7A, B).[1,19]

Another roadblock explored in addition to LacI was the *E. Coli* restriction endonuclease I (EcoR1) Q111 mutant. This endonuclease has a high affinity for the DNA sequence 5'-GAATTC-3' where it naturally cleaves the phosphodiester bonds of both strands in the double helix to produce ends with 5'-AATT overhangs. *In vivo,* restriction endonucleases cleave DNA sequences not found in the genome of the host organism to eliminate foreign DNA. *In vitro*, restriction digest reactions allow scientists to cut the DNA into blunt or single-stranded ends that can hybridize to join different DNA sequences.[20] To avoid cutting dsDNA here, this study's mutant form of EcoR1 recognized and bound to its palindromic binding sequence, 5'-GAATTC-3', with high affinity but left the dsDNA intact.[20] EcoR1 served as a non-interacting roadblock comparison to compare with LacI-O1, LacI-O2*,* and LacI-Os, since no studies documented an

interaction between RNAP and EcoR1: collaborator Dr. Irina Artsimovich demonstrated a LacI-

RNAP co-immunoprecipitation (unpublished; manuscript in preparation).

A.   **B.**   **C.**

Linear LacI-O1 (K_d = 0.13 nM)   Lopped LacI-Os   Linear EcoR1 Q111 *E. Coli* restriction
or LacI-O2 (K_d = 0.6 nM)   (K_d = 0.01 nM)   endonuclease I (EcoR1)
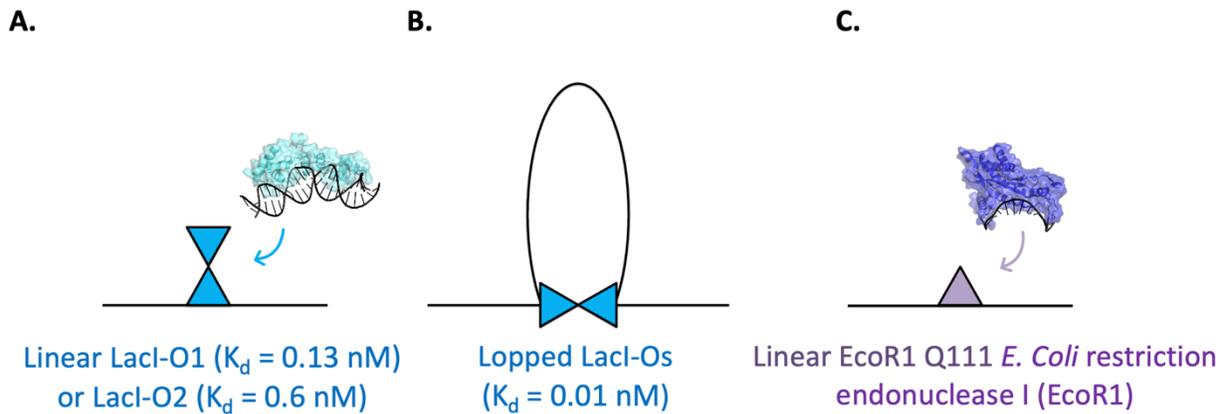
Figure 7. Binding of LacI and EcoR1 Roadblocks. (A) LacI (PDB: 1L1M) binds to single LacI-O1 or O2

sites along a linear DNA template. (B) A single LacI can simultaneously bind to two LacI-Os sites

producing a looped conformation. (C) EcoR1 (PDB: 2OXV) binds to the sequence 5'-GAATTC-3' of a

linear DNA template. Molecules not depicted to scale.

1.5   GreA Purpose

Protein transcription factors can modulate RNAP transcriptional activity. One such elongation

factor, growth factor A (GreA), has an endonuclease function that can cleave mRNA. Previous

studies demonstrate that GreA acts within the active site of RNAP to cleave backtracked

mRNA.[21] During elongation, RNAP may encounter obstacles and or sequences that induce

backtracking.[15] Backtracking pushes nascent mRNA forward into the primary channel, and

complexes stall because of the displacement between the 3' end of the nascent mRNA and

active site (Fig. 8A).[15] Curiously, introducing GreA mid-transcription may decrease how often

RNAP pauses but may also increase how far it backtracks.[22] However, past experiments

established adding GreA to backtracked RNAP could quickly rescue forward translocation at an

EcoR1 roadblock (Fig. 8 B, C).[21] Tracking the transcriptional progress of RNAP upon

encountering EcoR1 demonstrated that stalled complexes could now traverse roadblocks in the

presence of GreA (Fig. 8D).
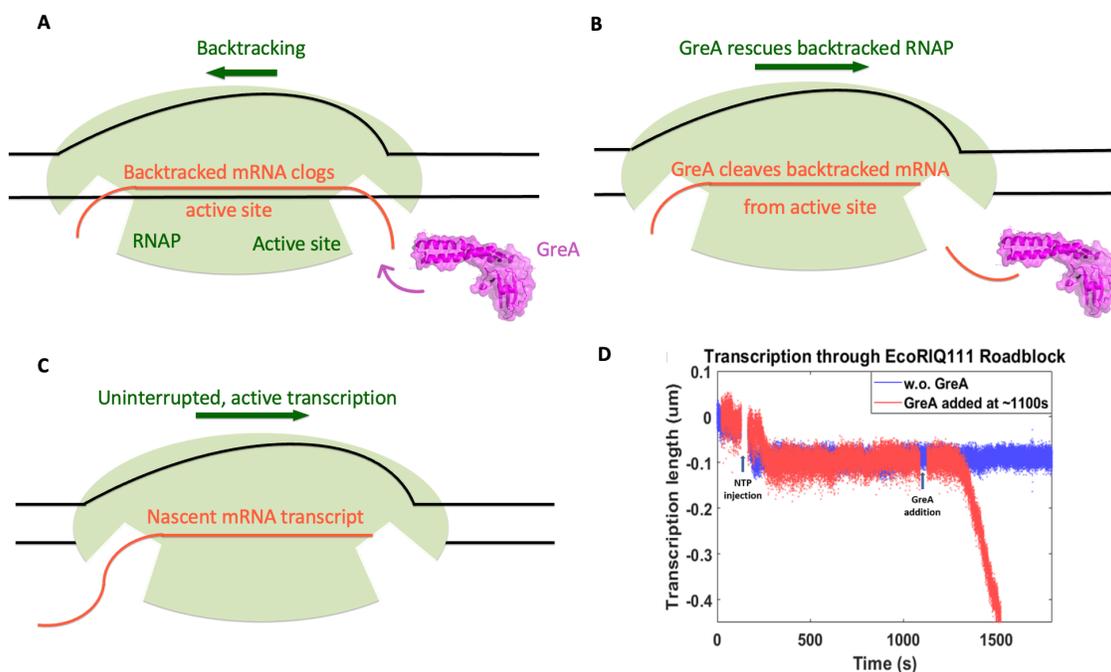
Figure 8. GreA Rescue of RNAP. (A) As RNAP backtracks, nascent mRNA protrudes into the

active site. (B) GreA cleaves mRNA protruding into the active site. (C) RNAP resumes

forward transcription. (D) RNAP translocates past EcoR1, as confirmed by the expected

height of the roadblock in solution (-0.1 μm), upon the addition of GreA (PDB: 1GRJ).[23]

Molecules not depicted to scale.

1.6   Single Molecule Approaches

The objective of this study was to observe the passage of RNAP through roadblocks using single

molecule techniques. Single-molecule techniques reveal the reaction trajectories of individual

molecules, as opposed to those of large ensembles used in bulk studies.[15] As a result, single-molecule techniques capture micro-scale kinetics, thermodynamics, and structures that might otherwise be lost in averages of population heterogeneity. To achieve such precision, experiments involve mechanical manipulation of individual molecules. Examples of these techniques include but are not limited to atomic force microscopy, optical trapping, and magnetic tweezer studies.[12] However, to determine whether findings are representative of a natural phenomenon, single-molecule studies require many assays, which can prove expensive or time-consuming. Moreover, single-molecule systems are quite sensitive to slight mechanical or chemical perturbations and generally employ reactant concentrations not found *in vivo*.[15] Nonetheless, single-molecule techniques are uniquely suited to investigate biomolecular interactions involving pico-newton scale transcriptional forces.

2   Chapter 2: Magnetic Tweezers (MT)

2.1   The MT Technique

Magnetic tweezers (MTs) are a powerful single-molecule manipulation tool with which to produce and detect real-time changes in the structure of tethered polymers to study thermodynamics or kinetics.[24] Overhead neodymium magnets generate a magnetic field ($\vec{B}$) that can exert upwards forces ranging from 10 fN to 100 pN on paramagnetic beads tethered to individual biopolymers without physical contact that might otherwise disrupt these delicate systems.[24] Rotating and raising or lowering magnets generates torques ($\vec{\tau}$) and modulates tensions on the tether holding the magnetic bead via upwards magnetic forces ($\overrightarrow{F_B}$).

A general MT schema uses rotating and vertically displacing magnets to exert torques and attractive forces on magnetic beads (Fig. 9). Tethered micron-diameter beads in solution experience simple Brownian motion, exhibited as stochastic movement.[25] In the absence of external fields, the range of Brownian motion is limited by the length of the polymer tether about the anchor point.
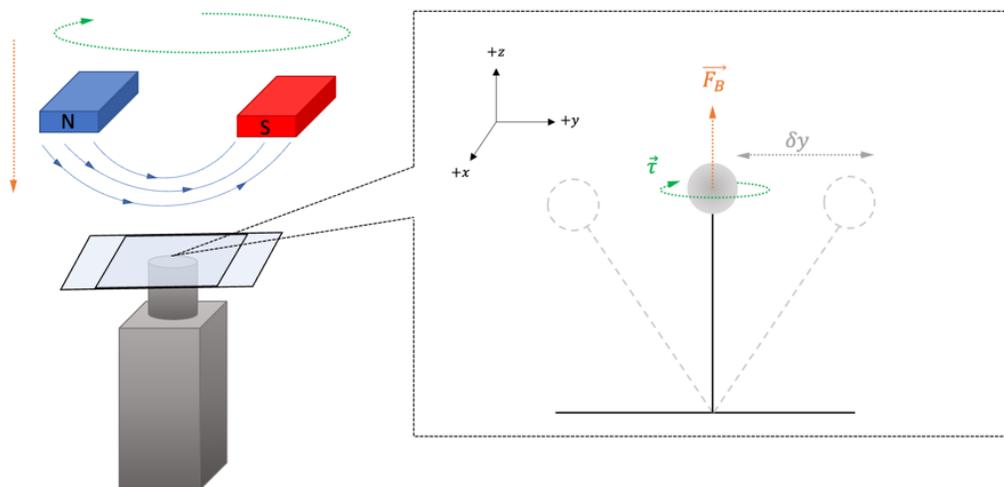
Figure 9. MT Schematic Illustration and Range of Motion. A tethered magnetic bead exhibits Brownian motion in all three directions, though only the +y-direction (δy) is shown as one example. Lowering magnets towards the sample generates increasingly larger attractive force while rotating the magnets rotates the magnetically coupled bead and exerts torque on the polymer tether. At the bottom, an objective and charged-coupled device (CCD) visualize beads within a fluid filled, glass chamber.

In experiments, an LED is used to illuminate beads in a field of view to produce radial diffraction patterns that appear as concentric circles. These patterns are visualized using a charge-coupled device (CCD) camera placed beneath samples and are tracked in real time (Fig. 10 A-C). To establish the *xyz*-position of a bead in each frame of a video sequence, the MT first tracks and averages the *xy*-Brownian motion of a bead. This average estimate of the anchor point serves as an origin for the analysis of motion due to the external field. Obtaining the *z*-position of beads, however, depends on the intensity variations and extent of the diffraction patterns relative to a look-up-table (LUT) calibration (Fig. 10D). LUT calibrations were recorded for each bead prior to experimentation and depict the *xy*-radial patterns at different focal distances as the focus of the objective is changed in *z*-space. Although figure 10D displays one experiment's radial profiles over a user-defined range of 49 to 59 micrometers (μm), only the relative displacements from a focal plane were considered.

Previous studies have characterized how radial patterns change as a bead moves out of focus further away from the Z=0 focal plane (Fig. 10).[24] Using an overfocused setting, beads stuck to

the bottom of the microchamber appear larger, fuzzier, and with more rings (Fig. 10C). These

beads close to the bottom of the chamber were selected as reference beads in the plane $Z = 0$.

