The Use of Single-Molecule DNA Nanomanipulation to Study Transcription Kinetics

By

CHEN-YU LIU

A dissertation submitted to the Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Computational Biology and Biophysics in

Written under the direction of

Dr. Richard H. Ebright

And Approved by

Dr. Richard H. Ebright

Dr. Ronald M. Levy

Dr. Wilma K. Olson

Dr. Richard D. Ludescher

New Brunswick, New Jersey

October, 2007

ABSTRACT OF THE DISSERTATION

The Use of Single-Molecule DNA Nanomanipulation to Study Transcription Kinetics

by CHEN-YU LIU

Dissertation Director:

Dr. Richard H. Ebright

Transcription involves many reaction steps and intermediates. Many phenomena in transcription kinetics are covered by ensemble average. Single-molecule DNA nanomanipulation techniques uncover these transcription kinetic events via determination of a transcription bubble in real time. In this dissertation, we focus on the transcription kinetics of bacterial RNAP. The study of transcription kinetics in this thesis can be divided into 4 main subjects: 1) The study of abortive initiation

mechanism: Through a single-molecule DNA nanomanipulation technique, we tested the three models proposed for the mechanism of abortive initiation - inchworming, scrunching and transient excursion - on T5 N25 promoter. Of the three models, only the scrunching model involves a change in the size of the transcription bubble during abortive initiation, which was observed by single-molecule DNA nanomanipulation technique. 2) The study of the kinetics of elongation and termination: By introducing varying transcribed region lengths into DNA templates, the kinetics of elongation and terminator rewinding were studied. The resulting kinetics, determined by the single-molecule nanomanipulation technique, was analyzed by different regression methods. In the normal regression, the elongation velocity is 10 b/s and terminator rewinding takes 3.4 s on a *tR2* terminator. Contrarily, through Poisson regression, the elongation velocity ranges from 6.4 b/s to 12.5 b/s and terminator rewinding takes 11.9 s on a *tR2* terminator. 3) The study of a promoter sequence's effect: The effect from sequence of the T5 N25 promoter and T5 N25 antiDSR promoter in transcription was evaluated. The transcription bubble sizes of open complex and initial transcribing complex using the T5 N25 antiDSR promoters are larger than the ones from the T5 N25 promoter. The difference in the two promoter sequences does not have an effect on the transcription bubble size of an elongation complex. And elongation and terminator rewinding kinetics are not

affected. On the other hand, abortive initiation and promoter escape are affected by the difference in the promoter sequence. 4) The study of the effect of transcription factor-GreB: The effect of transcription factor-GreB on abortive initiation was evaluated by the single-molecule DNA nanomanipulation technique. GreB does not affect the transcription bubble size during abortive initiation, but does reduce the lifetime of initial transcribing complex.

Abbreviations

ATP Adenosine triphosphate

bp Base pair

BSA Bovine serum albumin

CTP Cytidine triphosphate

Da Daltons

DNA Deoxyribonucleic acid

EDTA Ethylenediaminetetraacetic acid, sodium salt

GTP Guanosine triphosphate

HEPES N-[2-hydroxyethyl]-piperazine-N'-[2-ethane]-sulfonic acid

ITC Initial transcribing complex ITP

ITS Initial transcribed sequence

NMP Nucleoside monophosphate

nt Nucleotide(s)

NTP Nucleoside triphosphate

RNA Ribonucleic acid

RNAP RNA polymerase

UTP Uridine triphosphate

WLC model Worm-Like Chain Model

Acknowledgements

I am grateful for many people who assisted me to finish this dissertation. For the past several years, my work has become a part of my life. Many people who involved in my work also involved my work.

First, I would like to thank my advisor Dr. Richard Ebright. He supported me on many aspects from my daily life to study. He not only provided me the environment to study but also give me a lot of guidance on study. Without his support, I could not have such a great experience in life and study and my whole life will be quite different. I do not know how to express how grateful I am to him.

I would like to thank Dr. Yon Ebright. She solved many problems for me. Without her assistance, my study can not proceed.

I also thank Dr. Terence Strick. He taught me how to operate the single-molecule DNA nanomanipulation technique, which I used to study some interesting research objects. Also, he helped me to accommodate the life in Paris.

I thank Dr. Ronald Levy. Without his kindness, I could not join BioMaPS program and graduate smoothly.

I thank Dr. Richard Ludescher for spending time on correcting my dissertation and gave me the valuable comments.

I thank Dr. Wilma Olson. She also gave me some valuable comments on the dissertation.

I thank Jayanta Mukhopadhyay being kind and generous to me always.

I thank Charles O'Brien for being kind to provide any kind of assistances from correction of thesis, giving me accommodation, sharing experience with me…

I thank Dongye Wang, Kerri-Ann Norton, Jianxi Xiao, Hongjun Zhang and David Degen, for assistance in my dissertation and making my life more interesting during the past several years.

In the end, I would like to thank my family- my mother and my sister- for your

patience and support.

To the memory of my father,

Ming Liu

Contents

Abstract……ii

Abbreviations……v

Acknowledgements……...vi

Dedication……..ix

Table of contents……..……..x

Chapter 1

Introduction

1.1 RNA Polymerase ……1

1.1.1 Promoter……2

1.1.2 Introduction of Structure……...3

1.2. Transcription……4

1.2.1 Promoter search……..4

1.2.2 Initiation……..5

1.2.3 Elongation……..10

1.2.3.1 Pause state…….....12

1.2.3.2 Two functional Magnesium ions…….....13

1.2.4 Kinetics……..13

1.2.4.1 The thermal stability of TEC…….14

1.2.5 Termination……15

Summary……19

Reference……20

Chapter 2

Methods for the study of transcription with single-molecule technique

2.1 Previous studies of the transcription kinetics with the single-molecule technique……..27

2.2 Atomic force microscopy (AFM)……...27

2.3 Tethered Particle Motion Method……...28

2.4 Flow-control video microscopy…….29

2.5 Optical tweezers with rotation detection……30

2.6 The combination of total internal reflection fluorescence (TIRF) with optical trapping…….32

2.7 Dual optical tweezers…….33

Summary……..34

References……35

Chapter 3

Single-molecule DNA nanomanipulation

3.1 Experimental setup…….37

3.2 Preparation of DNA……...39

3.3 Determination of the position of a bead…….40

3.4 Force measurement and Manipulation……...42

3.5 Determination of DNA extension and force……...44

3.6 Stretching of torsionally free DNA……46

3.6.1 Stretching of supercoiled DNA……...47

Reference……..51

Chapter 4

Detection of initial transcription by single-molecule DNA nanomanipulation

4.1 Background and Objectives……53

4.2 Experimental design and methods……..54

4.2.1 Construct of N25 promoter DNA templates……55

4.2.3 Determine abortive initiation kinetics with the subset of NTPs..56

4.2.4 Basic experiment setup……56

4.2.5 Data analysis……...58

4.3 Results……58

4.4 Conclusions……60

References……61

Chapter 5

Detection of promoter escape by single-molecule DNA nanomanipulation

5.1 Background and Objectives……62

5.2 Experimental design and methods……..62

5.3 Results……63

5.4 Conclusions……65

5.4.1 Initiation-elongation transition is within the time resolution of nanomanipulation……66

References……66

Chapter 6

Detection of elongation and termination by single-molecule DNA nanomanipulation

6.1 Background and Objectives……68

6.2 Experimental design and methods……..68

6.3 Results……69

6.4 Conclusions……70

Chapter 7

Determination of kinetics of elongation and termination

7.1 Background and Objectives……72

7.1.1 Statistical model……..74

7.2 Experimental design and methods…….77

7.3 Results……78

7.3.1 Data analysis ……..78

7.3.1.1 The analysis based on linear regression……...78

7.3.1.2 The analysis based on Poisson Statistics……..80

7.5 Conclusions……81

References……82

Chapter 8

Determination of effects of promoter sequence on initial transcription, promoter escape, elongation, and termination

8.1 Background and Objectives…….84

8.2 Experimental design and methods……...85

8.3 Results……..87

8.3.1 The open complex on the N25 promoter vs. the N25 $_{antiDSR}$ promoter……..87

8.3.2 Transcription initiation on the N25 promoter and the $N25_{antiDSR}$ promoter …….88

8.3.3 Elongation and termination on the N25 promoter and $N25_{antiDSR}$ promoter……..91

8.4 Conclusions……...93

Reference ……94

Chapter 9

Determination of the effects of elongation factor GreB on initial transcription, promoter escape, elongation, and termination

9.1 Background and Objectives……..95

9.2 Experimental design and methods……95

9.3 Results……...96

9.3.1 Abortive initiation with/without GreB……96

9.3.2 Elongation and termination with and without GreB……...99

9.4 Conclusions……..100

References……..100

Chapter 10

Materials and Methods

10.1 The construction of plasmid DNA templates…….101

10.2 The construction of DNA fragment for nanomanipulation……102

10.3 Preparation of the Capillary……...104

10.4 Calibration and Data Acquisition……...……105

10.5 Transcription Assay……106

10.6 Buffer used in the experiments……...107

Appendix:

Appendix A: N25-160-tr2 time trace data for a complete transcription event……...108

Appendix B: N25-220-tr2 time trace data for a complete transcription event……...116

Appendix C: N25-280-tr2 time trace data for a complete transcription event……...118

Appendix D: N25-340-tr2 time trace data for a complete transcription event…….122

Appendix E: N25-100-tr2 without GreB time trace data for a complete transcription event…….126

Appendix F: N25antiDSR-160-tr2 time trace data for a complete transcription event……..127

Curriculum Vita

Chapter 1

Introduction

1.1 RNA Polymerase

RNAP (RNA polymerase) plays an essential role in gene expression; it can transcribe DNA sequences into RNA that can be translated into proteins that carry out many cellular functions. RNAP can be categorized into two types: multiple subunit RNAPs and single subunit RNAPs. Multiple subunit RNAPs are highly conserved in their sequences and structures and are found in prokaryote or eukaryote cells. By contrast, single subunit RNAPs are found among bacteriophages or mitochondria [1]. Although the sequences of the multiple subunit RNAP family are quite different from the sequences of the single subunit RNAP family, they still have very similar functions and structure. Many of the studies of multiple subunits RNAPs have relied on studies from the single subunit RNAP family.

Many papers have focused on *Escherichia coli* RNAP, even though the structure of *Escherichia coli* RNAP is not available yet. Most studies of the RNAP

structure – function relationship of *Escherichia coli* RNAP rely on the X-ray structure of *Thermus aquaticus* RNAP and *Thermus thermophilus* RNAP. This is because the structures of RNAPs of the two species have high similarity with *Escherichia coli* RNAP, so they can be predicted to have similar structure-function relationship.

Bacterial RNAP core enzyme is compose of five subunits: β, β',α^{I},α^{II} and ω. In order to carry out transcription on DNA, core RNAP needs to combine with a σ (sigma factor) to recognize a promoter DNA sequence and start transcription. A core RNAP incorporating with a sigma factor is called a holoenzyme. Different sets of sigma factors can initiate different sets of gene expression that help the cell accommodate to its environment.

Bacterial sigma factors can be classified into two categories- one is σ^{70} and other is σ^{54}. σ^{70} associates with most house keeping genes of *Escherichia coli* and is well characterized. The σ factor can be divided into four regions-σ1, σ2, σ3 and σ4 with flexible linkers in between some regions.

1.1.1 Promoter

A promoter is a DNA sequence by which an RNAP can initiate transcription. The promoter sequence can be composed of two hexamer elements : – 10 element and -35 element. The -10 and -35 refer to the center of two hexamer elements located at upstream 10 and 35 nucleotide of transcription start site [2,3]. Between the two elements is a spacer around 17 bp which can also determine the binding capability of RNAP on a promoter by its length [4,5]. In addition, a subclass of promoters lacks a –35 element but has an extra extended –10 element to assist RNAP binding [6,7]. For these promoters, they can also have an upstream element between -40 and -60 which can enhance the binding of RNAP and promoter by interaction with α subunit of RNAP [8-10].