During experiments, the magnetic field or other factors like tether length might change the

range of beads' 3D displacements. The *xy*-displacements are easily tracked within the field of

view. Establishing the *z*-displacement requires determination of the difference between the

heights of a user-selected reference bead and an experimental bead. Prior to beginning an

experiment, the focus is progressively changed by moving the objective through a known range

along the *z*-axis and registering the radial diffraction patterns with respect to objective height in

a calibration plot (Fig. 10D). Beads labeled A-C were given an estimated position along the LUT

calibration plot in panel 10D (Fig. 10). A more comprehensive, physical meaning for beads' A-C

position within the LUT is given in section 2.3 (Fig. 10). Generally, such plots can capture images

with nanometer precision at a user defined spacing over a range of 3-10 μm.[26] Comparing the

radial diffraction profiles of experimental and reference beads during real-time measurements

can determine the *z*-displacements of experimental beads with respect to the anchored

reference beads.
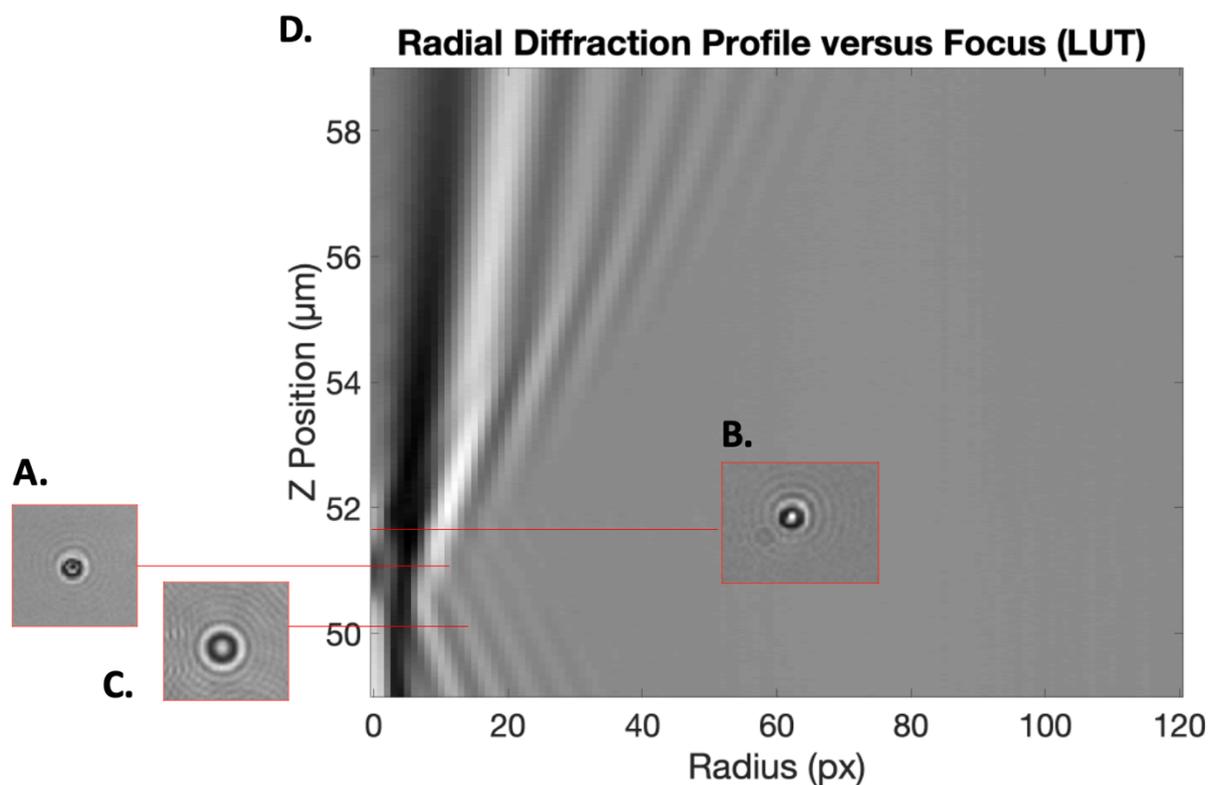
Figure 10. Experimental View to Track and Position a Magnetic Bead. (A-C) Beads' concentric

radial diffraction patterns change in size and fuzziness as the beads move relative from the focal

plane (See Section 2.3). The +z-direction is taken to go up the page. (D) An plot of a radial

diffraction profile versus focus (LUT) for a reference bead.

2.2 MT Experimental Constructs

Since our biomolecules of interest are not inherently magnetic, they were functionalized with

super paramagnetic beads (Invitrogen/ Life Technologies MyOne T1 Dynabeads) for MT

analysis. Complementary biomolecular attachments, streptavidin with biotin and digoxigenin

with anti-digoxigenin, were utilized to build constructs suitable for MT analysis (Fig. 11). Both

sets of attachments demonstrate remarkably high binding constants ($K_{d\ streptavidin} \sim 0.01$ pM, K_d

digoxigenin ~ 12 nM) and stability under our chemical and force conditions up to 100 pN. [24,27,28]

Using two different attachment methods increased the specificity of tether construction.

A.

B.
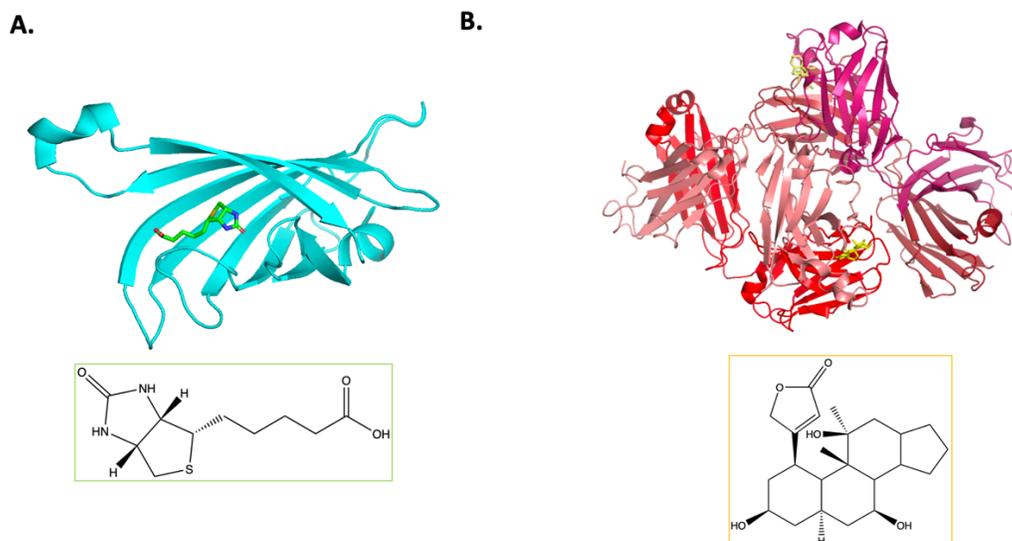
Figure 11. Molecular and Covalent Structures of Complementary Proteins and Ligands. (A) Streptavidin protein with its biotin ligand (PDB: 1STP). (B) Digoxigenin protein with its anti-digoxigenin ligand (PDB: 3RA7). Molecules not depicted to scale.

A construct illustrated in figure 12 was deemed optimal to study RNAP roadblock behavior. One end of a DNA molecule contained a single digoxigenin molecule that was allowed to bind to an anti-digoxigenin coated slide (Fig. 12). RNAP molecules labeled with a single biotin were introduced and allowed to bind to a promoter sequence in the tethered DNA. Streptavidin-coated beads were flushed into the system to bind to RNAP in a 1:1 manner. Finally, various roadblocks, LacI or EcoR1, and buffers with optimal salt and nucleotide conditions to support transcription were added. RNAP transcriptional activity was subsequentially monitored and as force was exerted with the MT to alter translocation by RNAP.
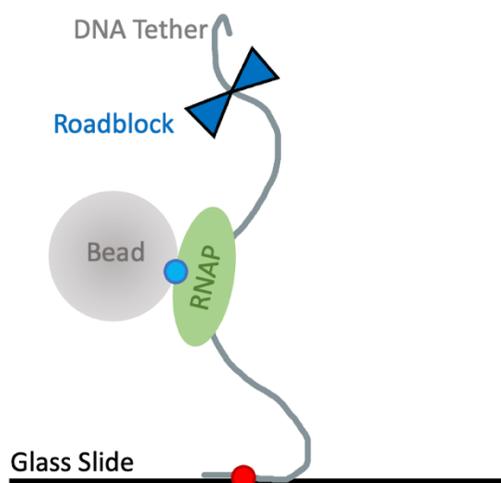
Figure 12. Experimental Bead, Tether, and RNAP Construct. A DNA tether is affixed to a glass slide

through a single digoxigenin and anti-digoxigenin bond (red dot) as well as a streptavidin-coated

bead linked to biotin-labeled RNAP (blue dot) prior to transcription. Construct not depicted to

scale.

2.3   RNAP Visualization and Altering Transcriptional Force

The MT single-molecule approach sought to visualize and influence RNAP progress along a DNA

tether upon encountering a roadblock such as LacI or EcoR1. As mentioned before,

transcription is a highly regulated, physical process, and transcribing RNAP motor enzymes

exert forces in the range of 20-25 pN *in vivo*.[17] Precisely applying forces to constructs illustrated

by figure 12 effectively altered transcriptional force against roadblocks. In this study, a MT

modulated RNAPs' transcription by ±0.2, ±0.7, ±2.0, and ±5.0 pN. As transcription is strictly a

unidirectional process, elongating RNAPs transcribed either up or down a tether (Fig. 13). An

RNAP transcribing upwards was assisted by upward magnetic force on the bead, considered a

positive assisting force (AF) (Fig. 13A). A RNAP transcribing downward was opposed by an

upward magnetic force on the bead, considered a negative opposing force (OF) (Fig. 13B).
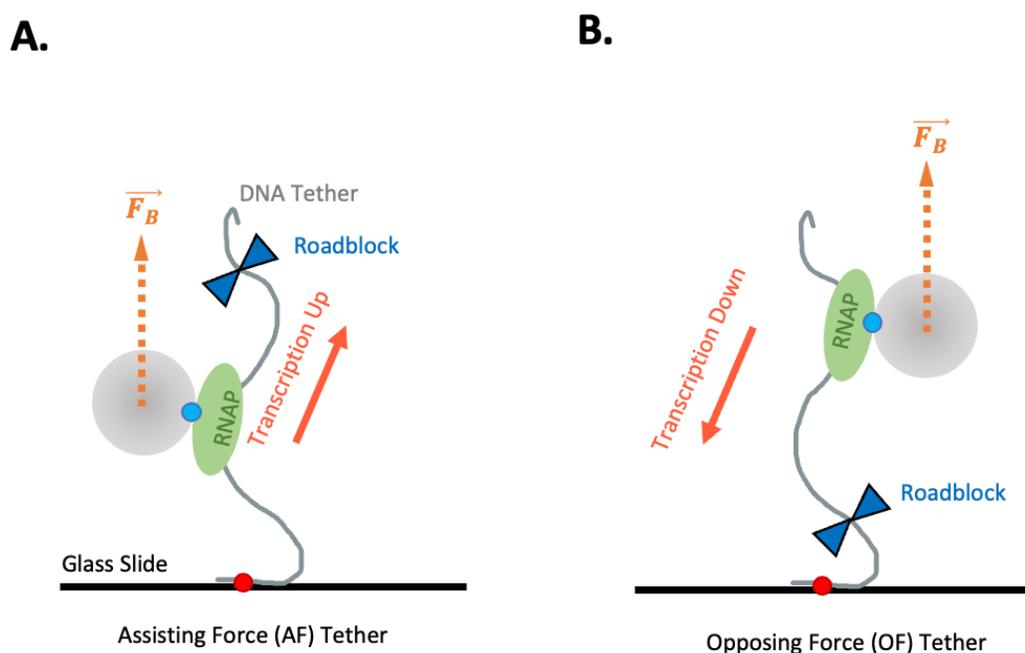
Figure 13. RNAPs Elongating along Tether Construct. Elongating RNAPs along a DNA tether

experienced upward magnetic forces. (A) In an assisting force tether, the upward magnetic force

assisted RNAP transcription. (B) In an opposing force tether, the upward magnetic force

opposed RNAP transcription. Constructs not depicted to scale.

Though figure 13 presumes beads attach to RNAPs on tethers in an intentional, specific

manner, some constructs were assembled incorrectly. Before the addition of all four

nucleotides initiated transcription elongation, it was expected that beads on correctly

assembled constructs would appear in roughly the same z-plane relative to each other. If the

beads were below the focal plane and too close to the bottom of the slide, a diffraction pattern

like figure 14C was observed. These beads were selected as reference beads. Contrary, if beads

fell above the focal plane, a diffraction pattern like figure 14B was observed. Beads like these were presumed to be non-specifically attached by a piece of DNA longer than expected. Patterns like figure 14A were considered optimal and were selected for in experiments, as they represented beads intermediate between these two extremes. Beads with these patterns had the best chance of being attached to an RNAP bound to a DNA of expected length.
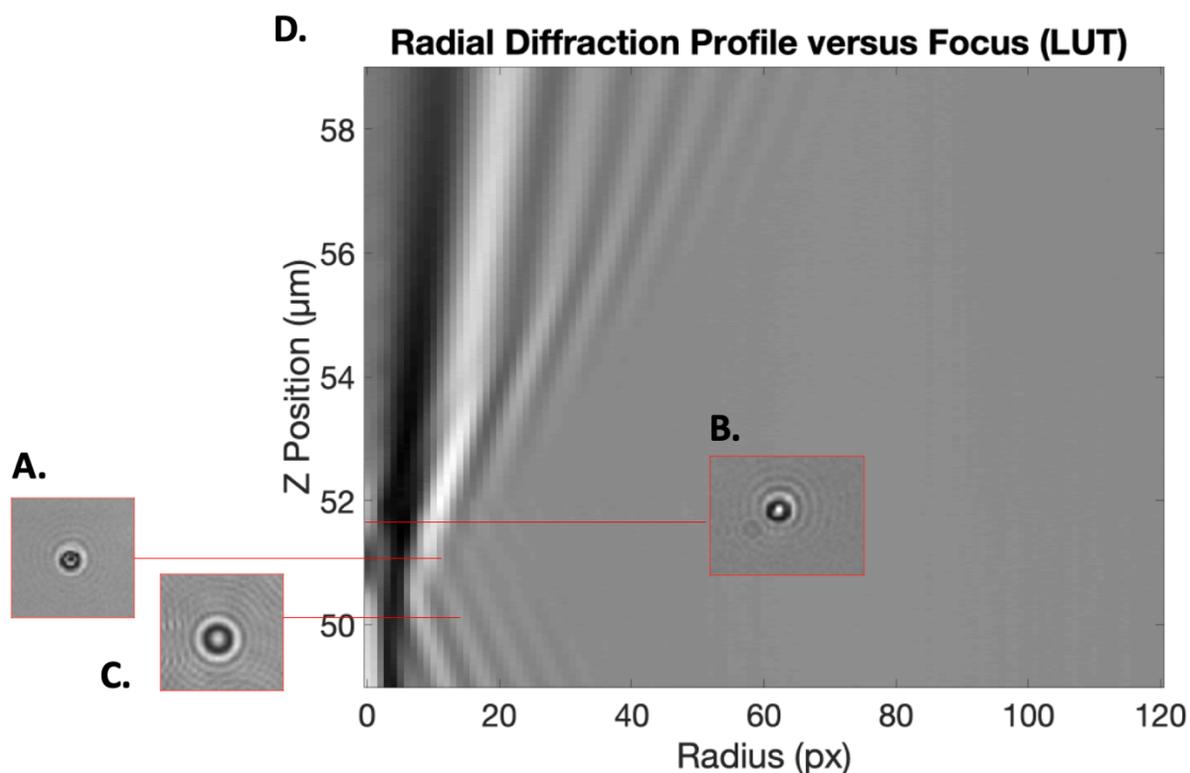
Figure 14. Illustrations of Beads' Radial Diffraction Pattern Relative to a LUT. (A) A sample bead, focused appropriately, was taken as an experimental measurement. (B) A sample bead focused above the objective plane was considered too high and potentially floating in solution. (C) A sample bead focused beneath the objective plane was considered to be too close to the bottom of the slide. (D) LUT comparison for beads A-C's *z*-position.

2.4 Experimental Procedure

Samples consisting of tether-bead-RNAP constructs depicted in figure 13 were placed within a

chemically prepared flow-chamber cell for MT analysis. Flow chamber cells were assembled

from a 24 X 40 mm glass slide, custom cut parafilm slice, and a 22 X 22 mm square glass

coverslip. Parafilm slices contained inlet and outlet channels through a triangular chamber that

held the samples (Fig. 15). Parafilm slices were first lightly pressed between the cleaned glass

slide and coverslip at room temperature. Chambers were then placed on a hot plate. Lightly

heating cells melted the parafilm between the glassware to produce a sealed flow-chamber.

These 10 μL volume sealed chambers were treated with a series of reagent-containing buffers

in order to assemble tethers for MT analysis.

Figure 15. Flow Cell Construct. A finished, parafilm flow cell labeled with the inlet and outlet

channel as well as the glass slide and parafilm components.

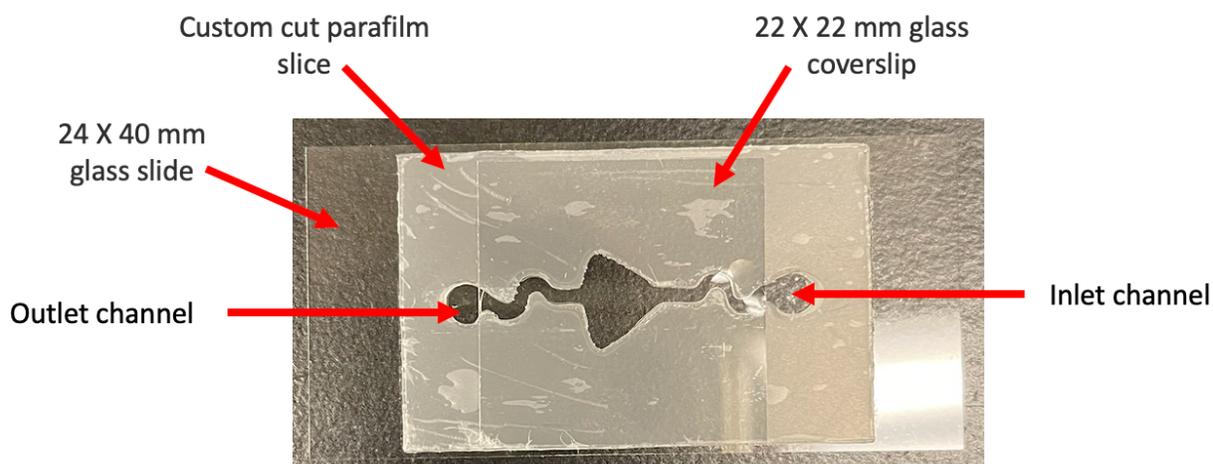Buffers such as phosphate buffer saline (PBS) and a dilution of anti-digoxigenin in bead-wash buffer (BWB) were pipetted through the inlet channel and drawn through chambers using twisted chem wipe to wick solution from the outlet channel. BWB contained 20 mM Tris-glutamate buffered at pH = 8.0, 50 mM potassium glutamate (KGlu), and 1 mM 100x dithiothreitol (dTT). Chambers first received 15 µL PBS (pH = 7.4) consisting of 2 mM chelating agent ethylenediaminetetraacetic acid (EDTA) and 2 mM sodium azide. EDTA removed any ions that could interfere with DNA's structure or RNAPs' transcription, and sodium azide prevented microbe contamination.[29] PBS was chosen for its similar physiological pH and osmolarity. Moreover, this initial PBS perfusion hydrated chambers, so that subsequent buffers would flow and spread evenly within the chamber. Immediately afterwards, a 4:95 dilution of anti-digoxigenin (8 µg/mL, Roche Diagnostics, Madison, WI) in BWB was flushed through chambers. BWB had the same pH as PBS but included KGlu to approximate physiological osmolarity. dTT was added to promote conditions favorable for transcription by preserving reduced sulfhydryl groups and thus protecting their function.[30] Ultimately, diluting anti-digoxigenin with BWB ensured the chamber bottom was coated with the appropriate amount of anti-digoxigenin to affix tethers: adding too little prevented tethers from sticking and adding too much prevented transcription entirely. Approximately 14-20 µL of this dilution was added to chambers. To let the dilution of anti-digoxigenin coat evenly along the chamber floor, chambers were incubated for either for 1.3 h at room temperature in a moist environment or overnight in a 4° fridge. Finally, chambers were incubated with 7 µL blocking buffer (PBS with 1% alpha-casein, GeneTex, Irvine, CA) for 1 h. Alpha-casein in the initial perfusion with PBS prevented reagents from sticking to glass surfaces.

2.5   Advantages and Disadvantages to MT Analysis

As previously mentioned, single-molecule studies may utilize optical trapping or MTs in order to exert forces or torques on systems of nucleic acids and or proteins. As opposed to MTs, optical trapping utilizes focused laser beams instead of magnetic fields to manipulate individual beads.[15] While MTs exert a nearly homogenous magnetic field on many tethered magnetic particles at once, optical trapping manipulates single beads.[31] Though optical trapping allows exquisite translational control of individual beads, intense radiation may photodamage samples during observation. [31, 26] In comparison, MTs produce no radiative damage, can simultaneously manipulate tens to hundreds of individual molecules, and can exert torque which is difficult with optical trapping.[26]

On the other hand, MTs also suffer from mechanical vibrations that may move molecular systems: motors that turn or lower the permanent magnets may cause slight mechanical vibrations that can affect data measurement if experiments require simultaneous collection and magnet movement. For example, experiments that measure bead height best record data when the magnets are stationary. Although some studies cite cost barriers to more widespread usage of stationary electromagnetic MTs, some of which can overcome this limitation, other studies report reasonably priced constructs that can effectively twist, stretch, and measure tether height without mechanical vibrations. [31, 32] However, such technologies are under development. To prevent motor vibration interference, this study first positioned the magnets at an appropriate vertical height then recorded measurements.

Despite some limitations, MTs provide a sufficient means to analyze RNAP transcription along a

DNA strand. Studies report MTs with a spatial resolution of a few nanometers and a force range

of a few tenths to tens of pNs.[26] This study's MT technique provided precise control over

multiple constructs without significant concerns of sample degradation or measurement

error.[31]

3 Chapter 3: Materials and Methods

3.1 DNA Tether Construction Approach

We generated linear DNA tether molecules with two orientations: one such that RNAP would

transcribe in the same direction as that of the applied force (assisting force) and another where

the force would oppose transcription (opposing force) (Fig. 16). All tethers contained a T7A1

promoter, EcoR1 or LacI binding site, and a λt1 terminator in sequential order (Fig. 16). The

tethers were obtained from appropriate plasmids, circular DNA molecules and bacterial

chromosomes, containing the elements of interest at specified distances (Fig. 16). Tethers,

linear segments selected from plasmids, of precise length were amplified via a polymerase
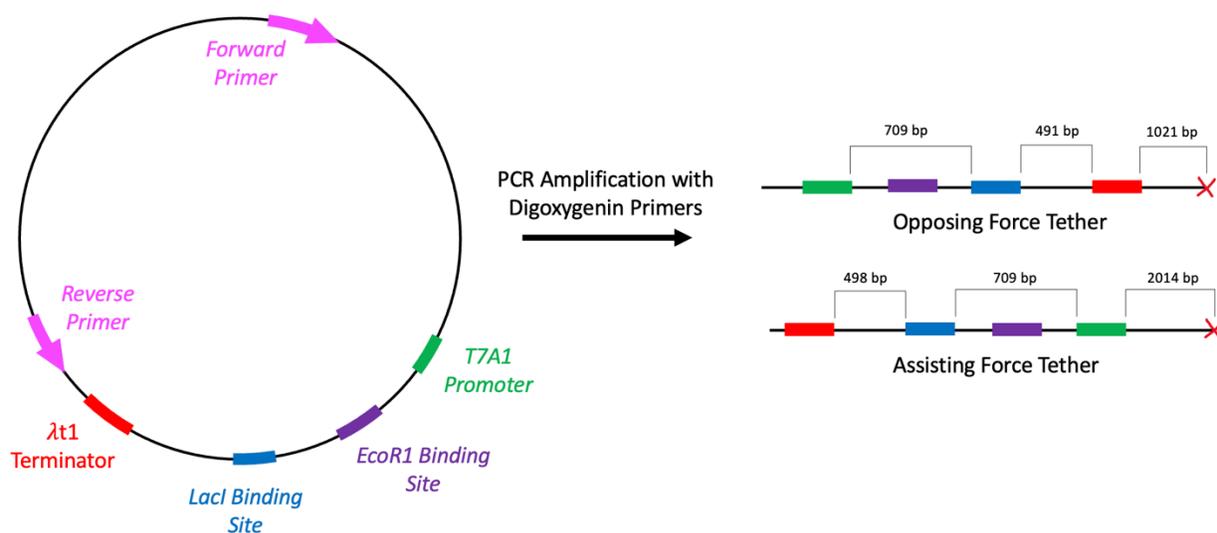
chain reaction (PCR).