1.1.2 Introduction of Structure

The crab-claw shape like multiple subunit bacterial RNAP holoenzyme has dimensions of ~150 Å x ~100 Å x ~100 Å. The beta and beta primer domain form two pincers of the claw. The active site is the cleft between the two pincers. The structure of RNAP holoenzyme is shown in (Figure 1-1).

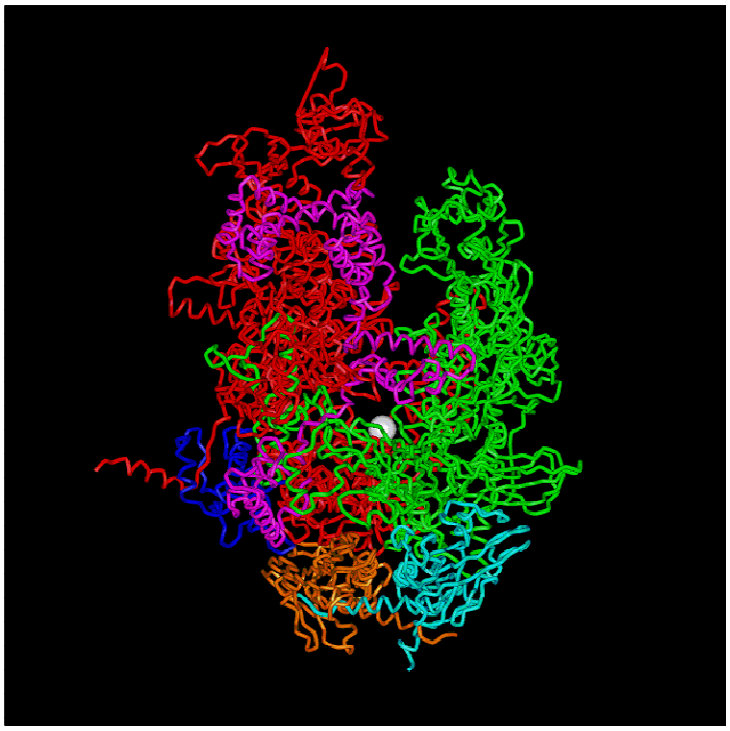


Figure 1-1.
The Crystal structure of *Thermus aquaticus* RNAP Holoenzyme:
The α Subunits: Orange and Light Blue, the β Subunit: Green, the β' Subunit: Red, the ω Subunit: Blue, the σ Subunit: Magenta. The magnesium ion in the active site: White. The beta and beta primer domain form two pincers of the claw. The active site is the cleft between the two pincers [11].

1.2. Transcription

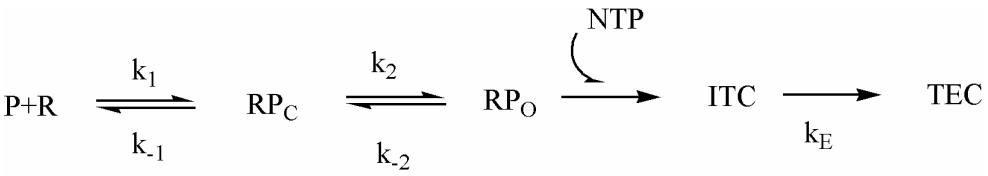
Generally speaking, the transcription process can be divided: promoter search, initiation, elongation and termination.

1.2.1 Promoter search

Since RNAP is responsible for transcribing genes in vivo, it needs to bind onto a promoter. Through diffusion, RNAP can target a promoter in solution. There are different modes of diffusion that can take place, these include 1D diffusion-sliding along DNA and 3D diffusion - hopping and intersegment transfer. It has been demonstrated that RNAP can slide along DNA [12], however, 1D diffusion may not be an efficient way for RNAP carrying out transcription of a gene because thermal diffusion has a probability to go in the reverse or the forward direction. Moreover, 1D diffusion works for short range movement. On the other hand, the protein-ligand association rate is on the order of $1\times10^{8} M^{-1}s^{-1}$ which is close to the diffusion limit [13]. If 3D diffusion is the only mechanism for a protein to target a specific site on a DNA, the protein-ligand association rate would be much lower than $1\times10^{8} M^{-1}s^{-1}$ [14]. Therefore, the promoter search of RNAP could be achieved by both 3D and 1D diffusion among DNA [15,16]. At low salt condition, a protein may only traverse along the DNA, while in vivo, it may have two diffusion modes [17]. The RNAP searches a promoter through 1D and 3D diffusion and when it finds the promoter site, initiation starts.

1.2.2 Initiation

The kinetic scheme of RNAP transcription initiation is shown in scheme 1-1.



Scheme 1-1

When the RNAP holoenzyme (R) arrives at a promoter (P); it can associate with a promoter, forming a closed complex- RPc, it does this at rate k_1. The closed complex can dissociate from the promoter, which happens at rate k_{-1}. Thus, the formation of closed complex is a reversible reaction and can be characterized by K_B (k_1/k_{-1}) on the order of $3\times10^{7} M^{-1}s^{-1} \sim 3\times10^{8} M^{-1}s^{-1}$ [18]. Following the formation of closed complex, σ2 region on the holoenzyme nucleates DNA unwinding at -10 element of a promoter and separates the double stranded DNA from -11/-12 to +3/+4, isomerizing the closed complex into an open complex- RPo [19]. This step can be described with kinetic rate of k_2. Usually, for some strong promoters, the open complex formation is very stable, the reverse reaction being negligible (k_{-2} is very small.). The formation of open complex causes the change of RNAP

holoenzyme conformation. The DNA enters the two pincers of RNAP. The two pincers then close down, clamping DNA in place. The DNA between position -11 to + 3 unwinds and the strands separate. Then each strand passes through the channels formed by β and β' domain [20]. The σ1.1 domain- a DNA mimic- is displaced from the DNA binding domain which is formed by the two pincers [21,22]. A model of the structure of open complex of *Thermus thermophilus* is shown in Figure 1-2.

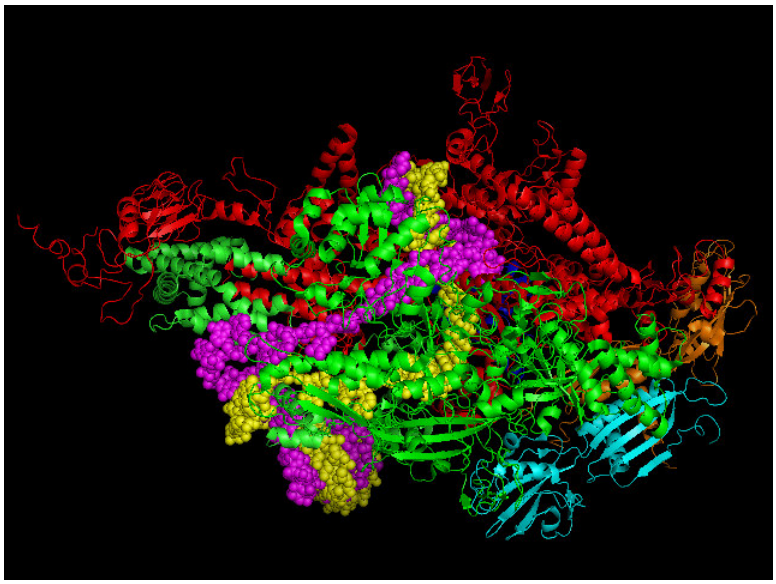


Figure 1-2
The model structure of *Thermus aquaticus* RNAP Holoenzyme: The template DNA: Yellow, Magenta, the α Subunits: Orange and Light Blue, the β Subunit: Green, the β' Subunit: Red, the ω Subunit: Blue. The two pincers of β and β' clamp and bend the promoter DNA, the promoter DNA is separated into two single strands of DNA [23].

Upon the formation of open complex, binding of NTPs which are complementary to the initiating sequence can further stabilize the open complex

without formation of phosphodiester bond [24]. With a supply of NTPs, the RNAP can start initiation transcription.

When transcription initiation begins, RNAP enters into a cycle of synthesis of nascent RNA fragments and aborts RNA products from 2 nt~10 nt long(or up to 15 nt at some promoters) until the RNAP escapes from the promoter- a stage called abortive initiation state. The rate in which RNAP leaves from the abortive initiation state is characterized as k_E (scheme 1-2). The complex is called initial transcribing complex (ITC). Not all ITC can generate full length of RNA transcripts, only some of these ITCs can transcribe full length of RNA after many cycles of transcribing and releasing nascent RNA [27,28]. NTPs diffuse through the secondary channel into the active site of RNAP, annealing with the template DNA sequence. Afterward the pyrophosphate group on the NTP is hydrolyzed and the NMP is incorporated into the RNA transcript. This active center will keep moving relative to the template DNA and incorporating nucleotides on the RNA products. As the length of the nascent RNA transcript increases, it encounters the linker between σ3 and σ4 [25]. Either the linker of σ3 and σ4 releases from the RNA exit channel or the RNAP releases the nascent RNA through secondary channel [26].

Many factors can affect or regulate the transcription at the abortive initiation stage like the sequence of a promoter, transcription factors, concentration of NTP and concentration of salt, etc. These factors may also play a role in gene regulation. For instance, by transcribing with ITS (initial transcribed sequence), RNAP can produce various lengths of nascent RNA products with different abortive probability and productive ratio. The productive ratio can be increased by the presence of some cleavage factors like GreA and GreB or by an increase in NTPs concentration [29-32].

There have been three proposed mechanisms for RNAP starting abortive initiation- the inchworming model, transient excursion model and scrunching model in Figure 1-3 [33-35].

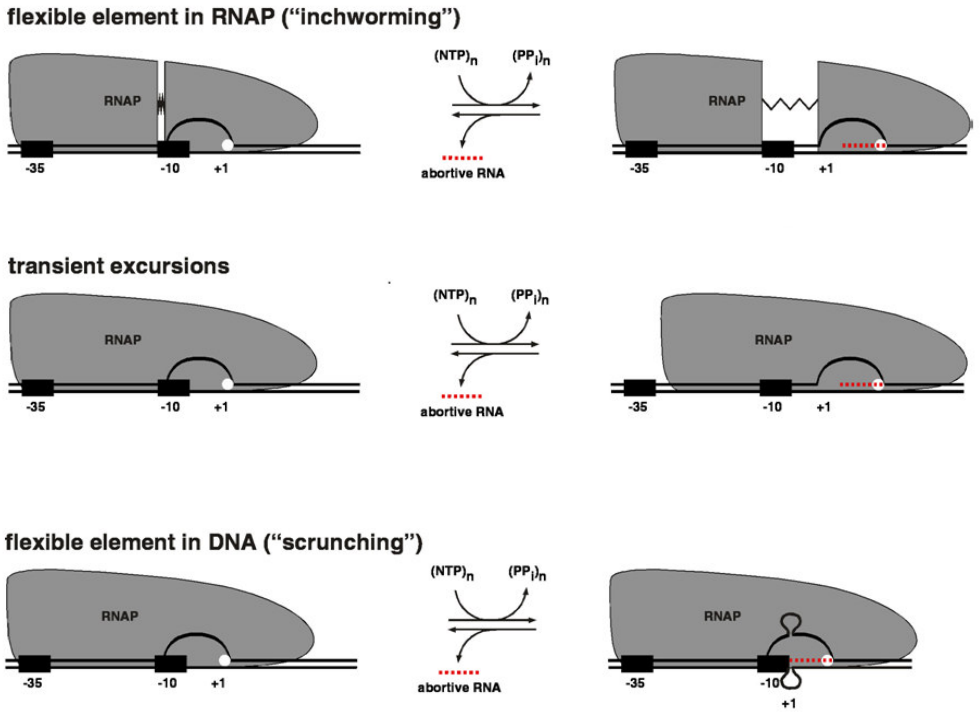


Figure 1-3
The proposed mechanisms for abortive initiation and promoter escape: Figure from reference [36].

These models are based on the observations that abortive initiation nascent RNA has a length of 9~10nt and that upstream boundary DNA is protected during abortive initiation. That is found from the footprint experiments. The inchworming model states that abortive initiation is involved in the movement of RNAP active center without release of the binding of RNAP at -35 element while the RNAP synthesizes nascent RNA. The transient excursion model proposes that RNAP synthesizes abortive RNA transcripts and leaves promoter, aborting the RNA transcript and moving back to the promoter. As for the scrunching model, it

proposes that the DNA may scrunch into RNAP without breaking the upstream contacts of the RNAP with the promoter during abortive initiation [37].