Figure 16. General Scheme of Tether Construction. This study used multiple rounds of PCR to

produce linear DNA segments containing a promoter, roadblock binding site, and terminator.

Base-pair distances and features not depicted to scale.

Overall, assisting force constructs consisted of approximately 4 kbps while opposing force constructs were 3 kbps. These lengths were sufficiently long to monitor RNAPs' location and progress through each relevant feature. The additional 1 kbp length in assisting force constructs between the T7A1 promoter and digoxygenin labeled end was added to ensure RNAPs would not stick to the anti-digoxygenin glass slide during initiation. However, the distance between the T7A1 promoter and λt1 terminator was the same in both tethers, approximately 1.2 kbp (Fig. 16).

3.2 DNA Tether Construction with PCR and Purification

Many scientific disciplines depend on PCR as a reliable means to exponentially replicate dsDNA from a template strand and to greatly increase the amount of the starting dsDNA. This study leveraged PCR's ability to amplify plasmid segments with appropriate features and lengths in order to generate experimental linear assisting and opposing force tethers (Fig. 16). By repeating three basic steps a number (n) of times (cycles), PCR can ultimately generate 2^n copies of dsDNA (Fig. 17). The first step heats the templates at 94-5˚C. This breaks the hydrogen-bonds in DNA's secondary structure, turning dsDNA to ssDNA which allows primers to bind to their complementary sites in the subsequent annealing phase. This second step, the annealing phase, is conducted at lower temperatures (50-6˚C) to prevent non-specific primer binding. In the third and last step, the replication phase, *Thermus aquaticus* DNA polymerase (*Taq*), a bacterial enzyme with optimal functioning at 72˚C, elongates the DNA strand complementary to the template strand by adding deoxynucleotide triphosphates (dNTPs) to the

3' end. The three steps of denaturation, annealing, and replication were performed around 30-

40 times in order to produce a sufficiently concentrated dsDNA solution.
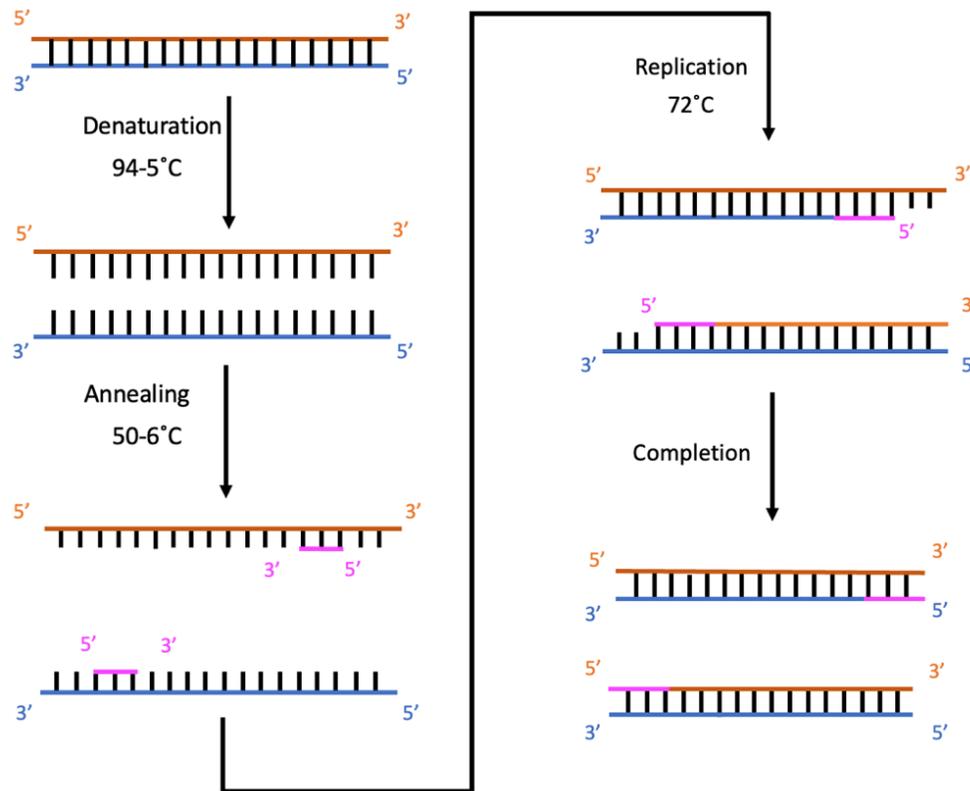
Figure 17. General PCR Scheme. Reactions occur in three sequential steps: denaturation,

annealing, and replication. Repeating this cycle n times generates 2^n dsDNA copies.

For assisting force LacI and all EcoR1 trials, plasmid pDM_E1_400, forward primer D-s/YY_400-

103 (5'-GCTTGGTTATGCCGGTACTG-3'), and reverse primer A/pUC18-nuB104/2043 (5'-

ACGACCTACACCGAACTGAG-3') were used. For the remaining opposing force LacI constructs,

plasmids pDM_N1_400, pZV_NI_400, or pDM_N2_400, forward primer S/JBOID01- 400/2086

(5'-AGCTTGTCTGTAAGCGGATG-3'), and reverse primer Dig-A/JBOID01-400/5096 (5'-

ATCGTTGGGAACCGGAG-3') were used. PCR reactions used 0.5 µL plasmid, 2.5 µM forward primer, 2.5 µL reverse primer, 25 µL Q5-High-Fidelity 2X Master Mix (New England Biolabs Ipswich, MA), and 19.5 µL autoclaved H_2O. The Q5 High-Fidelity 2X Master Mix contained *Taq,* full spectrum dNTPs for elongation, and stabilizing Mg^{2+} ions within a general-purpose buffer. To dilute the transcription solution and run more chambers, the 80 µL volume was split evenly and underwent approximately 30-40 PCR cycles.

For the first assisting force LacI and EcoR1 construct, a 4149 bp expected product was verified with a 1% agarose electrophoresis gel (Fig. 18 A, B). To remove post-PCR impurities like excess dNTPs, Mg^{2+}, buffer, or *Taq,* purification following Thermo Scientific GeneJET protocol was performed.[33] However, a significant loss of post-PCR dsDNA product was observed (Fig. 18C). As a result, approximately 1-1.5 µL DNA was used in assisting force experiments requiring this construct as opposed to the standard 0.75 µL DNA in other experimental conditions within this study.
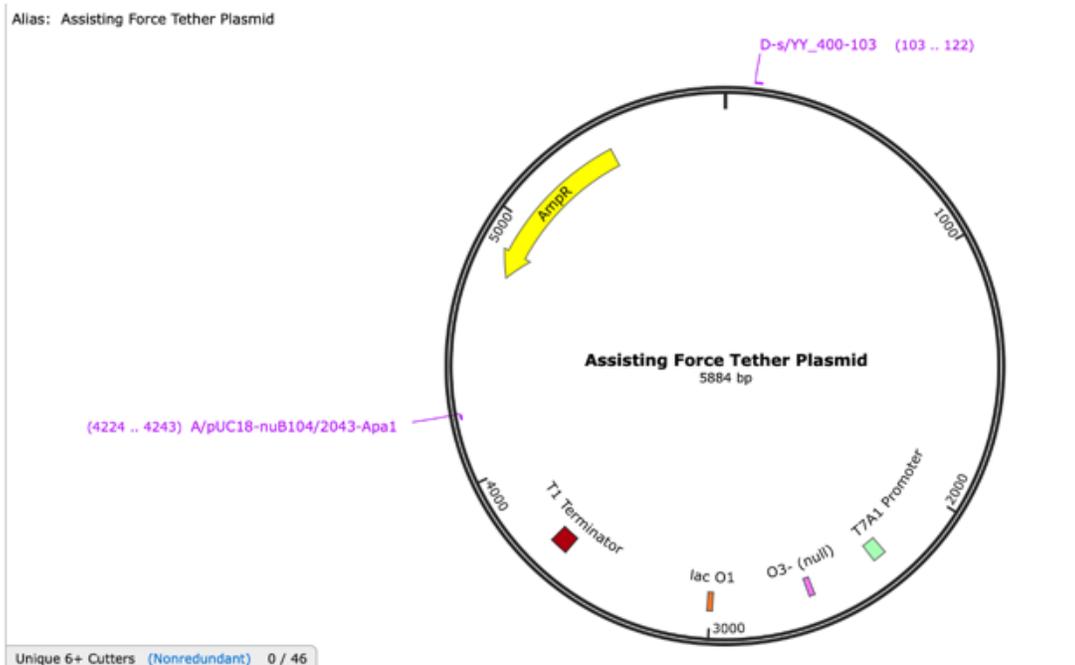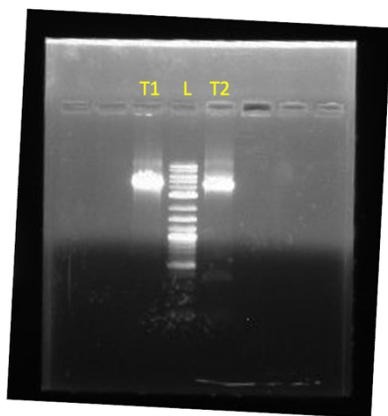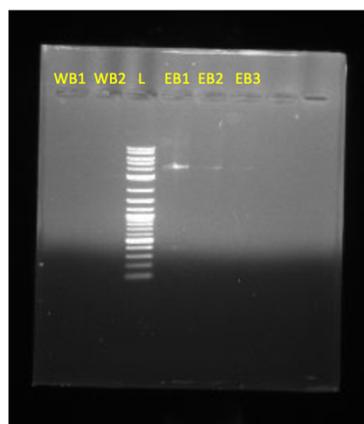
Figure 18. SnapGene EcoR1 Construct Verified by Experimental Gels. (A) Assisting force SnapGene construct for EcoR1 built within this study. (B) A 1% agarose electrophoresis gel verified the expected 4149 bp construct was properly synthesized via comparison with a New England Biolab ladder standard, not pictured due to copyright permissions. Both trials resulted in the expected length. (C) Thermo Scientific GeneJET purification resulted in a significant loss of product in the first elution buffer step.

3.3 Chamber Preparation for Magnetic Tweezers

15 µL of 0.2µM PBS, followed by 14-20 µL of a 4:95 dilution of anti-digoxigenin to bead wash

buffer (BWB, See Section 2.4) were introduced into chambers (See Section 2.4). Chambers were

left to incubate either at room temperature for 1.3 h, or in a 4° fridge overnight in a closed box

with a damp paper wick placed at the bottom to prevent samples from drying out. After the

anti-digoxigenin incubation, 7 µL blocking buffer (See Section 2.4) was added to prevent non-

specific binding of beads to the glass chamber. Chambers were left to incubate at room

temperature for 1 h in the same box.

Next, a 1% dilution of DTT in transcription buffer (TBX) was prepared (TB+DTT). TBX contained

20 mM tris-glutamate buffered at pH = 8.0, 50 mM KGlu, 10 mM magnesium glutamate

(MgGlu$_2$), 1 mM 100x DTT, 0.2 mg/mL alpha-caesin, and 13.2 mL autoclaved H$_2$O. Tris-

glutamate served to maintain a near-physiological pH, and salts KGlu and MgGlu$_2$ served to

maintain physiological osmolarity. To protect the buffer's fidelity and preserve biological

molecules, it was always kept on ice along as well as other biomolecular solutions.

To prepare the tether-RNAP construct, 1 µL of a 30 nM biotinylated RNAP holoenzyme was

diluted with 9 µL TBX. We placed 1 µL of this RNAP dilution, 18 µL TB+DTT, and approximately

0.75-1.50 µL DNA into a 37° heater for 10 min. Heating at physiological temperatures allowed

RNAPs to bind to the promoter in initiation. To begin elongation up to a certain point, 1 µL

guanosine phosphate adenine (GPA, Trilink, San Diego, CA), a covalently linked dinucleotide

which served as the first transcriptional nucleotides and 2 µL of 5 mM AUG dNTPs (New

England Biolabs) were added (Fig. 19). Since no C nucleotides were added, all RNAPs stalled at

the first C in the DNA that required a missing dCTP. This 20 µL transcriptional mix was placed

again for 10 min in a 37˚ heater to facilitate transcription elongation.
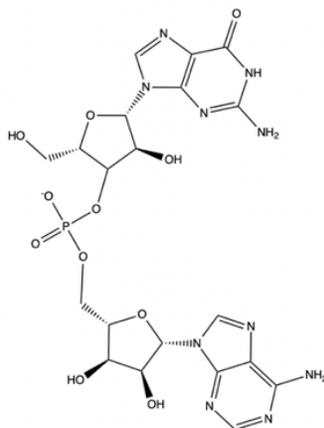
Figure 19. GPA Dinucleotide Structure. Guanosine phosphate adenosine (GPA) RNA dinucleotide
served as the first nucleotide incorporated by RNAPs entering elongation.

In between incubations, 0.75 µL magnetic beads were taken from an aliquot and resuspended

in approximately 100 µL BWB. To separate beads stuck to each other, the solution was

thoroughly vortexed until cloudy. One side of the test tube was put in contact with a magnet

for five minutes. This time frame was considered optimal for the magnetic beads to stick to this

side and for the non-magnetic beads to fall to the bottom in solution. After the five minutes,

these inert beads were removed along with approximately 60-90 µL of solution. Depending on

the qualitative concentration of the solution, an additional 120-150 µL PBS was added. Again,

beads were thoroughly vortexed, placed against the magnet, and drained after 5 minutes.

Beads finally received approximately 90 µL TB+DTT and were placed on ice until needed.

Once the transcriptional complex finished heating in its second incubation, it was diluted in half with the addition of 20 μL TB+DTT to a final volume of 40 μL. Diluting in half allowed more chambers to run with relatively the same data yield. Approximately 10 μL of this solution was drawn through chambers and sat for 10 minutes at room temperature. Next, 10 μL of the prepared magnetic beads were added and left for 3 minutes at room temperature. A single streptavidin-coated bead bound specifically to a biotinylated RNAP. To remove all other, unnecessary solutes or molecules from previous reactions, chambers were washed with 20 μL TB+DTT. One chamber was selected for immediate roadblock addition and MT analysis, and the rest were placed within their damp, sealed box in a 4˚ fridge. Roadblocks were introduced either within a solution with full spectrum dinucleotides or diluted by themselves, alone, depending on the roadblock's experimental procedure (See Sections 3.4 and 3.5).

3.4   Chamber Preparation for LacI

Experiments used *in vivo* protein roadblock LacI both with and without mRNA endonuclease growth factor A (GreA). To affix LacI to its binding site, a 1 μL aliquot was diluted by a factor of 200 in TB+DTT to a final concentration of 20 nM. Next, 5 μL of this diluted LacI solution received 5 μL NTPs (New England Biolabs) and 90 μL TB+DTT. In particular, 5 μL each of 100 mM dATP, dUTP, dGTP, and dCTP (New England Biolabs) were mixed together so that each nucleotide had final concentration 25 mM. Though RNAPs already elongated using the first three nucleotides, the intentional injection of dCTP would begin RNAPs' transcription past the first C nucleotide in the DNA. However, if trials required GreA, a 1 μL GreA aliquot (TriLink), 5 μL LacI, 5 μL NTPs, and 89 μL TB+DTT solution was prepared instead.

3.5   Chamber Preparation for EcoR1 Under High Salt Conditions

The novel roadblock this study explored was mutant endonuclease EcoR1 which served as a non-interacting comparison to the LacI roadblock. The LacI protein had been observed to associate and dissociate from its binding site with an expected frequency (See Section 1.4) so that discrete RNAP pauses could be observed and measured. Early experiments with EcoR1 followed the LacI procedure. However, RNAPs transcribing against EcoR1 either presented an infinite stall (on the scale of 2 h measurements) as if the roadblock was either very strongly bound to its site or presented no pause as if EcoR1 was apparently missing from the DNA. Previous work demonstrated EcoR1 stalled around 80% of a RNAP population for times up to 2 h but appeared to dissociate for the remaining 20% allowing uninterrupted transcription (Fig. 8D).[23] This roadblock feature posed a challenge to measure RNAP pauses induced by EcoR1 efficiently, without relying on GreA to promote a readthrough. Thus, the main objective for this study was to bypass this problem.

Tracking RNAPs' downwards transcription in an opposing force tether demonstrated that traces stalled indefinitely at EcoR1 without GreA (Fig. 20A). However, adding GreA appeared to rescue stalled complexes–perhaps by cleaving mRNA pushed forward into RNAP's active site (Fig. 8). GreA thus allowed RNAPs to immediately move past the roadblock obstacle without an observable pause (Fig. 20A). While GreA only substantially helped some stalled complexes in the presence of forces opposing transcription, it had negligible effect on readthrough measured under forces assisting transcription (Fig. 20B).
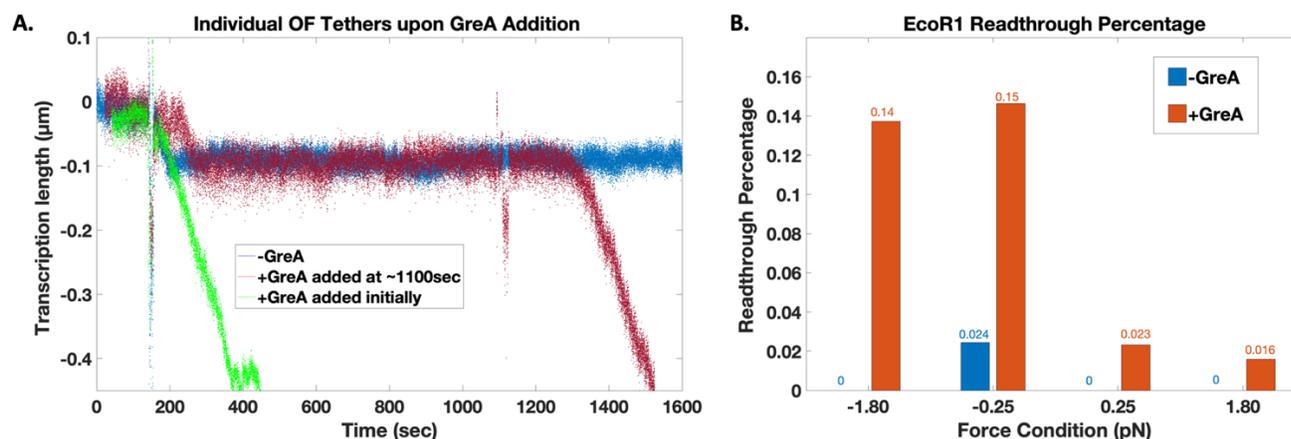
Figure 20. Initial Results with EcoR1. An RNAP transcribes downwards from its initial position at 0 μm. (A) A series of EcoR1 opposing force (OF) traces with various GreA conditions. First, GreA was never added in the blue trace, so the complex stalls indefinitely even for times exceeding 2 h. Next, GreA was added at the beginning in the green trace, so the complex reads through without pausing. Finally, in the maroon trace, GreA was added approximately 1100s into an experimental run, and complexes were observed to readthrough upon its immediate addition. (B) Despite GreA's singular effect on RNAPs stalled at EcoR1, it only boosted readthrough percentages significantly (at most 15%) in opposing force trials. Unpublished data collected by Jin Qian.

However, this study successfully induced EcoR1's association and dissociation from its binding site at an appropriate frequency for an elongating RNAP to visibly pause then read through the obstacle. A previous study observed sufficient EcoR1 dissociation independent of nascent oligonucleotides at 180 mM NaCl, so salt concentrations around this benchmark were explored first.[34] So, 50, 150, and 200 mM KCl TBXs were prepared from the standard procedure (See Section 3.3). During preliminary trials, the 200 mM TBX was tested first to prepare EcoR1 the same as LacI with the following exception: this high salt TBX was used solely in place of regular

TBX for chamber preparation as well as roadblock and nucleotide preparation (See Section 3.3, 3.4). Some traces with discrete visible pause times were successfully observed, and discrete RNAP pausing at EcoR1 could indeed be measured, as opposed to the indefinite stalls observed before (Fig. 21A).

Yet, the data quality and yield significantly declined over time as chambers sat longer in the presence of increased KCl concentrations. We speculated that longer salt exposure caused beads to float towards the surface and or chamber components to dissociate. As a result, a higher concentration 500 mM KCl TBX was prepared and added in the two steps preceding MT analysis: in the EcoR1 aliquot dilution and in a similarly diluted NTP solution. Foremost, the molecular fidelity of chambers not selected for immediate analysis was preserved. But most importantly, exposing chambers to high salt TBX through the NTP and EcoR1 dilutions for shorter times yielded more efficient, viable data not subject to dissociation over the experiment's duration (Fig. 21B).
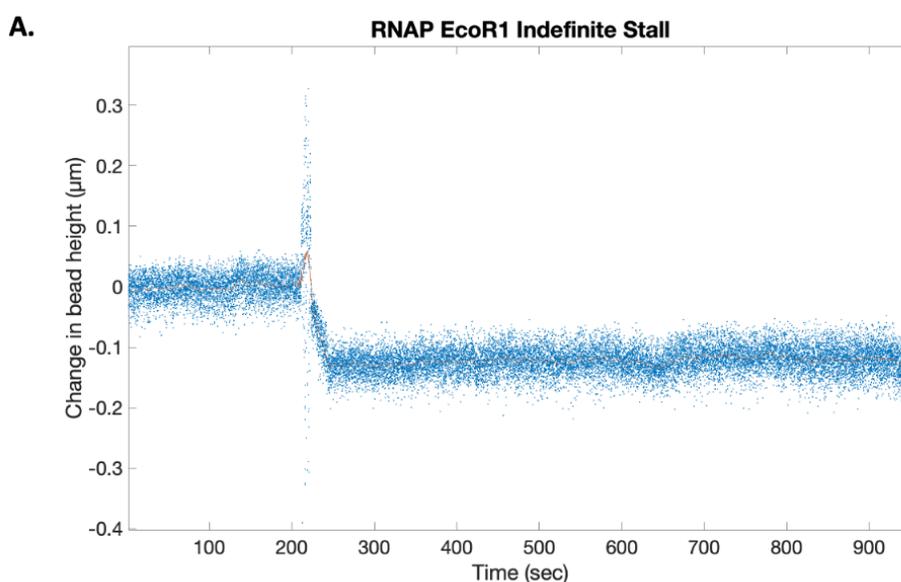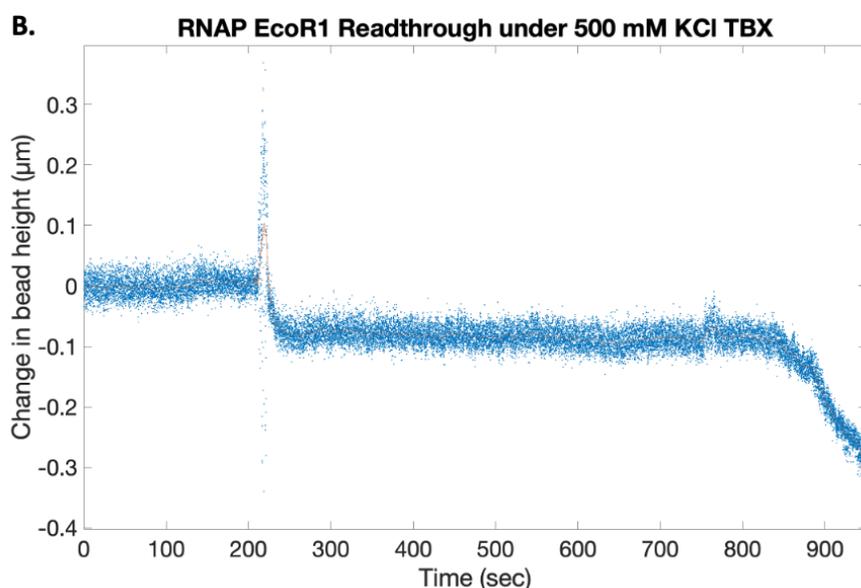
Figure 21. EcoR1 Success under High Salt Conditions. RNAP transcribed downwards against an EcoR1 roadblock (-0.1 μm site) in these opposing force traces. (A) Some RNAPs stalled indefinitely at EcoR1 without the high salter salt TBX treatment. (B) Others demonstrated a discrete pause time, a novel finding of this study, using high salt 500 mM KCl TBX.