1.2.3 Elongation

After RNAP escapes from a promoter and starts elongation, RNAP can extend the RNA product when entering elongation stage [38]. A single NTP is brought into the active site through the secondary channel and binds at the substrate binding sub-site next to the 3' terminal of RNA. The RNAP removes the pyrophosphate from the NTP and incorporates the NMP into the RNA by the formation of phosphodiester bond with the 3' terminal of the RNA. Afterward, the translocation of 3' terminal of the RNA is triggered by the G trigger loop and the bridge-helix of RNAP [39]. A DNA base pair rewinds at the upstream of the transcription bubble and a DNA base pair is flipped out on the downstream of the bubble when the translocation of the RNAP occurs (Figure1-4). After the translocation of transcription bubble, the substrate binding sub-site is empty again for the next coming NTP. The single nucleotide addition cycle will continue until RNAP terminates RNA synthesis [40].

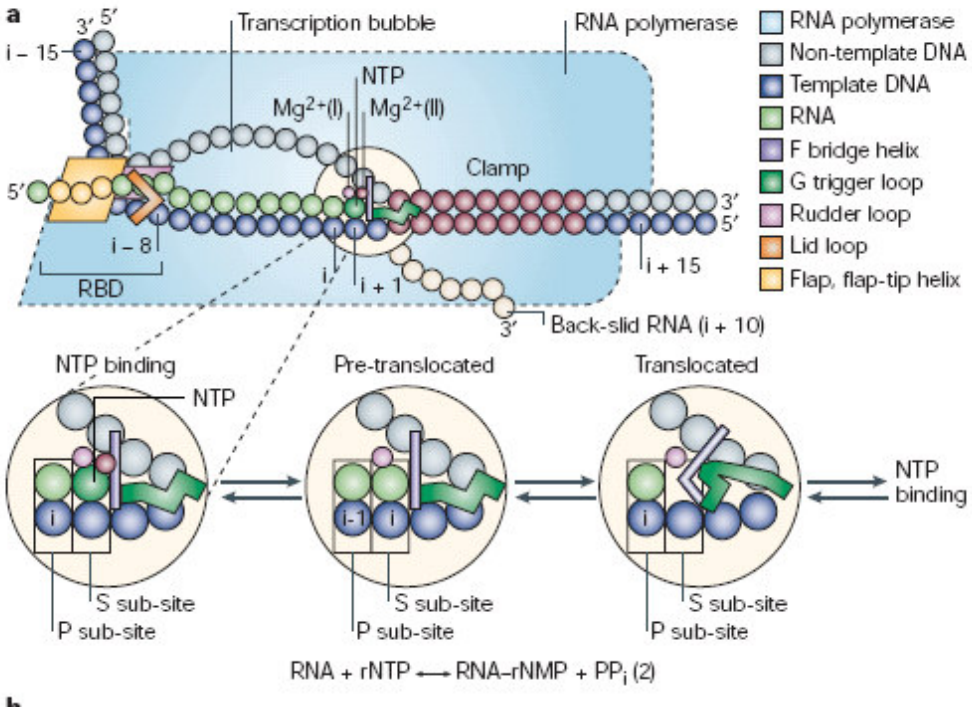


Figure 1- 4

The single nucleotide binds at the substrate binding sub-site followed by the hydrolysis of NTP. Afterward, the G loop triggers the bending of the bridge –helix, and the RNAP translocates forward, starting next cycle of RNA synthesis. Figure from ref [41].

1.2.3.1 Pause state

In the elongation state, RNAP sometimes stalls at template DNA for several seconds to one minute, which is called pause. These pause events in the RNAP elongation can be put into two categories- Class I pauses and Class II pauses. The Class I pauses are caused by a RNA hairpin. Thus, the Class I pauses are resistant to GreA and pyrophosphate [42]. Addition of an oligonucleotide which is

complimentary to the pause hairpin can reduce Class I pauses. The Class II pauses are caused by weak DNA:RNA hybrids, which are sensitive to GreA and pyrophosphate. The Class II pauses may result in the backtracking of RNA transcript of TEC. And the Class II pauses are also called an arrested state [43]. When the backtracking of RNA transcript occurs, the RNA protrudes downstream from the active site into secondary channel. The protruded RNA can be cleaved by transcription factors like GreA and GreB.

1.2.3.2 Two functional Magnesium ions

The two functional Mg^{2+} need to be bound to the active site of RNAP. MgI is tightly bound to the active site but not MgII. Without MgII the RNA can not be synthesized properly. The GreA and GreB serve a role in helping chelation of the MgII in the active site of RNAP.

1.2.4 Kinetics

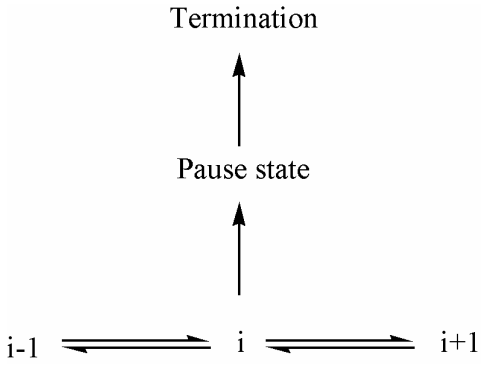
Both a power stroke model and a Brownian ratchet model have been proposed for transcription elongation [44]. The power stroke model emphasizes that the conformation change of RNAP in elongation is involved in the translocation of the

RNAP along the DNA, which converts the energy of NTP incorporating into RNA transcript to the mechanic driving force of RNAP translocation [45].

On the other hand, the Brownian ratchet model proposes that TEC (transcription elongation complex) can slide along DNA by thermal fluctuation and is biased by nucleotide incorporation.

In the Brownian ratchet model, the transcription kinetics for each template position of RNAP can be regarded as a competition among various reaction pathways (Scheme 1-2). RNAP must reach a state at which RNAP can choose to enter a pause/backtracking state, pyrophosphorolysis or an editing state.

The energy difference between these states can be small. Therefore, an increase of the pyrophosphate concentration may reverse the elongation reaction. However, shortening a RNA product does require 100 fold increase of the concentration of pyrophosphate in vivo [46].



Scheme 1-2

1.2.4.1 The thermal stability of TEC

The thermal stability of TEC can be applied in the ratchet model quantitatively. The thermodynamic stability of TEC can be expressed as the sum of interaction energy between DNA, RNA and protein. The $\Delta G^{\circ}_{(bubble)}$ is responsible for opening the double stranded DNA. The $\Delta G^{\circ}_{(hybrid)}$ represents the stabilization from an 8-9 bp DNA-RNA hybrid [47]. As for the $\Delta G^{\circ}_{(RNA\ polymerase)}$, it comes from the interaction between the RNAP and the DNA or the RNA [48]. By an assumption of the $\Delta G^{\circ}_{(RNA\ polymerase)}$ being constant during elongation, the variation of thermal stability of TEC along the template DNA can be determined by the $\Delta G^{\circ}_{(bubble)}$ and the $\Delta G^{\circ}_{(hybrid)}$, which determines the kinetic behavior of transcription in a manner of sequence dependence like termination and backtracking [49-51].

This assumption of a constant $\Delta G^{\circ}_{(\text{RNA polymerase})}$ is based upon two factors: the first one is nonspecific interaction of the RNAP with the DNA and the RNA and the second one is that two terms: $\Delta G^{\circ}_{(\text{bubble})}$ and $\Delta G^{\circ}_{(\text{hybrid})}$ have to be comparable with $\Delta G^{\circ}_{(\text{RNA polymerase})}$.

1.2.5 Termination

The transcription termination can occur with a termination factor or intrinsically. The transcription termination which requires termination factors can be divided as Rho-dependent transcription termination or Mfd-dependent transcription termination. Rho-dependent transcription termination relies on an ATP driven termination factor called Rho factor. Rho factor is a RNA dependent ATPase and RNA translocase which can unwind the DNA-RNA hybrid and terminate the synthesis of RNA [52,53]. By loading onto the cytosine rich site at the RNA transcript and moving in 5' to 3' direction, the Rho factor can catch up when the TEC stalls at pause site [54]. The Rho factor pulls out the RNA transcript from the TEC, and the RNAP dissociates from the DNA.

Mfd-dependent transcription termination also depends on an ATP driven

termination factor, which is called Mfd ATP dependent DNA translocase. Mdf can rescue a backtrack or arrested complex into TEC in the presence of full set of NTP. In the absence of NTP, Mdf can cause the dissociation of RNAP or the release of transcript [55]. Mdf has two domains related to its RNAP release activity: a RNAP binding domain and a DNA translocase domain. The DNA translocase domain contacts with around 25 bp of the upstream duplex DNA of transcription bubble. A study has suggested that Mfd-RNAP complex may exert force on transcription bubble, which causes transcription bubble move forward [56].

In the case of intrinsic termination, when RNAP arrives at a particular DNA sequence called a terminator, the RNAP is likely to stop elongation and release itself from the template DNA otherwise the RNAP may run off the terminator. A terminator has an invert repeat sequence and a poly A sequence. The invert repeat sequence can be transcribed into a hairpin and the sequence is followed by a poly A tract which triggers transcription termination. Different terminators have different efficiencies in termination of RNA synthesis [57].

In the case of deleting half of the invert repeat sequence of a terminator, the termination event will not occur while RNAP transcribes the modified terminator.

However, upon addition of complimentary oligonucleotide, the termination occurs by annealing of the oligonucleotide with the RNA transcript. The annealing of the oligonucleotide and the RNA transcript functions like a termination hairpin.

Two models have been proposed for termination mechanism: forward translocation model [58], allosteric model [59,60](Figure 1-5).

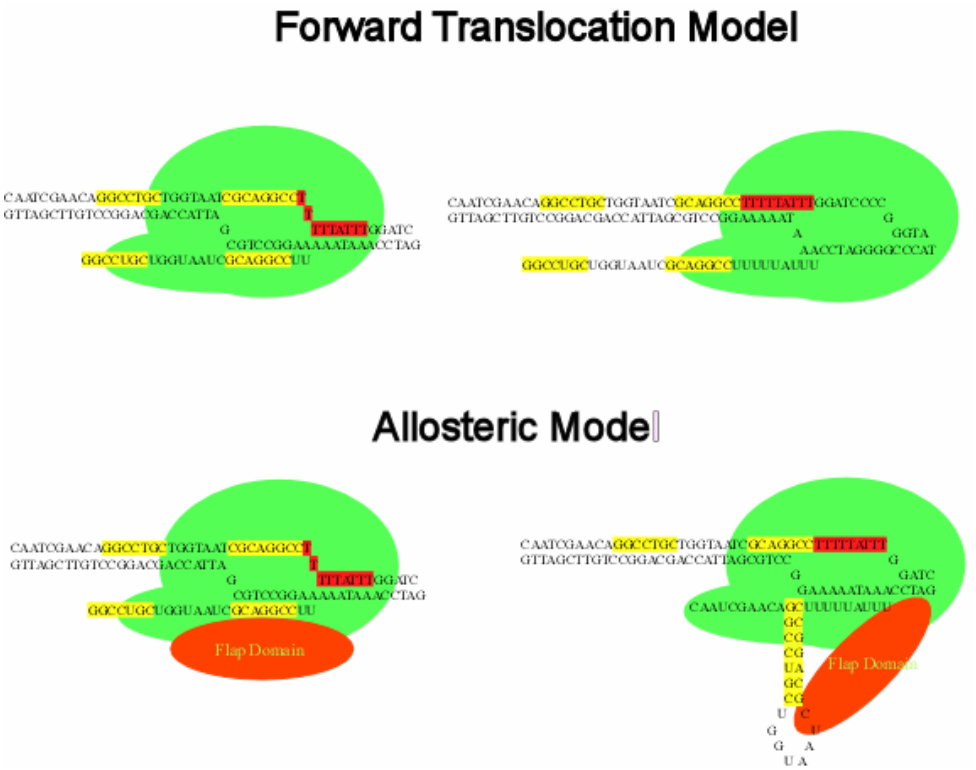


Figure 1-5
In the forward translocation model, termination occurs when RNAP stops RNA synthesis and the transcription bubble translocates forward. In the allosteric model, the termination hairpin causes the long range conformation change of RNAP, which terminates the synthesis of RNA.