3.6 Magnetic Tweezer Field of Capture

Post-preparation, a chamber was selected for immediate MT analysis with either assisting or opposing force tether constructs (See Section 3.2). A drop of oil was added on top of the MT's 63X objective, and the stage was quickly dropped so that the bottom of the sample slide pressed against the oil (Fig. 22A). The oil between the objective and sample provided a focused path for the camera's light to illuminate and visualize beads. This produced a field of approximately 30-60 beads from which a user could scan fields by moving the stage along the xy-plane (Fig. 22B). The MT software drew bright boxes of random, striking colors around

experimental or reference beads (Fig. 22B). The software also generated a label next to each

bead in red text to denote it as either "Meas" or "Ref," respectively (Fig. 22B). The MT

determined each bead's initial *xy*-position from its *xy*-Brownian motion and marked it with a

red cross. The initial *z*-position calibrated from a reference bead, was displayed as a number

beside it. The preferred types of beads were selected as stated in the procedure (See Section

2.3). A singular field of view per chamber was determined based on how many viable tethers

were observed.

A.

B.

Figure 22. MT Procedure and Field of Capture. (A) Samples were contained within a sealed glass-parafilm chamber. The bottom slide was placed in contact with a drop of oil on the objective. (B) An experimental field of view with some beads selected in brightly colored boxes. The color of the box was determined by the MT code randomly, but the code labeled each bead as either a reference (Ref) or experimental (Meas).

3.7 Magnetic Tweezer Force Application

Next, the permanent magnets were lowered to particular heights corresponding to known upwards magnetic force magnitudes. Decreasing the distance between the sample and the magnet increased the upwards magnetic force (Fig. 23). Twisting magnets induced torque (Fig. 23). Past calibrations noted that the heights 13, 14.5, 18.5, and 20 mm corresponded to 0.2, 0.7, 2.0, and 5.0 pN, respectively. The orientation of tether DNA and thus direction of transcription were pre-determined by the type of DNA used (See Section 3.2).
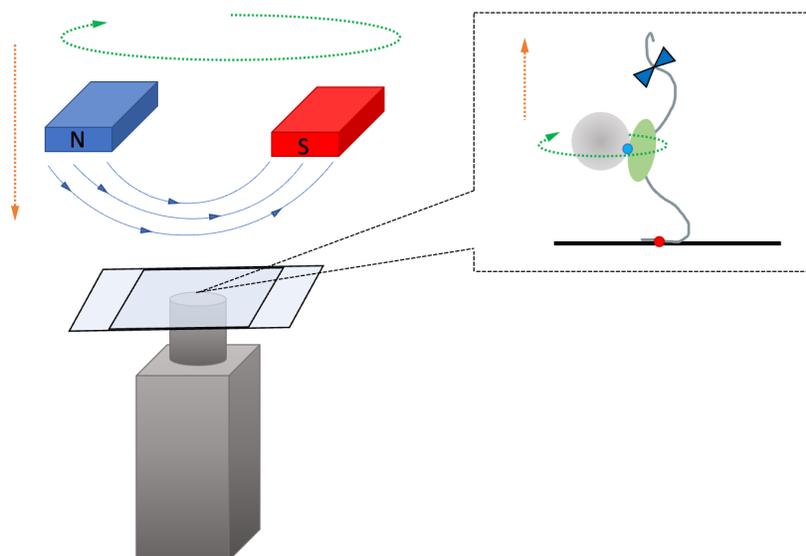
Figure 23. Forces and Torques Produced by MTs. Lowering the MT's permanent magnets towards the sample increased the magnitude of upwards magnetic force while twisting them induced torque.

3.8 Observed Pause Times as Individual Histograms

After selecting the appropriate magnet height, a series of radial calibrations were performed to obtain radial profiles consistent with figure 10D. This study utilized a range of 6 μm: from a qualitatively determined Z = 0 focal plane, a radial diffraction profile was taken -1.5 μm beneath the objective height and 4.5 μm above the objective height. In LacI experiments, the MT recorded approximately 3 minutes of RNAP stalled at the first, missing C nucleotide. Next, after injecting the diluted LacI and NTP solution, RNAP transcriptional behavior begun and was measured for 30 minutes to 1 h. In EcoR1 experiments, 10 μL of the 500 mM TB+DTT and EcoR1 dilution was flushed through chambers which incubated for at least 10 minutes on the MT

stage before adding dCTPs. Similar to LacI experiments, stalled RNAPs were observed for a few minutes before chambers received a high salt NTP solution, then RNAP transcription was observed for 30 min to 1h. Both roadblock conditions received either no or 1 μL GreA in their NTP solution.

Within some range of z-distance, the MTs tracked RNAPs' transcriptional progress and displacement in μm from an initial 0 μm position over time (s) (Fig. 24). During experimental measurements, RNAPs transcribed a bit until they reached the first C nucleotide, intentionally withheld so that all RNAPs stalled at relatively the same location (Fig. 24). Upon the manual injection of C nucleotides, RNAPs transcribed towards the roadblock, in particular LacI at O1 for figure 24. As transcription is strictly a unidirectional process, RNAPs transcribed either down or up a tether based on the DNA added preparing chambers beforehand (Fig. 24). Figure 24's flat plateaus represented the average roadblock pauses. Plateaus were selected based on the roadblock's expected height in solution. These values were ±0.2 μm for LacI assisting and opposing conditions, respectively, and ±0.1 μm for EcoR1 assisting and opposing conditions, respectively.[23] After the duration of the roadblock pause, RNAPs overcame the roadblock and continued their transcriptional progress to a terminator expected at the ±0.4 μm height for assisting and opposing conditions, respectively (Fig. 24). Overall, at least 30 of these traces were recorded for each roadblock, force direction and magnitude, and GreA treatment. Force magnitudes explored were 0.2, 0.7, 2.0, and 5.0 pN.
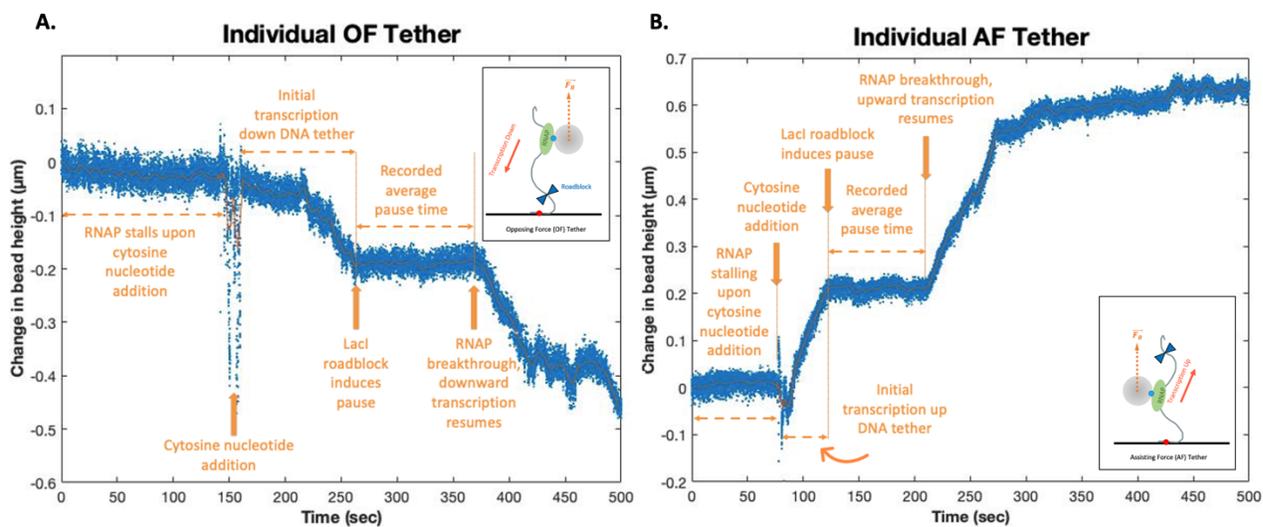
Figure 24. RNAP Transcriptional Progress against LacI-O1. Trace characteristics looked similar for EcoR1 using high salt conditions. RNAP behavior and experimental procedures are annotated for each feature along a trace. (A) Downwards opposing force tether. (B) Upwards assisting force tether.

4   Chapter 4: Results and Discussion
4.1   Comparing RNAP Transit for EcoR1 and LacI-Os

Traces of three types were considered: those exhibiting a finite roadblock pause evidentiated

by an observable plateau, those with an indefinite stall with an indefinite plateau, and those

with no observable pause, where the roadblock appears to be absent. Data processing by

graduate student Jin Qian yielded ratios of transit or readthrough ratios for insurmountable

obstacles EcoR1 and LacI-Os roadblocks (Fig. 25). The ratio of transit for each roadblock was

found by dividing the number of readthroughs–whether a discrete pause was noticed or not–by

the total number of all traces. Only the relatively strong roadblocks EcoR1 and LacI-Os were

investigated in this manner because all RNAPs eventually transcribed past relatively weaker

roadblocks LacI-O1 and LacI-O2; in other words, unlike LacI-O1 and LacI-O2, EcoR1 and LacI-Os

had the potential to stall some population of RNAP indefinitely for at least 2 h.[23]

Readthrough ratios were collected, and it was noted for LacI-Os that findings seemed to

confirm results obtained in our group by Jin Qian where readthrough ratios depended on the

direction of force rather than its magnitude: similar transit ratios were observed between

assisting and opposing forces, but the addition of GreA appeared to boost readthrough ratios

further in opposing forces (Fig. 25A). Since it was speculated opposing forces and GreA promote

backtracking, it was thought that backtracking may help RNAP successfully overcome LacI-Os

(Fig. 25A). Additional experiments were performed with *in vitro* mutant endonuclease EcoR1, a

roadblock for which there is no evidence of interaction with RNAP interaction. We presumed

EcoR1 was slightly stronger than LacI-Os due to EcoR1's decreased readthrough ratio under 50 mM KCl (Fig. 25B).

First, EcoR1 under normal salt (50 mM KCl TBX) yielded low readthrough ratios, indicating that the roadblock effectively blocked nearly all transcribing RNAPs (Fig. 25B). Yet, the novel procedure with high salt 500 mM KCl TBX buffer increased readthrough ratios compared to standard 50 mM KCl TBX salt conditions. Within the high salt population, EcoR1 displayed a similar trend to LacI-Os where the readthrough ratio was affected more so by force direction rather than magnitude. The addition of GreA in opposing force conditions boosted readthrough ratios the most in both EcoR1 and LacI-Os, indicating that backtracking may play a vital role for RNAP transcription through relatively strong roadblocks. How force conditions and GreA alter RNAPs' transit ratio led us to speculate that RNAP may take different mechanisms to overcome roadblocks of comparatively different strengths.
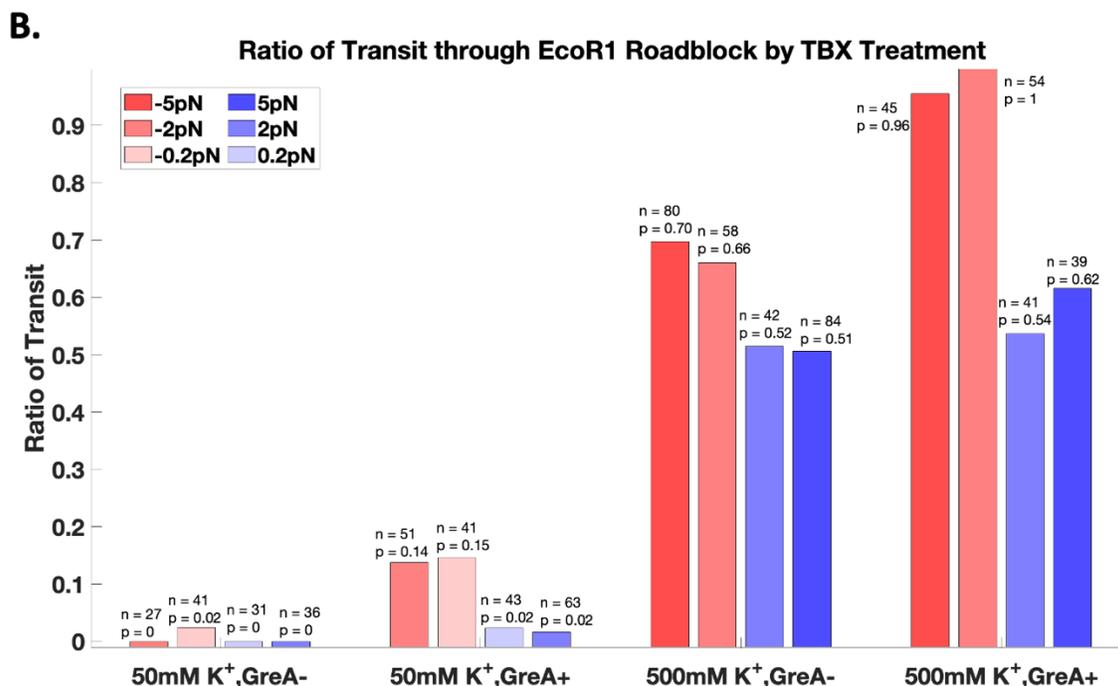
A.

B.
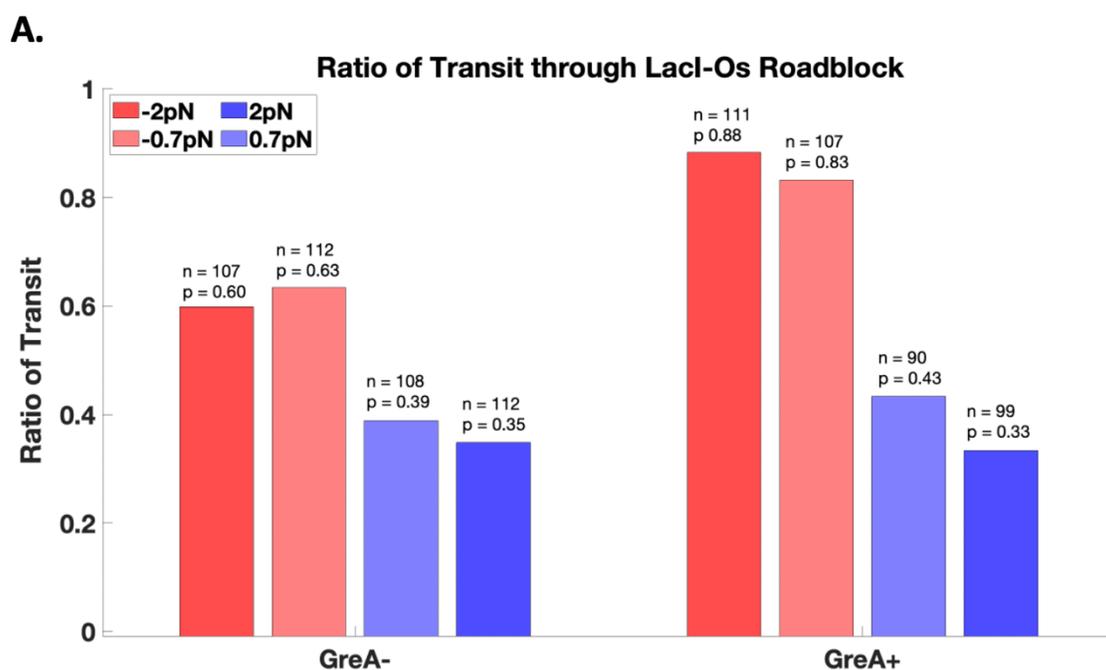
Figure 25. RNAP Readthrough past Relatively Strong Roadblocks LacI-Os and EcoR1.[23] (A)

Readthrough ratios collected for LacI-Os under standard, 50 mM KCl transcription buffer (TBX). (B)

Readthrough ratios collected for EcoR1 under standard, 50 mM then 500 mM KCl TBX. Across

both roadblocks, readthrough ratio tended to depend more on the direction of the force rather

than its magnitude. Adding GreA to opposing force conditions increased readthrough ratios above

the opposing force no GreA baseline.

4.2   Pause Time Comparison between LacI and EcoR1 Roadblocks

The numerical, average pause durations of RNAPs stalling at LacI and EcoR1 under high-salt

conditions were calculated. Previously, the readthrough ratios only proved applicable to

relatively strong roadblocks EcoR1 and LacI-Os because all RNAPs eventually transited through

the relatively weaker LacI-O1 and O2 in 2 h measurements. However, the RNAPs' average pause

times at roadblocks were collected for all of this study's roadblocks (Fig. 26).

Across the LacI system, the duration of pause time followed the expected strength of

roadblocking (Fig. 26 A-C) (See Section 1.4). Compared to an assisting force baseline without

GreA, adding GreA under opposing forces increased average pause times by 50 seconds (s) (Fig.

24 A-C). For LacI-O1 and LacI-O2, adding GreA appeared to level out all pause times regardless

of force direction or magnitude (Fig. 26 B, C). Thus, while adding GreA brought all average

pause times down to the assisting force, GreA independent baseline in these instances, pause

times rose noticeably for opposing forces without GreA (Fig. 26 B, C). Assuming backtracking

lengthens average roadblock pauses, the observation that GreA levels out the average pause

times across all force conditions in the LacI-O1 and LacI-O2 system indicates two key concepts

(Fig. 26 B, C). First, opposing forces may promote backtracking because the singular action of

GreA–to cut mRNA clogged in RNAP's active site– can effectively lower pause times to the

assisting force baseline. Second and consequentially, assisting forces must hinder backtracking,

as GreA had no observable effect on assisting force pause times. Taken together, it was

speculated that backtracking may not prove vital to RNAPs overcoming LacI-O1 and LacI-O2.

Instead, RNAPs may passively wait for these two, relatively weaker obstacles to dissociate

before transiting through.

For LacI-Os, adding GreA under assisting forces leveled out pause times as with LacI-O1 and

LacI-O2, but GreA's addition to opposing forces actually lowered average pause times beneath

its assisting force baseline (Fig. 26A). The finding that RNAPs pause for shorter times, on

average, for opposing forces in the presence of GreA supports the notion that backtracking may

prove more efficient for overcoming stronger obstacles like LacI-Os. Consequentially, it was

suggested that RNAPs may use repetitive cycles of backtracking and GreA rescue to overcome

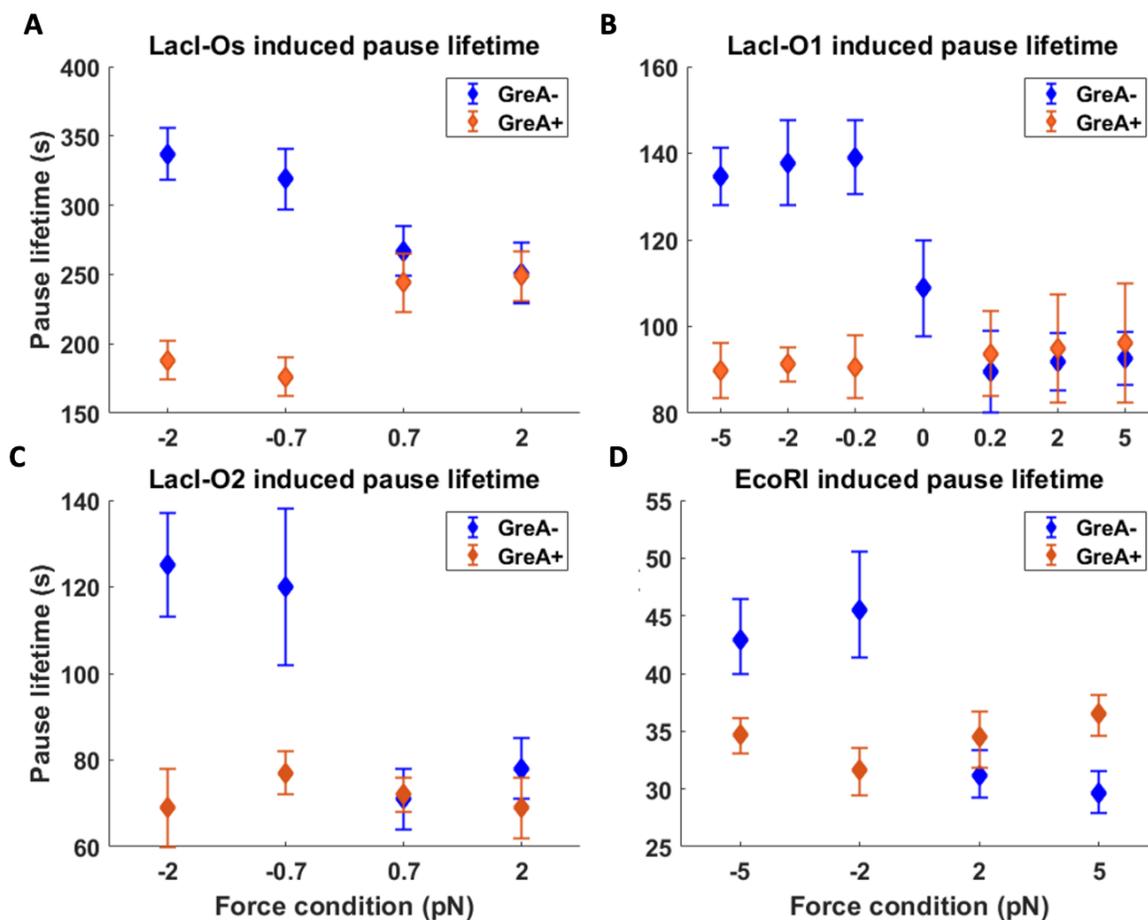the relatively stronger obstacles LacI-Os and EcoR1.

Figure 26. Exponentially Distributed Trace Summaries under Force and GreA Conditions. (A-C)

RNAP roadblock average pause times decrease with decreasing LacI roadblock affinity. (D) RNAP

roadblock average pause times for EcoR1 with the 500 mM KCl TBX procedure.

4.3   Proposing a Hybrid Model

The patterns obtained in figures 25 and 26 elucidated a mechanism where RNAP may proceed

with either an active or passive mechanism depending on the roadblock strength. In the active

mechanism, the most efficient pathway manifests itself as repetitive cycles of backtrack and

recovery. Such a pathway would help RNAP overcome stronger roadblocks LacI-Os and EcoR1,

where readthrough was indeed boosted by opposing forces and GreA as well as where pause

times were brought beneath a standard assisting force baseline (Fig. 25). On the other hand,

the passive mechanism may involve RNAP waiting for less strong roadblocks to dissociate

before proceeding. For example, RNAP overcoming LacI-O1 and LacI-O2 demonstrated assisting

force pause times unaffected by GreA, indicating that backtracking does not occur to a

significant extent (Fig. 26).

From this qualitative understanding, Jin Qian formulated a quantitative model consisting of

three kinetic parameters (s^{-1}) and a probability constant:

$$k_1 = bactracking\ rate,$$

$$k_2 = backtracking\ recovery\ rate,$$

$$k_3 = roadblock\ dissociation\ rate,$$

$$P1 = probability\ of\ removing\ roadblock\ during\ a\ single\ encounter.$$

The roadblock dissociation rate k_3 was determined by the pause lifetimes under assisting force

experiments.[23] The backtracking rate k_1, backtracking recovery rate k_2, and roadblock removal

probability $P1$ were also fitted from an experimental pause lifetime distribution, with the

assumption that $k_2 = 0.46k_1$, an estimate of the energy of the RNAP at the roadblock

obstacle.[23] These three parameters and constant were applied for RNAPs transcribing past this study's roadblocks and are illustrated figure 27A as the following states:

a: an actively transcribing RNAP,

b: an RNAP encountering a roadblock,

c: an RNAP backtracking upon encountering a roadblock,

d: the roadblock passively dissociates from its DNA site with rate k3,

e: an actively transcribing RNAP dislodges the roadblock with probability P1,

f: an RNAP transcribing past the roadblock.