In the forward translocation model, when transcription termination event occurs,

RNAP stops incorporating nucleotides into RNA transcript and the upstream DNA of transcription bubble rewinds and the downstream of transcription bubble unwinds. The transcription bubble moves forward. Since in the model the unwinding of the downstream of transcription bubble requires a torque, termination factor may play a role in providing the torque. From a recent study, the substitution of nucleotide on the non template DNA will affect the Rho and Mdf to mediate RNA release [61]. It indicates that the substitutions prevent rewinding of the transcription bubble by the termination factors.

In the case of intrinsic termination, the DNA rewinding torque can be generated by the formation of a termination hairpin or the annealing of oligonucleotide with the RNA transcript of a terminator in which the half stem sequence of a termination hairpin has been removed. It was found that blocking downstream double stranded DNA of a terminator can prevent termination [62]. The termination can be caused by altering the sequence of non template DNA of a terminator or the upstream sequence of the non template DNA of a terminator [63,64].

In the allosteric model, the transcription termination involves the long range conformation change of RNAP by some terminator factors to trigger the release of the

RNA transcript from TEC [65,66].

Summary

RNAP can carry out transcription. The stages of transcription can be divided into promoter search, initiation, elongation and termination. In the initiation stage, RNAP forms a closed complex at a promoter. Afterward, the closed complex forms an open complex, which is a complex having double stranded promoter DNA separated by RNAP. Upon providing NTPs, the open complex can transform to ITC. The ITC can transforms into TEC while the RNAP escapes from the promoter. The efficiency of promoter escape can be determined by the ITS. During elongation, RNAP may be trapped in an arrested state and pause state. The kinetics of elongation can be described in terms of the thermal stability of TEC, which is affected by the sequence of transcribed region, the concentration of nucleotides and the transcription factors.

References

1. Jeruzalmi D. and Steitz T.A. (1998). Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme**.** EMBO J. *17***:** 4101-4113.

2. Hawley D.K. and McClure W.R. (1983). Compilation and analysis of Escherichia coli promoter DNA sequences**.** Nucleic Acids Res. *11***:** 2237-2255.

3. Harley C.B. and Reynolds R.P. (1987). Analysis of E. coli promoter sequences**.** Nucleic Acids Res. *15***:** 2343-2361.

4. Mulligan M.E., Brosius J., and McClure W.R. (1985). Characterization in vitro of the effect of spacer length on the activity of Escherichia coli RNA polymerase at the TAC promoter**.** J. Biol. Chem. *260***:** 3529-3538.

5. Stefano J.E. and Gralla J.D. (1982). Mutation-induced changes in RNA polymerase-lac ps promoter interactions**.** J. Biol. Chem. *257***:** 13924-13929.

6. Dombroski A.J. (1997). Recognition of the -10 promoter sequence by a partial polypeptide of sigma70 in vitro**.** J. Biol. Chem. *272***:** 3487-3494.

7. Sabelnikov A.G., Greenberg B., and Lacks S.A. (1995). An extended -10 promoter alone directs transcription of the DpnII operon of Streptococcus pneumoniae**.** J. Mol. Biol. *250***:** 144-155.

8. Landini P. and Volkert M.R. (1995). Transcriptional activation of the Escherichia coli adaptive response gene aidB is mediated by binding of methylated Ada protein. Evidence for a new consensus sequence for Ada-binding sites**.** J. Biol. Chem. *270***:** 8285-8289.

9. Rao L., Ross W., Appleman J.A., Gaal T., Leirmo S., Schlax P.J., Record M.T., Jr., and Gourse R.L. (1994). Factor independent activation of rrnB P1. An "extended" promoter with an upstream element that dramatically increases promoter strength**.** J. Mol. Biol. *235***:** 1421-1435.

10. Ross W., Gosink K.K., Salomon J., Igarashi K., Zou C., Ishihama A., Severinov K., and Gourse R.L. (1993). A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase**.** Science *262***:** 1407-1413.

11. Vassylyev, Sekine, Laptenko, Lee, Vassylyeva, Borukhov and Yokoyama. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 angstrom resolution Nature. 417: 712-719

12. Kabata H., Kurosawa O., Arai I., Washizu M., Margarson S.A., Glass R.E., and Shimamoto N. (1993). Visualization of single molecules of RNA polymerase

sliding along DNA. Science *262*: 1561-1563.

13. Halford S.E. and Marko J.F. (2004). How do site-specific DNA-binding proteins find their targets? Nucleic Acids Res. *32*: 3040-3052.

14. von Hippel P.H. and Berg O.G. (1989). Facilitated target location in biological systems. J. Biol. Chem. *264*: 675-678.

15. Gerland U., Moroz J.D., and Hwa T. (2002). Physical constraints and functional characteristics of transcription factor-DNA interaction. Proc. Natl. Acad. Sci. U. S. A *99*: 12015-12020.

16. Kabata H., Kurosawa O., Arai I., Washizu M., Margarson S.A., Glass R.E., and Shimamoto N. (1993). Visualization of single molecules of RNA polymerase sliding along DNA. Science *262*: 1561-1563.

17. Gowers D.M., Wilson G.G., and Halford S.E. (2005). Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. Proc. Natl. Acad. Sci. U. S. A *102*: 15883-15888.

18. Lonetto M., Gribskov M., and Gross C.A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. J. Bacteriol. *174*: 3843-3849.

19. Murakami K.S., Masuda S., Campbell E.A., Muzzin O., and Darst S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science *296*: 1285-1290.

20. Kuznedelov K., Korzheva N., Mustaev A., and Severinov K. (2002). Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. EMBO J. *21*: 1369-1378.

21. Wilson C. and Dombroski A.J. (1997). Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. J. Mol. Biol. *267*: 60-74.

22. Vuthoori S., Bowers C.W., McCracken A., Dombroski A.J., and Hinton D.M. (2001). Domain 1.1 of the sigma(70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. J. Mol. Biol. *309*: 561-572.

23. Vassylyev D.G. and Artsimovitch I. (2005). Tracking RNA polymerase, one step at a time**.** Cell *123***:** 977-979.

24. Gaal T., Bartlett M.S., Ross W., Turnbough C.L., Jr., and Gourse R.L. (1997). Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria**.** Science *278***:** 2092-2097.

25. Murakami K.S., Masuda S., and Darst S.A. (2002). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution**.** Science *296***:** 1280-1284.

26. Chan C.L. and Gross C.A. (2001). The anti-initial transcribed sequence, a portable sequence that impedes promoter escape, requires sigma70 for function**.** J. Biol. Chem. *276***:** 38201-38209.

27. Susa M., Sen R., and Shimamoto N. (2002). Generality of the branched pathway in transcription initiation by Escherichia coli RNA polymerase**.** J. Biol. Chem. *277***:** 15407-15412.

28. Sen R., Nagai H., and Shimamoto N. (2001). Conformational switching of Escherichia coli RNA polymerase-promoter binary complex is facilitated by elongation factor GreA and GreB**.** Genes Cells *6***:** 389-401.

29. Gralla J.D., Carpousis A.J., and Stefano J.E. (1980). Productive and abortive initiation of transcription in vitro at the lac UV5 promoter**.** Biochemistry *19***:** 5864-5869.

30. Feng G.H., Lee D.N., Wang D., Chan C.L., and Landick R. (1994). GreA-induced transcript cleavage in transcription complexes containing Escherichia coli RNA polymerase is controlled by multiple factors, including nascent transcript location and structure**.** J. Biol. Chem. *269***:** 22282-22294.

31. Hsu L.M., Vo N.V., and Chamberlin M.J. (1995). Escherichia coli transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression in vivo and in vitro**.** Proc. Natl. Acad. Sci. U. S. A *92***:** 11588-11592.

32. Hsu L.M., Vo N.V., Kane C.M., and Chamberlin M.J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters**.** Biochemistry *42***:** 3777-3786.

33. Straney D.C. and Crothers D.M. (1987). Comparison of the open complexes

formed by RNA polymerase at the Escherichia coli lac UV5 promoter**.** J. Mol. Biol. *193***:** 279-292.

34. Krummel B. and Chamberlin M.J. (1989). RNA chain initiation by Escherichia coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes**.** Biochemistry *28***:** 7829-7842.

35. Spassky A. (1986). Visualization of the movement of the Escherichia coli RNA polymerase along the lac UV5 promoter during the initiation of the transcription**.** J. Mol. Biol. *188***:** 99-103.

36. Revyakin A., Liu C., Ebright R.H., and Strick T.R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching**.** Science *314***:** 1139-1143.

37. Spassky A. (1986). Visualization of the movement of the Escherichia coli RNA polymerase along the lac UV5 promoter during the initiation of the transcription**.** J. Mol. Biol. *188***:** 99-103.

38. Mooney R.A., Darst S.A., and Landick R. (2005). Sigma and RNA polymerase: an on-again, off-again relationship? Mol. Cell *20***:** 335-345.

39. Tuske S., Sarafianos S.G., Wang X., Hudson B., Sineva E., Mukhopadhyay J., Birktoft J.J., Leroy O., Ismail S., Clark A.D., Jr. et al. (2005). Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation**.** Cell *122***:** 541-552.

40. Greive S.J. and von Hippel P.H. (2005). Thinking quantitatively about transcriptional regulation**.** Nat. Rev. Mol. Cell Biol. *6***:** 221-232.

41. Greive S.J. and von Hippel P.H. (2005). Thinking quantitatively about transcriptional regulation**.** Nat. Rev. Mol. Cell Biol. *6***:** 221-232.

42. Artsimovitch I. and Landick R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals**.** Proc. Natl. Acad. Sci. U. S. A *97***:** 7090-7095.

43. Artsimovitch I. and Landick R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals**.** Proc. Natl. Acad. Sci. U. S. A *97***:** 7090-7095.

44. Steitz T.A. (2006). Visualizing polynucleotide polymerase machines at work**.**

EMBO J. *25*: 3458-3468.

45. Wang H. and Oster G. (2002). Ratchets, power strokes, and molecular motors. Applied Physics A-Materials Science & Processing *75*: 315-323.

46. Erie D.A., Yager T.D., and von Hippel P.H. (1992). The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective. Annu. Rev. Biophys. Biomol. Struct. *21*: 379-415.

47. Nudler E., Mustaev A., Lukhtanov E., and Goldfarb A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell *89*: 33-41.

48. Yager T.D. and von Hippel P.H. (1991). A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. Biochemistry *30*: 1097-1118.

49. Yager T.D. and von Hippel P.H. (1991). A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. Biochemistry *30*: 1097-1118.

50. Bai L., Shundrovsky A., and Wang M.D. (2004). Sequence-dependent kinetic model for transcription elongation by RNA polymerase. J. Mol. Biol. *344*: 335-349.

51. Komissarova N. and Kashlev M. (1997). RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. J. Biol. Chem. *272*: 15329-15338.

52. Jin D.J., Burgess R.R., Richardson J.P., and Gross C.A. (1992). Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. Proc. Natl. Acad. Sci. U. S. A *89*: 1453-1457.

53. Walstrom K.M., Dozono J.M., Robic S., and von Hippel P.H. (1997). Kinetics of the RNA-DNA helicase activity of Escherichia coli transcription termination factor rho. 1. Characterization and analysis of the reaction. Biochemistry *36*: 7980-7992.

54. Brennan C.A., Dombroski A.J., and Platt T. (1987). Transcription termination factor rho is an RNA-DNA helicase. Cell *48*: 945-952.

55. Park J.S., Marr M.T., and Roberts J.W. (2002). E. coli Transcription repair

coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation**.** Cell *109***:** 757-767.

56. Chambers A.L., Smith A.J., and Savery N.J. (2003). A DNA translocation motif in the bacterial transcription--repair coupling factor, Mfd**.** Nucleic Acids Res. *31***:** 6409-6418.