Overall, a polymerase taking the active path would follow states $a \to (b \leftrightarrow c)^n \to e \to f$ over n cycles of backtrack and repeat, where rates k_1 and k_2 would influence the transitions between states b and c (Fig. 27A). A RNAP utilizing a passive pathway might take route $a \to b \to d \to f$ (Fig. 27A). For the active and passive pathways, Jin Qian determined,

$$k_{active} = P1 \, \frac{k_1}{1+\frac{k_1}{k_2}},$$

$$k_{passive} = k_3.$$

The equation for k_{active} was fit from the data, and the equation of $k_{passive}$ was taken simply as dependent on a weaker roadblock's dissociation.[23] To capture our system of roadblocks and

experimental conditions, the following relationships between k_{active} and $k_{passive}$ were explored:

$$(1)\ k_{active} \sim k_{passive},$$

$$(2)\ k_{active} \ll k_{passive},$$

$$(3)\ k_{active} \gg k_{passive}.$$

The first instance (1) was presumed for stronger roadblock LacI-Os where it was predicted the active pathway would prove more efficient than the passive one. Model fitting to the experimental data compared favorably in predicting the longest average pause would occur in the opposing force, no-GreA condition, followed by roughly equal pause times in the assisting force conditions independent of GreA, and the shortest in the opposing force, GreA condition (Fig. 27B). The next case (2) was presumed for less strong roadblocks LacI-O1 and LacI-O2 where it was predicted the passive pathway would prove more efficient than the active one. Model fitting to LacI-O1 also predicated longest average pause for the opposing force, no-GreA condition followed by relatively equal pause times along the other instances (Fig. 27C). Finally, the last case (3) was speculated for extremely strong roadblocks, such as EcoR1 in the 50 mM KCl TBX. In this study, however, readthrough ratios instead of pause times were used as a comparison, as the pause times likely exceeded the observational duration. This model would expect to see fewer RNAPs transit through the obstacle under assisting forces than opposing forces, and experimental measurements validated this model's prediction (Fig. 26B).
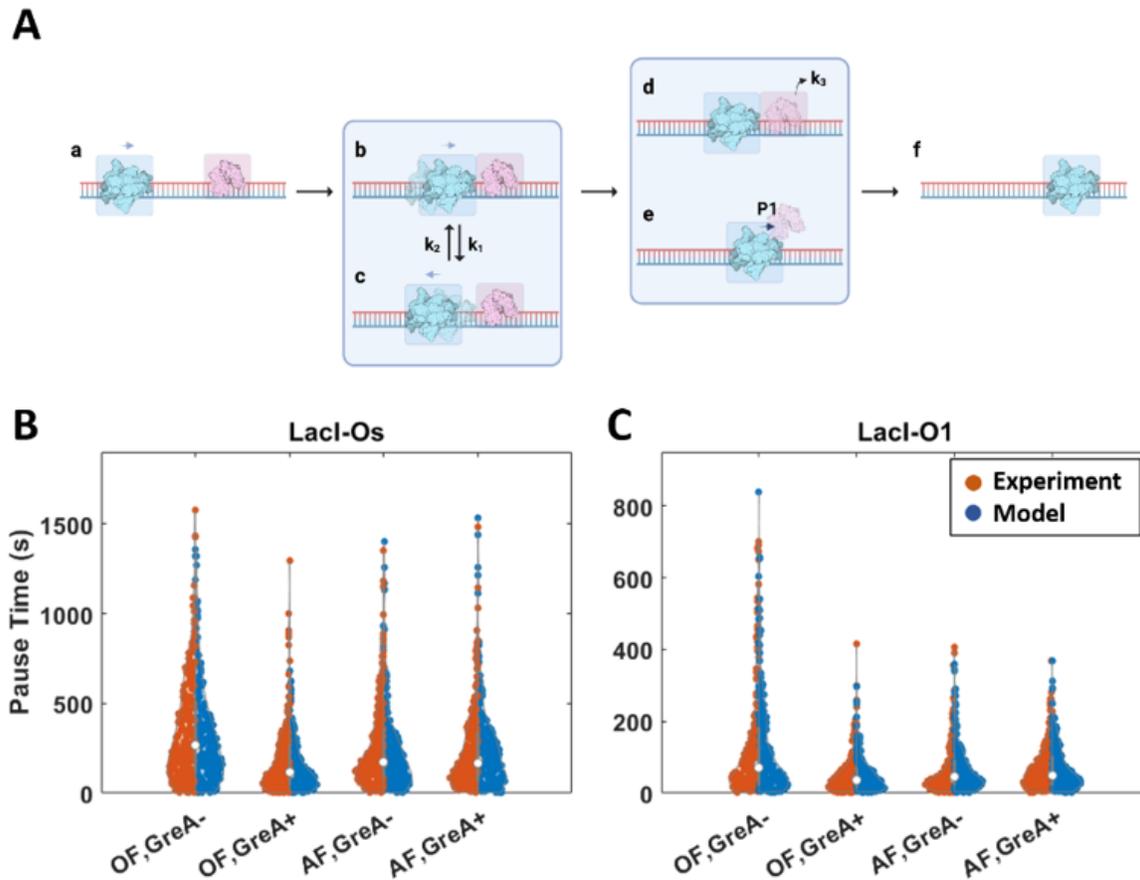
Figure 27. Hybrid Mechanism Schematic and Fitting Agreement.[23] (A) Pathways of states for a transcribing RNAP in blue against a roadblock in pink. (B) A comparison between pause times generated from LacI-Os experiments and the model. The model used $k_{active} \sim k_{passive}$ for a predominately active transit pathway. (C) A comparison between pause times generated for LacI-O1 experiments and the model. The model used $k_{active} \ll k_{passive}$ for a predominately passive transit pathway.

4.4 Future Directions

The findings drawn by this study regarding a dihybrid mechanism beget future work to explore and strengthen its conclusions. First, adding 500 mM KCl TBX to EcoR1 experiments successfully

dissociated EcoR1 at an appropriate frequency. However, it is uncertain whether, or to what extent, salt affected other components of the process, such as RNAP transcription, DNA twisting, or GreA action beyond EcoR1's dissociation. More importantly, unpublished data by collaborator Dr. Irina Artsimovich suggests LacI co-immunoprecipitates with RNAP which is why EcoR1 was explored as a non-interacting control. Though the EcoR1 and LacI-Os readthrough ratios compared favorably, the LacI RNAP interaction should be further characterized to determine whether it could disrupt this study's potentially universal mechanisms (Fig. 27). Finally, the difference between roadblock induced versus sequence induced backtracking should be explored, as this study demonstrated roadblock pausing independent of force magnitude but not direction. Additional structural support of RNAP's paused configuration might shed light on this question.

5 Chapter 5: Acknowledgements

6   Chapter 6: References

1.   Lu, Y. *et al.* Transcription disrupts DNA-scaffolded bacteriophage repressor complexes. doi:10.1101/2021.10.29.466366.

2.   Nikitina, T., Norouzi, D., Grigoryev, S. A. & Zhurkin, V. B. *DNA topology in chromatin is defined by nucleosome spacing*. https://www.science.org (2017) doi:10.1126/sciadv.1700957.

3.   Epshtein, V., Toulmé, F., Rahmouni, A. R., Borukhov, S. & Nudler, E. Transcription Through the Roadblocks: the Role of RNA Polymerase Cooperation. *EMBO J* **22**, 4719–4727 (2003).

4.   Nudler, E. RNA polymerase backtracking in gene regulation and genome instability. *Cell* vol. 149 1438–1445 Preprint at https://doi.org/10.1016/j.cell.2012.06.003 (2012).

5.   Ramji, D. P. & Foka, P. *CCAAT/enhancer-binding proteins : structure, function and regulation*. *Biochem. J* vol. 365 (2002).

6.   Lewis, D. E. A., Komissarova, N., Le, P., Kashlev, M. & Adhya, S. DNA Sequences in gal Operon Override Transcription Elongation Blocks. *J Mol Biol* **382**, 843–858 (2008).

7.   Garcia, H. G. & Phillips, R. Quantitative dissection of the simple repression input-output function. *Proc Natl Acad Sci U S A* **108**, 12173–12178 (2011).

8.   Gama-Castro, S. *et al.* RegulonDB (version 6.0): Gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids Res* **36**, (2008).

9.   Wang, L., Watters, J. W., Ju, X., Lu, G. & Liu, S. Head-on and co-directional RNA polymerase collisions orchestrate bidirectional transcription termination. *Mol Cell* (2023) doi:10.1016/j.molcel.2023.02.017.

10.   Bera, S. C. *et al.* Quantitative parameters of bacterial RNA polymerase open-complex formation, stabilization and disruption on a consensus promoter. *Nucleic Acids Res* **50**, 7511–7528 (2022).

11.   Vörös, Z., Yan, Y., Kovari, D. T., Finzi, L. & Dunlap, D. Proteins mediating DNA loops effectively block transcription. *Protein Science* **26**, 1427–1438 (2017).

12.   Qian, J., Xu, W., Dunlap, D. & Finzi, L. Single-molecule insights into torsion and roadblocks in bacterial transcript elongation. *Transcription* vol. 12 219–231 Preprint at https://doi.org/10.1080/21541264.2021.1997315 (2021).

13.   Basu, A., Bobrovnikov, D. G. & Ha, T. DNA mechanics and its biological impact. *Journal of Molecular Biology* vol. 433 Preprint at https://doi.org/10.1016/j.jmb.2021.166861 (2021).

14.   Verma, S. C., Qian, Z. & Adhya, S. L. Architecture of the Escherichia coli nucleoid. *PLoS Genet* **15**, (2019).

15.   Bai, L., Santangelo, T. J. & Wang, M. D. Single-molecule analysis of RNA polymerase transcription. *Annual Review of Biophysics and Biomolecular Structure* vol. 35 343–360 Preprint at https://doi.org/10.1146/annurev.biophys.35.010406.150153 (2006).

16.   Sutherland, C. & Murakami, K. S. An Introduction to the Structure and Function of the Catalytic Core Enzyme of Escherichia coli RNA Polymerase. *EcoSal Plus* **8**, (2018).

17.   Wang, M. D. *et al. Force and Velocity Measured for Single Molecules of RNA Polymerase*. www.sciencemag.org (1998) doi:10.1126/science.282.5390.902.

18.   Ren, B. *et al. Genome-Wide Location and Function of DNA Binding Proteins*. www.sciencemag.org.

19.   Sadler, J. R., Sasmor, H. & Betzt, J. L. *A perfectly symmetric lac operator binds the lac repressor very tightly (palindrome/protein-DNA interaction). Biochemistry* vol. 80 (1983).

20.   Sapienza, P. J., Rosenberg, J. M. & Jen-Jacobson, L. Structural and Thermodynamic Basis for Enhanced DNA Binding by a Promiscuous Mutant EcoRI Endonuclease. *Structure* **15**, 1368–1382 (2007).

21.   Fernández-Coll, L., Potrykus, K., Cashel, M. & Balsalobre, C. Mutational analysis of Escherichia coli GreA protein reveals new functional activity independent of antipause and lethal when overexpressed. *Sci Rep* **10**, (2020).

22.   Shaevitz, J. W., Abbondanzieri, E. A., Landick, R. & Block, S. M. *Backtracking by single RNA polymerase molecules observed at near-base-pair resolution Escherichia coli RNA polymerase (RNAP) synthesizes RNA with remarkable fidelity in vivo*. www.nature.com/nature (2003).

23.   Qian, J., Cartee, A. G., Dunlap, D., Artsimovitch, I. & Finzi, L. Transcription Through Roadblock Systems Reveals A Hybrid Transit Mechanism. doi:10.1101/2023.01.04.522798.

24.   Vilfan, I. D., Lipfert, J., Koster, D. A., Lemay, S. G. & Dekker, N. H. Magnetic Tweezers for Single-Molecule Experiments. in *Handbook of Single-Molecule Biophysics* 371–395 (Springer US, 2009). doi:10.1007/978-0-387-76497-9_13.

25.   Metcalfe, G., Speetjens, M. F. M., Lester, D. R. & Clercx, H. J. H. Beyond passive. Chaotic transport in stirred fluids. in *Advances in Applied Mechanics* vol. 45 109–188 (Academic Press Inc., 2012).

26.   Gosse, C. & Croquette, V. *Magnetic Tweezers: Micromanipulation and Force Measurement at the Molecular Level*. (2002).

27.   Duan, X. *et al.* Quantification of the affinities and kinetics of protein interactions using silicon nanowire biosensors. *Nat Nanotechnol* **7**, 401–407 (2012).

28.   Tetin, S. Y., Swift, K. M. & Matayoshi, E. D. *Measuring antibody affinity and performing immunoassay at the single molecule level*. www.academicpress.com.

29.   Cabrol, L., Quéméneur, M. & Misson, B. Inhibitory effects of sodium azide on microbial growth in experimental resuspension of marine sediment. *J Microbiol Methods* **133**, 62–65 (2017).

30.   Wufuer, R. *et al.* Distinct Roles of Nrf1 and Nrf2 in Monitoring the Reductive Stress Response to Dithiothreitol (DTT). *Antioxidants* **11**, (2022).

31.   Sarkar, R. & Rybenkov, V. V. A guide to magnetic tweezers and their applications. *Frontiers in Physics* vol. 4 Preprint at https://doi.org/10.3389/fphy.2016.00048 (2016).

32.   Piccolo, J. G., Harper, J. M., Kovari, D., Dunlap, D. & Finzi, L. Force spectroscopy with electromagnetic tweezers. (2021) doi:10.1063/5.0060276.

33.   Zilinskiene, J. *Thermo ScientificGeneJET PCR Purification Kit #K0701, #K0702*.

34.   Sidorova, N. Y., Scott, T. & Rau, D. C. DNA concentration-dependent dissociation of EcoRI: Direct transfer or reaction during hopping. *Biophys J* **104**, 1296–1303 (2013).