57. Toulokhonov I. and Landick R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination**.** Mol. Cell *12***:** 1125-1136.

58. Yarnell W.S. and Roberts J.W. (1999). Mechanism of intrinsic transcription termination and antitermination**.** Science *284***:** 611-615.

59. Toulokhonov I. and Landick R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination**.** Mol. Cell *12***:** 1125-1136.

60. Toulokhonov I., Artsimovitch I., and Landick R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins**.** Science *292***:** 730-733.

61. Park J.S. and Roberts J.W. (2006). Role of DNA bubble rewinding in enzymatic transcription termination**.** Proc. Natl. Acad. Sci. U. S. A *103***:** 4870-4875.

62. Yarnell W.S. and Roberts J.W. (1999). Mechanism of intrinsic transcription termination and antitermination**.** Science *284***:** 611-615.

63. Park J.S. and Roberts J.W. (2006). Role of DNA bubble rewinding in enzymatic transcription termination**.** Proc. Natl. Acad. Sci. U. S. A *103***:** 4870-4875.

64. Yarnell W.S. and Roberts J.W. (1999). Mechanism of intrinsic transcription termination and antitermination**.** Science *284***:** 611-615.

65. Toulokhonov I., Artsimovitch I., and Landick R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins**.** Science *292***:** 730-733.

66. King R.A., Markov D., Sen R., Severinov K., and Weisberg R.A. (2004). A conserved zinc binding domain in the largest subunit of DNA-dependent RNA polymerase modulates intrinsic transcription termination and antitermination but does not stabilize the elongation complex**.** J. Mol. Biol. *342***:** 1143-1154.

Chapter 2

Methods for the study of transcription with single-molecule technique

2.1 Previous studies of the transcription kinetics with the single-molecule technique

Transcription reaction involves many steps and reaction pathways, and is highly heterogeneous and asynchronous. Most results from bulk measurements only reflect an ensemble average of intermediate states and parallel reactions, many phenomena are included in the ensemble average. Thus, there is a need for the use of single molecule technique to uncover the reaction intermediates and the reactions in transcription. The single-molecule technique provides an opportunity to observe the distribution of the reaction intermediates and study the transcription mechanism of each single reaction step. Below is the brief review for the single-molecule techniques used in the study of RNAP transcription.

2.2 Atomic force microscopy (AFM)

AFM has been adopted in determining the position of RNAP along DNA as a function of time. However, surface adsorption in the AFM measurement affects the kinetics of RNAP on DNA molecules. In addition, AFM also has been used to measure the bending of DNA by formation of open complex and elongation complex of RNAP. It has been shown that in an open promoter complex, the DNA bend angle varies from 55-88° and contour length is reduced by ~30 nm [1;2]

2.3 Tethered Particle Motion Method

The tethered particle motion method has been used in study of transcription elongation [3;4], where RNAP is immobilized on a glass surface, and the RNAP is bound to a DNA molecule at one end, and a streptavidin coated particle (or an avidin coated bead) is tethered by the other end of the DNA molecule through biotinylated nucleotides. The RNAP transcribes in the direction of the tethered particle or the opposite direction, depending on the orientation of a promoter. As the RNAP transcribes in the direction of the tethered particle, the contour length of DNA between the tethered particle and the immobilized RNAP is reduced, whereas if

RNAP transcribes in the opposite direction it is increased. The magnitude of Brownian motion of the tethered particle is correlated with the length of the tethering DNA. Thus, by monitoring the Brownian motion of the tethered particle with light microscopy and digital image processing, the tethering DNA length is evaluated. However, the temporal and spatial resolution of this method is limited by the Brownian motion. At the tether length of 308 bp and 1915 bp, it has an accuracy of 108 bp and 258 bp r.m.s., respectively. The shorter the tether length, the more precision can be obtained and less acquisition is required. The method has been used in determination of elongation velocity of RNAP at 15 bp s^{-1} (Figure 2-1) [4].

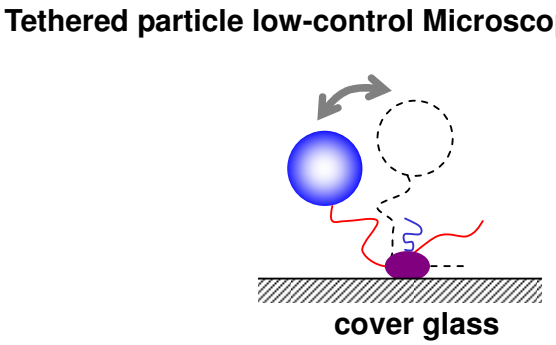


Figure 2-1
The RNAP is immobilized on the glass surface and bound to a DNA molecule, which tethers a streptavidin coated particle (or an avidin coated bead). Through the Brownian motion of the particle or bead, the tethered DNA length is evaluated.

2.4 Flow-control video microscopy

Tethered particle method could be modified by introducing an external force on a tethered bead. That can be done by immobilizing RNAP on a bead, capturing the bead by a micropipette and introducing a constant flow to apply an external force on the DNA tethered bead (Figure 2-2) [5;6].

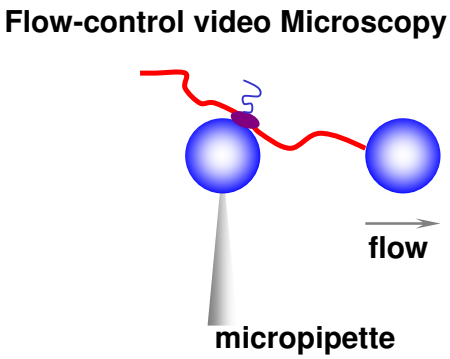


Figure 2-2
The RNAP is immobilized on a bead which is captured by a micropipette. The RNAP is also bound to one end of a DNA molecule which tethers another bead in the other end. An external force is applied on the DNA linking bead through a flow.

2.5 Optical tweezers with rotation detection

By attaching fluorescence labeled beads to an optically trapped DNA tethered particle and immobilizing the RNAP on a glass surface, it is observed that the tethered

particle rotates while the RNAP transcribes the tethering DNA. The results suggest the rotation of tethered particle is driven by RNAP groove-tracking while sliding along a DNA molecule. Since the sliding of RNAP along a DNA molecule is involved with groove-tracking, RNAP rotates along the helical axis of the DNA molecule and thus the movement of RNAP along a DNA molecule is longitudinal and rotational, too.

However, it is difficult to observe these events by this method because of Brownian motion of tethered particles, low binding possibility of RNAP, short lifetime of promoter search and two different rotation directions of tethered particles (Figure 2-3) [7]. The observation can be deviated by the Brownian motion. The low binding possibility of RNAP, short lifetime promoter search and two different rotation directions of tethered particles results in a few distinguishable events in observation.

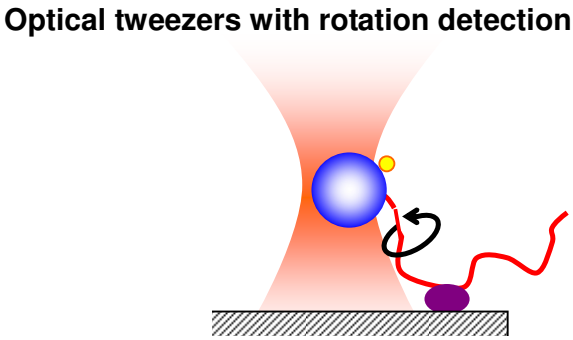


Figure 2-3
The RNAP is immobilized on a glass surface and bound to a DNA molecule. The DNA molecule tethers a bead at one end, and the fluorescence labeled beads are attached to the DNA tethered bead. An external force is applied on the DNA tethered bead through optical trapping.

2.6 The combination of total internal reflection fluorescence (TIRF) with optical trapping

In TIRF, there are two beads which are attached on the two ends of a DNA molecule which is aligned by fixing two beads with two optical traps. The position of fluorescence labeled RNAP sliding along the DNA molecule is determined. It is found that RNAP prefers to bind AT rich sequences rather than GC rich sequences of DNA. The association constant for the DNA binding is around $10^{-3} bp^{-1} M^{-1} s^{-1}$. From the rate of dissociation and association to a DNA molecule, the diffusion constant of RNAP along the DNA is obtained as $10^{-10} cm^2 s^{-1}$. Nevertheless, the method could not distinguish other binding species like open complex or closed

complex. The estimated diffusion constant is 3 order lower than other studies [8-11]. This may be due to the low spatial resolution (0.2μm) which causes underestimation of the diffusion constant (Figure 2-4).

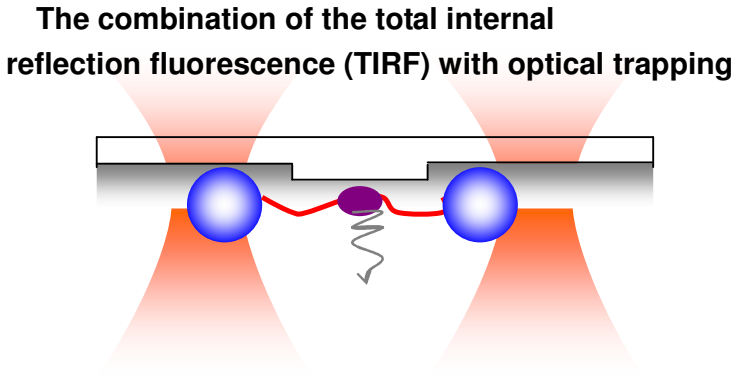


Figure 2-4
In TIRF, a DNA molecule is bound to two beads which are optical trapped. The position of fluorescence labeled RNAP can be determined. The diffusion constant is obtained through the position of RNAP with time.

2.7 Dual optical tweezers

In the dual optical tweezers method, one end of the DNA is anchored on a bead; another end of the DNA tethers a RNAP molecule which is immobilized on a bead. This could also be achieved by optically trapping the tethered beads [3]. In order to reduce the Brownian noise and increase the position stability, the whole experimental apparatus can be placed in helium rather than in the air. The helium

gas can reduce the deflection of laser beam. The lower deflection of laser beam increases the positional stability. In this method, the spatial resolution is able to reach the angstrom level.

Through this experiment setup, it has been demonstrated that RNAP is able to incorporate single nucleotides into the RNA product one by one. The elongation dynamics is determined by applying assistant load and hindering load on the RNAP. By fitting the dynamics into both a Brownian ratchet model and a power stroke model, it has been demonstrated that RNAP elongates along DNA in the Brownian ratchet mechanism (Figure 2-5) [12].

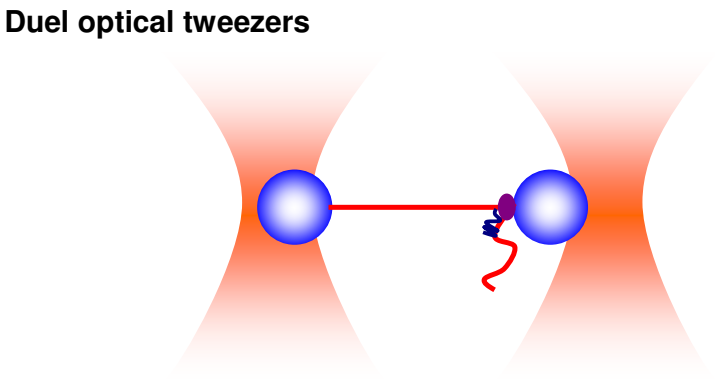


Figure 2-5
In the dual optical tweezers method, one end of the DNA is anchored on a bead. Another end of the DNA tethers a RNAP molecule which is immobilized on a bead.

Summary

The single-molecule technique has been used in uncovering the phenomena behind the ensemble average. Transcription is highly heterogeneous and involves many reaction pathways and intermediates. Atomic force microscopy is used in determining the position of RNAP and the conformation of DNA molecules, but it has a disadvantage of surface interaction between surface and RNAP. Tethered particle motion method is used in determining the tracking of RNAP along DNA molecule. Brownian motion causes deviations for the determination of RNAP position along the DNA molecule. The flow-control video microscopy can overcome the Brownian motion problem by exerting an external force on the bead. Or optical trapping can be used to fix the bead in detection of RNAP's groove-tracking and reduce the Brownian motion. In addition, the combination of total internal reflection fluorescence (TIRF) with optical trapping has been used in determination of diffusion of RNAP along the DNA molecule. The dual optical tweezers can be used in determining the kinetics of elongation in the resolution of angstrom.

References

[1] Rees,W.A., Keller,R.W., Vesenka,J.P., Yang,G., & Bustamante,C. (1993) Evidence of DNA bending in transcription complexes imaged by scanning

force microscopy. *Science* **260,** 1646-1649.

[2] Rivetti,C., Guthold,M., & Bustamante,C. (1999) Wrapping of DNA around the E.coli RNA polymerase open promoter complex. *EMBO J.* **18,** 4464-4475.

[3] Shaevitz,J.W., Abbondanzieri,E.A., Landick,R., & Block,S.M. (2003) Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature* **426,** 684-687.

[4] Yin,H., Landick,R., & Gelles,J. (1994) Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophys. J.* **67,** 2468-2478.

[5] Davenport,R.J., Wuite,G.J., Landick,R., & Bustamante,C. (2000) Single-molecule study of transcriptional pausing and arrest by E. coli RNA polymerase. *Science* **287,** 2497-2500.

[6] Forde,N.R., Izhaky,D., Woodcock,G.R., Wuite,G.J., & 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,** 11682-11687.

[7] Sakata-Sogawa,K. & Shimamoto,N. (2004) RNA polymerase can track a DNA groove during promoter search. *Proc. Natl. Acad. Sci. U. S. A* **101,** 14731-14735.

[8] Singer,P. & Wu,C.W. (1987) Promoter search by Escherichia coli RNA polymerase on a circular DNA template. *J. Biol. Chem.* **262,** 14178-14189.

[9] Neuman,K.C., Abbondanzieri,E.A., Landick,R., Gelles,J., & Block,S.M. (2003) Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* **115,** 437-447.

[10] Wang,M.D., Schnitzer,M.J., Yin,H., Landick,R., Gelles,J., & Block,S.M. (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* **282,** 902-907.

[11] Yin,H., Wang,M.D., Svoboda,K., Landick,R., Block,S.M., & Gelles,J. (1995) Transcription against an applied force. *Science* **270,** 1653-1657.

[12] Abbondanzieri,E.A., Greenleaf,W.J., Shaevitz,J.W., Landick,R., & Block,S.M. (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature*

438, 460-465.

Chapter 3

Single-molecule DNA nanomanipulation

Single-molecule DNA nanomanipulation is a technique used to study DNA-protein interactions upon the change of topology of a DNA molecule. The single-molecule DNA nanomanipulation experiment geometry is shown in Figure 3-1.

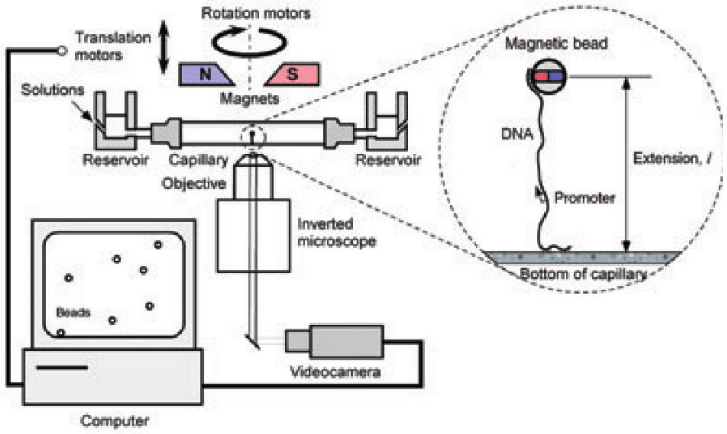


Figure 3-1
The instrument setup of the single-molecule DNA nanomanipulation experiment
Figure from ref [1]

3.1 Experimental setup

A microscope with CCD camera is used to collect the images of beads that are processed and analyzed by a computer after collection. The camera collects images

at a rate of 33 Hz. The objective of the microscopy is bound onto a piezo driver, which is able to adjust the object's z axis position with tens of nanometers precision (Figure 3-1). After injection of the polystyrene beads through a pair of reservoirs, the beads are deposited and attached to the surface of the capillary. The polystyrene beads serve as a reference for relocating the focal plane on the polystyrene coated glass surface.

DNA-paramagnetic beads are injected after the polystyrene beads and then sedimentation occurs. After sedimentation, the antidigoxigenin on the plastic surface grasps one end of DNA which is labeled with digoxigenin (Figure 3-2). The other end of this DNA is biotinylated and attached on the streptavidin coating paramagnetic bead.

Above the paramagnetic bead-anchored DNA is a pair of magnets which apply an external mechanical force on the DNA molecule through the paramagnetic bead. The magnets are mounted on a rotor, which could change the orientation of the magnets. The rotor is attached to a motor which is able to adjust the z axis position of the rotor and magnets (Figure 3-1).

Since the gradient of the magnetic field created by the magnets changes little on the micrometer scale and is parallel to the surface, the change in force due to the change of the DNA extension can be neglected. However, at the difference of millimeter scale, the gradient of magnetic field decays exponentially, which has an effect on the magnitude of force applied on the magnetic beads. Thus, changing the z axis position of the magnets in millimeter scale alters the magnitude of the magnetic force on the paramagnetic beads. In addition, the magnetic force also depends on the size of the paramagnetic beads. Generally, a bead with a radius of 1.5 micrometers can generate a force up to 50 pN.

3.2 Preparation of DNA

The DNA molecule used in our current experiment setup is around 2kb and is obtained by PCR from a plasmid. The two ends of this DNA fragment contain restriction enzyme sites. After treating the DNA with restriction enzyme, the two ends of the 2kb DNA are ligated with two 1kb DNA fragments, which have biotin and digoxigenin labeled nucleotides, respectively (Figure 3-2). After ligation, the two ends of the DNA are bound to the surface of the capillary and the paramagnetic bead with antidigoxigenin and streptavidin, respectively.

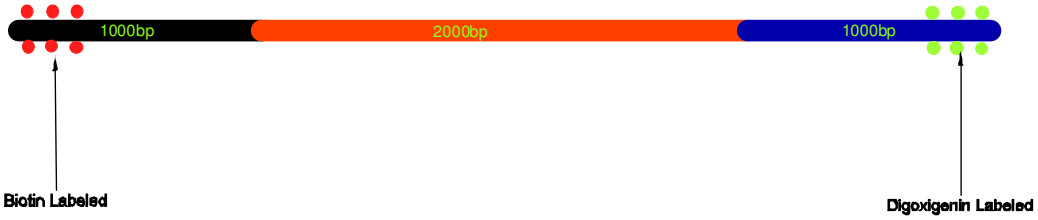


Figure 3-2
The basic DNA construct

3.3 Determination of the position of a bead

In order to determine the position of a bead, an image is recorded and this image is analyzed. The image for a bead in the capillary has concentric optical diffraction rings (Figure 3-3-a). The x, y position of the bead is determined by the center of the diffraction ring, which is obtained by the following procedures. The image has a diffraction ring located at X(0), Y(0) at time equal to zero (t=0). The X(t') position of the bead at time equal to t' (t=t') can be obtained by carrying a FFT (fast Fourier transform) correction of inverse image intensity profile along X=X(0) and relative to Y=Y(0) (Figure 3-3c) with the new image intensity profile. The FFT correction of two image intensity profiles can have a maximum value by translating the inverse intensity profile by δy. $Y(t')=Y(0)+\frac{1}{2}\delta y$. Using this method, one can also obtain X(t'). The bead's new position X(t'), Y(t') is determined from a

previous position of X(0) and Y(0).

Before determining the z position of a bead, it is required to have a reference for the change of radius of optical diffraction ring along the z axis of focal plane which is relative to the objective of the microscopy. In order to make the reference, the objective has to move down 300nm for a couple cycles and take an image at the focus plane with various z positions. This step is necessary to determine the radii of rings with various z positions of the object. The radii of the diffraction ring are linearly correlated with z positions of object. In other word, the radii of the diffraction rings are proportional to the distance between the objective and the bead which we are observing. Therefore, as the objective is fixed, the z position of a bead which is relative to objective could be determined by the radii of its diffraction rings from this linear correlation (Figure3-3-b).

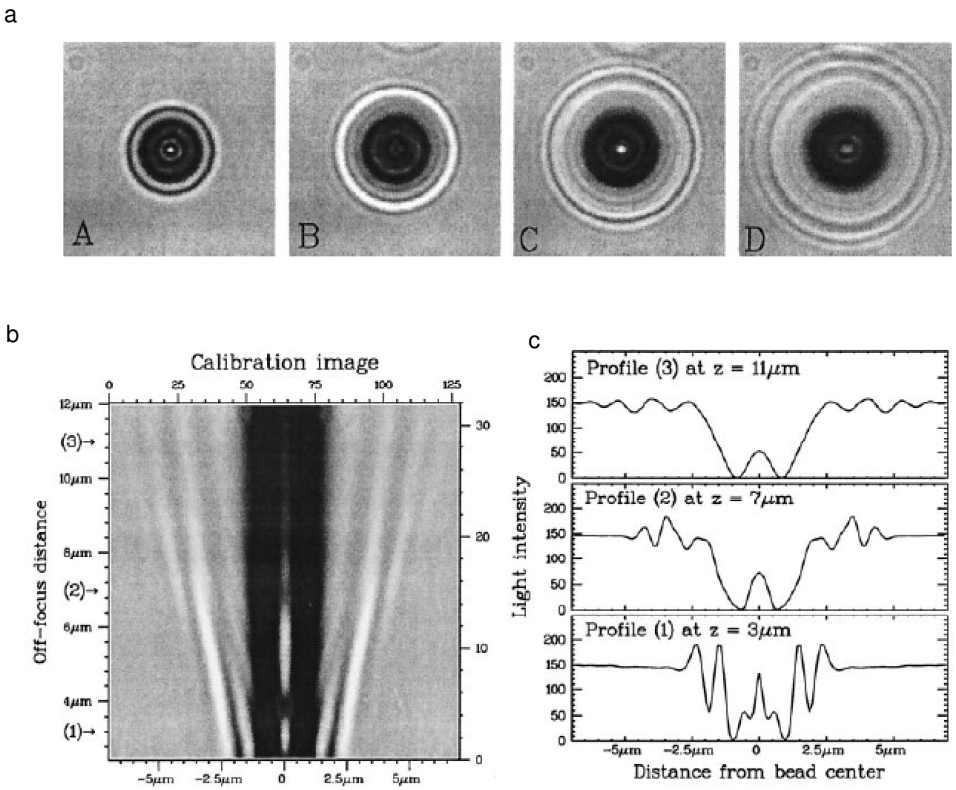


Figure 3-3

The image of the a paramagnetic bead on the capillary surface with various z positions

(a) Image of 4.5 μm paramagnetic bead at various positions of focal plane (A) z=1μm (B) z=3μm (C) z=5μm (D) z=7μm

(b) Calibration of 4.5 μm paramagnetic bead on the glass surface. The lines were obtained from radius profile with various z axis position of object.

(c) Intensity profile corresponds to z=3μm z=7μm z=11μm

Figure from ref[2]

3.4 Force measurement and manipulation

With rotation of the magnets, the paramagnetic bead rotates synchronously in a step-locked register manner, which introduces super-helical turns into the torsionally constrained DNA.

The magnitude of the magnetic force applied on the DNA can be adjusted by moving the magnets along z-axis. For the DNA in the experimental setup, the magnitude of the magnetic force applied on the DNA can be monitored by Brownian motion of the paramagnetic beads. That follows the relation:

$$m\frac{d^2x(t)}{dt^2} = -6\pi\eta R\frac{dx(t)}{dt} - k_x x(t) + f_a(t) \quad \text{(equation 3.1)}$$

Here, the R is radius of the bead and η is the friction coefficient. $-6\pi\eta R\frac{dx(t)}{dt}$ is the friction force in the system. $-k_x$ is the spring constant for the restoring force of the DNA molecule which tethers the bead. The random force can be characterized as

$$\left\langle \overrightarrow{f(t)_a} \right\rangle = 0 \quad \text{and} \quad \left\langle \overrightarrow{f(t)_a}\,\overrightarrow{f(t')_a} \right\rangle = 4k_B T(6\pi\eta R)\delta(t-t') \quad \text{(equation 3.2)}$$

The k_B is Boltzmann constant. The Fourier transform of X(t)and f_a(t) are:

$$X(\varpi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)\exp(-iwt)dt \quad \text{(equation 3.3)}$$

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f_a(t)\exp(-iwt)dt \quad \text{(equation 3.4)}$$

Then one can obtain:

$$X(\omega) = \frac{F(\omega)}{-m\omega^2 + 6\pi i\eta R\omega + k_x} \quad \text{(equation 3.5)}$$

However, in the system, $m \approx 10^{-15}kg$, $\eta = 10^{-3}$, $k_x = 10^{-9}N/m$ and $R = 10^{-6}m$,the term $6\pi\eta R/m$ is close to $\approx 3\times10^7 Hz$. It can not be observed by the CCD camera which collects images at a pace of $33Hz$ in the experiment setup. Thus, the moving equation of the bead can be rewritten as:

$$-6\pi\eta R\frac{dx(t)}{dt}-k_x x(t)+f_a(t)=0 \quad \text{(equation 3.6)}$$

The square of $X(t)$ in the ω space obtained from the equation above is

$$X^2(\omega)=\frac{F(\omega)}{k_x^2((\frac{\omega}{\omega_c})^2+1)}=\frac{2/\pi\, k_B T\cdot 6\pi R\eta}{k_x^2}\frac{1}{(\frac{\omega}{\omega_c})^2+1} \quad \text{(equation 3.7)}$$

Here ω_c is cut off frequency and $\omega_c = k_x/6\pi\eta R$.

$$\left\langle \delta x^2\right\rangle=\int_0^{\omega} X^2(\omega)d\omega=\frac{2/\pi\, k_B T\cdot 6\pi R\eta}{k_x^2}\frac{\omega_c}{2}=\frac{k_B T}{k_x} \quad \text{(equation 3.8)}$$

From equipartition theory: $\frac{1}{2}k_x\delta x^2=k_B T$ and $F/l=k_x$

Then the force due to the Brownian collision is $F=\frac{2k_B T}{\left\langle \delta x^2\right\rangle}l$ (equation 3.9)

Therefore, the magnitude of the magnetic force can be determined by a measurement of $\left\langle \delta x^2\right\rangle$- the fluctuation of the paramagnetic bead and l- the tether length of the DNA molecule (**Figure 3-4**).

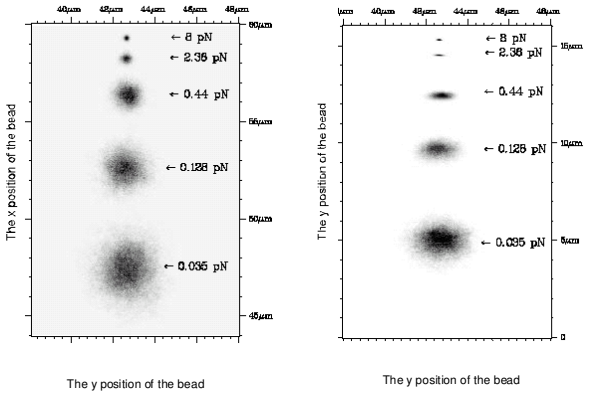


Figure 3-4
The position of a paramagnetic bead under the various magnitude of force
Position of a paramagnetic bead bound to λ DNA molecule (16.5 μm) and subjected to different stretching force
Left plane: The bead motion in microscope vertical plane x,y
Right plane: The bead motion in microscope vertical plane y,z
Figure from ref [11].

3.5 Determination of DNA extension and force

The distance between the paramagnetic bead and the surface of a capillary in the z-axis is defined as the extension of the DNA molecule- l_{obs} (Figure 3-1). The l_{obs} here is equal to the l listed (equation 3.9) above. The extension of the DNA molecule, l_{obs}, can be obtained in real time by determination of the z position of the paramagnetic bead and a reference bead on the surface of the capillary.

The moving amplitude of the paramagnetic bead-$\langle \delta x^2 \rangle$ is obtained by taking the Fourier transform of $X(t)^2$ - $X(\omega)^2$ and, after removing first few modes

of $X(\omega)^2$, integrating equation 3.8. Then the magnitude of the force is determined by equation 3.9 (Figure 3-5).

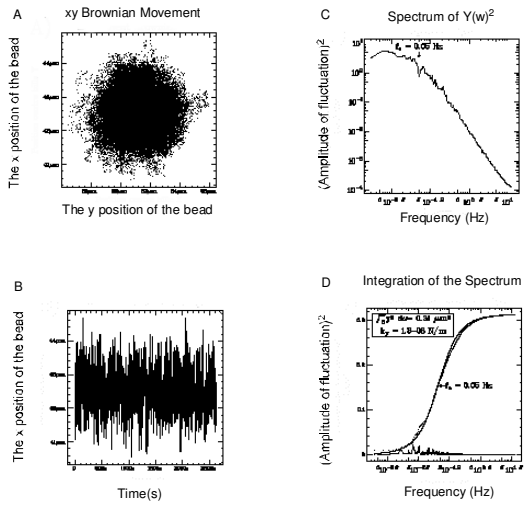


Figure 3-5

The time variance of z position of the paramagnetic bead processed with FFT
(A) The raw Brownian motion in the microscope's focal place
(B) The position of the bead along Y axis as function of time
(C) Taking the Fourier transform of the bead's position - $Y(t)$ **gives** $Y(\omega)^2$ **as function of** ω**.** $\omega = 2\pi f$ **and the cutoff frequency:** $fc = \omega_c / 2\pi$**. The amplitude of cutoff frequency below** $fc = 0.05Hz$ **is constant. But the amplitude of cutoff frequency above** $fc = 0.05Hz$ **drops off as** $1/f^2$**.**

(D)Integrating of spectrum gives Lorentzian with $\langle y^2 \rangle = 0.31\mu m^2$

Figure from ref [11].

Therefore, the relationship of l_{obs} and stretching force or the degree of supercoiling of a DNA molecule can be obtained, since we can control the two parameters- the magnitude of stretching force and the degree of supercoiling of the

DNA. At a proper buffer condition and temperature, like 10 mM phosphate buffer and 25℃, we can determine the relationship of stretching force and l_{obs} of the DNA molecule.

3.6 Stretching of torsionally free DNA

For the torsionally free DNA, the force and extension curve can be categorized into three regimes of elasticity: entropy regime, enthalpy regime, regime of the transition from B-form into S-form The low force regime is of the order of 0.1 pN ($F \approx k_B T / \xi$), where ζ is the persistence length of a DNA molecule. In this regime, the configurational entropy of the DNA molecule is reduced by stretching; thus, it is also called the entropy regime [3]. If the force increases to the order of 10~70pN, the DNA molecule enters into another regime called enthalpy regime. In this regime, the double strand-helix of DNA is deformed. The relationship of stretching force and extension can be described as $F = EA(\frac{l - l_0}{l_0})$ [4-6]. Here, E is Young's modulus and A is the cross-sectional area of the DNA molecule. When the force arrives at 70pN, the DNA can transform from B-form into S-form (stretching form). A DNA in S-form has a 70% longer length than its crystallographic length in B-form [4]. The Worm-Like Chain model can describe the stretching behavior of DNA

below 10 pN: $F = \frac{k_B T}{\xi}[\frac{1}{4(1-l/l_0)^2} - \frac{1}{4} + \frac{l}{l_0} + \sum_{i=2}^{i=7} a_i (\frac{l}{l_0})^i]$ [7].

(Figure 3-6)

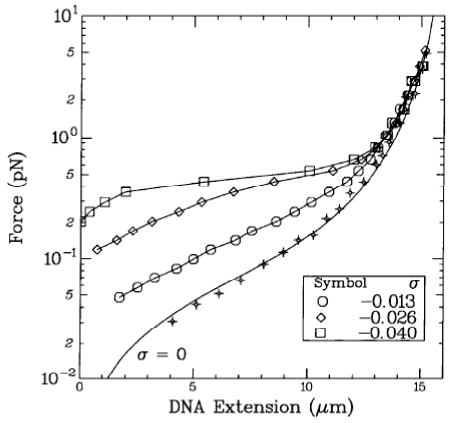


Figure 3-6

The Force vs. the DNA extension at various σ, ($\sigma = \frac{n}{Lk_0}$)

Force vs. DNA extension curves for a DNA at various σ at 10mBP buffer. σ=0 can fit into WLC model.
Figure from ref [8].

3.6.1 Stretching of supercoiled DNA

Under a fixed stretching force at which the Worm-Like Chain model can describe the system well, l_{obs} can vary with the degree of supercoiling of the DNA

($\sigma = \frac{n}{Lk_0}$, Lk_0 is the number of turns in the relaxed DNA molecule, is determined by dividing the total base pairs of the molecule by the relaxed bp/turn which is 10.4 in B form DNA usually.). For a double-stranded DNA molecule under low force (ie. $F = 0.2\,pN$), it follows a topological relation: *Lk*= *Tw*+*Wr*. *Lk* is the linking number and equal to the sum of the supercoiling number (*Wr*, writhe number) and the twist number (*Tw*). A torsionally constrained DNA molecule has invariant *Lk*. Therefore, when rotation of the paramagnetic bead introduces superhelical turns into the torsionally constrained DNA molecule, the supercoiling number of the DNA molecule has to change and compensate the change in twist of the DNA molecule (Figure 3-7).

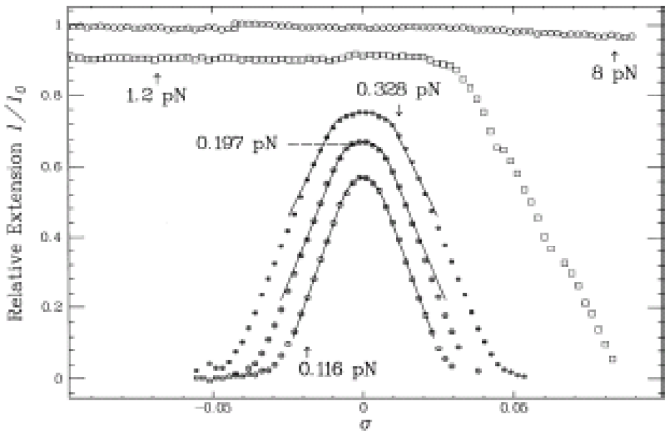


Figure 3-7
The extension of DNA vs. σ at various stretching forces
At forces below 0.328 pN, the curves are symmetric with $\sigma \to -\sigma$. At a force around 1pN, the extension of DNA does not change as σ < 0. When the force is 8pN, the extension of DNA does not change with σ.
Figure from ref [9].

At intermediary force (F=1 pN), the double stranded DNA stops formation of supercoils when underwound, instead the double stranded DNA starts to denature. The DNA molecule, when overwound, still forms supercoils and follows the topological relationship - $\Delta Lk = \Delta Tw + \Delta Wr$ (Figure 3-7). At high force (F=8pN), unwiding or overwinding of the DNA causes denaturation of the DNA molecule (Figure 3-7).

At low force, at the beginning of the twisting of the DNA, the torque applied on the DNA increases linearly with twist angle: $\Gamma = k_B T \frac{C}{l_0} \Omega$. Here, the Ω is twist angle. The $C \cdot k_B T$ is twist stiffness of the DNA molecule at a force F which

stretches the DNA molecule to l_0. After a number of superhelical turns are introduced into the DNA molecule, the torque stops increase and becomes saturated. The DNA begins to bend and forms a loop with radius R. The torque can be described as $\Gamma = \sqrt{2BF}$ since in this regime the torque is constant. Here, B is bending stiffness constant of the DNA molecule and F is the force applied on the DNA molecule. Also, the extension of the DNA varies linearly with the number of superhelical turns which have been introduced into the DNA. It is suited to use the model at low force. At high force or intermediary force regime, the force destroys the plectonemics – tertiary structure of the DNA.

Because the DNA molecule with different degrees of supercoiling corresponds to a particular l_{obs} when the DNA is in a regime of low force (F=0.1pN), the calibration can be made by introducing various numbers of superhelical turns into the DNA molecule and determining each corresponding l_{obs}.

The calibration curve may cover a range of superhelical turns. There is a constant torque regime, in which the variation of l_{obs} linearly corresponds to the number of superhelical turns and the torque is constant. At 10 mM phosphate buffer and 25℃, in this regime, a change of one superhelical turn can cause a change of

56±5nm in l_{obs} [10] (Figure 3-8). This kind of property can help the study of DNA melting by enzymes. For instance, with the experiment set up as mentioned above, for a negatively supercoiled DNA molecule with a promoter, approximately one turn of unwinding can be introduced by promoter DNA melting with RNAP, which causes a decrease in *Tw*. In order to compensate the loss in *Tw* of the DNA, the supercoiling number of the DNA molecule must increase. For a negative supercoiled DNA molecule, the increase in supercoiling number can increase the extension of the DNA (Figure 3-9). Thus, from the change of l_{obs}, we can evaluate how many turns of double stranded DNA (or DNA base pairs) have been melted. Based on the experiment construction, with 2kb DNA, the spatial resolution can be achieved at 1 bp and temporal resolution can reach 1 second.

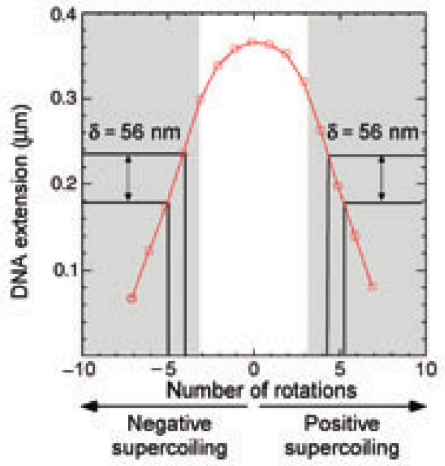


Figure 3-8

Calibration of the DNA extension vs. superhelical turns

At constant torque regime, the change of extension per-superhelical turn is 56.5 nm Figure from ref [1].

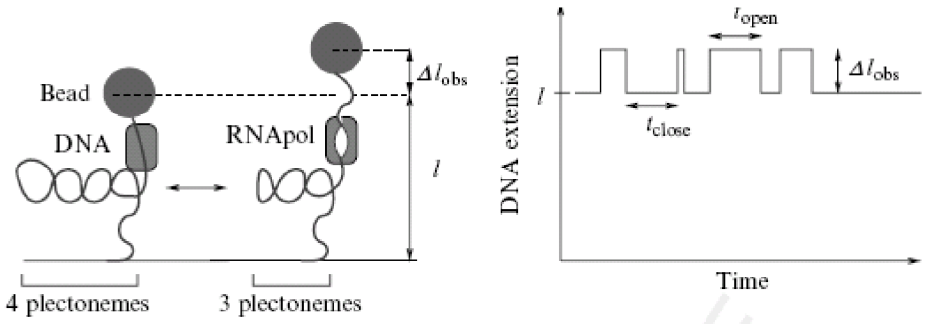


Figure 3-9

The DNA extension changes when RNAP melts the promoter of the DNA

Figure from ref [10]

References

1. Revyakin A., Ebright R.H., and Strick T.R. (2005). Single-molecule DNA nanomanipulation: improved resolution through use of shorter DNA fragments**.** Nat. Methods *2***:** 127-138.

2. Gosse C. and Croquette V. (2002). Magnetic tweezers: micromanipulation and force measurement at the molecular level**.** Biophys. J. *82***:** 3314-3329.

3. Bouchiat C., Wang M.D., Allemand J., Strick T., Block S.M., and Croquette V. (1999). Estimating the persistence length of a worm-like chain molecule from force-extension measurements**.** Biophys. J. *76***:** 409-413.

4. Smith S.B., Cui Y., and Bustamante C. (1996). Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules**.** Science *271***:** 795-799.

5. Cluzel P., Lebrun A., Heller C., Lavery R., Viovy J.L., Chatenay D., and Caron F. (1996). DNA: an extensible molecule**.** Science *271***:** 792-794.

6. Hogan M.E., Rooney T.F., and Austin R.H. (1987). Evidence for kinks in DNA folding in the nucleosome**.** Nature *328***:** 554-557.

7. Bouchiat C. and Mezard M. (1998). Elasticity model of a supercoiled DNA molecule**.** Physical Review Letters *80***:** 1556-1559.

8. Strick T.R., Bensimon D., and Croquette V. (1999). Micro-mechanical measurement of the torsional modulus of DNA**.** Genetica *106***:** 57-62.

9. Strick T.R., Allemand J.F., Bensimon D., and Croquette V. (2000). Stress-induced structural transitions in DNA and proteins**.** Annu. Rev. Biophys. Biomol. Struct. *29***:** 523-543.

10. Revyakin A., Allemand J.F., Croquette V., Ebright R.H., and Strick T.R. (2003). Single-molecule DNA nanomanipulation: detection of promoter-unwinding events by RNA polymerase**.** Methods Enzymol. *370***:** 577-598.

11. Strick,T.R. (1999) Mechanical supercoiling of DNA and its relaxation by topoisomerase . Ph.D. Thesis

Chapter 4

Detection of initial transcription by single-molecule DNA nanomanipulation

4.1 Background and Objectives

Transcription starts when RNAP binds to a promoter and forms a closed complex. An open complex state follows the closed complex state, in which RNAP unwinds a double stranded promoter DNA into two single stranded DNAs. The previous study on promoter unwinding has demonstrated the capability of nanomanipulation technique to determine lifetime and frequency of the open complex unwinding events occurring on *lacCONS* and *rrnB*P1 promoter [1].

In this object, we used a similar approach to study initial transcription on the N25 $_{\text{anti DSR}}$ promoter. Upon addition of a subset of NTPs, RNAP starts abortive initiation on the N25 $_{\text{anti DSR}}$ promoter. Different subsets of NTPs determine the different lengths of the nascent RNA. However, if the supply of subsets of NTPs is more complete, it is supposed that RNAP can make longer nascent RNA or go

through more nucleotide addition steps. Thus, when provided a more complete subset of NTPs, the lifetime of the abortive initiation state will be longer. To test this hypothesis, we determine the lifetime of the abortive initiation state with different subsets of NTPs and various concentrations of a single nucleotide.

4.2 Experimental design and methods

In the experimental setup, the 2kb long DNA fragments was used. Each DNA molecule tethers a paramagnetic bead at the end. The DNA fragment contains a promoter, a transcribed region with or without a terminator. RNAP binds to the promoter and starts RNA synthesis on the transcribed region. RNAP melts the double stranded DNA and forms a transcription bubble near the transcription start site. Since the size of the transcription bubble affect the *Tw* of the tethering DNA in the low force regime, in order to compensate the change in *Tw*, the |*Wr* | of the DNA molecule must change as well. Thus, the l_{obs} of the DNA molecule must decrease if the |*Wr*| increases or decrease The melting of a couple of base pairs of the tethering DNA can result in a change in l_{obs} on the order tens of nanometers in the constant torque regime. The number of melted base pairs of the tethering DNA reflects the size of the transcription bubble. Thus, during transcription, each l_{obs} represents a

different transcription stage having a specific size of the transcription bubble. In the following chapters, we used the same experimental setup to study transcription with different DNA templates.

4.2.1 Construct of N25 promoter DNA templates

Here, we used the DNA construct N25 -100 -tr2. The N25 -100 -tr2 has a N25 promoter, a 100 bp transcribed region and *tR2* terminator (Figure 4-1). We choose the N25 promoter as this promoter was characterized in the bulk measurement [2].

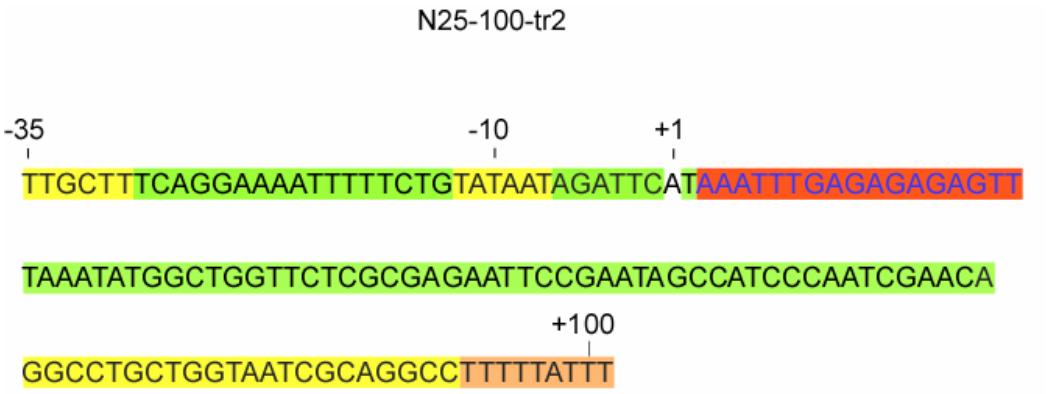


Figure 4-1
The constructs of the N25-100-tr2 (Top strand)
The promoter element -10 and -35 are in yellow. The ITS is in the red. The inverse sequence of *tR2* terminator is in yellow and the poly U corresponding sequence is in orange. (Top strand)

4.2.3 Determine abortive initiation kinetics with the subset of NTPs

By providing subsets of NTPs-ATP or ATP/UTP to the system with the N25 -100 -tr2 construct, RNAP can start initiation; and the lifetime of the ITC with each subset of NTPs can be determined. This information may help in building up a model of the transcription abortive initiation kinetics. We also used a subset of NTPs with low concentration of UTP (2 μM) and high concentration of ATP (500μM) in determining the change of DNA extension (Δl_{obs}) during abortive initiation on the N25 -100-tr2 construct.

4.2.4 Basic experiment setup

The DNA construct is incorporated into a 2kb DNA, a fragment of Rpo C gene of *Thermus thermophilus from 2050 to 4050 bp*. The two ends of the 2kb DNA are ligated with biotinylated 1kb DNA and digoxigenin labeled 1 kb DNA, respectively (Figure 4-2). In each of the following experiments, we use the same method to construct tethering DNA.

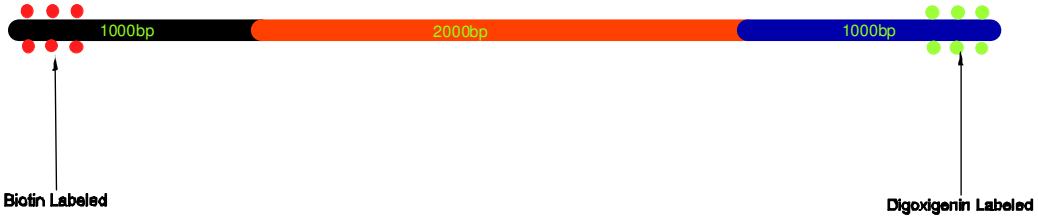


Figure 4-2

The design for the tethering DNA

The two ends of the 2kb DNA were ligated with biotinylated 1kb DNA and digoxigenin labeled 1 kb DNA.

The whole construct is used as a tethering DNA in the nanomanipulation experiment as follows (Figure 4-3).

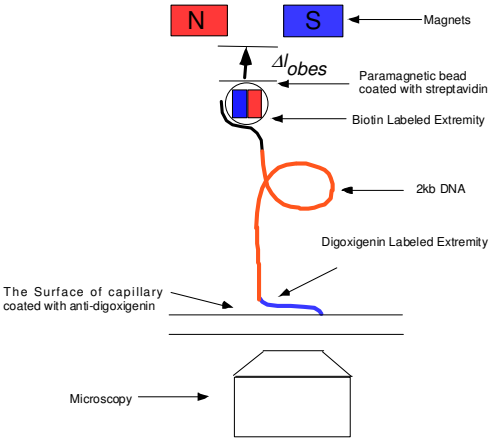


Figure 4-3

The DNA construct in Figure 4-2 used in the nanomanipulation experiment

The DNA tethers the paramagnetic bead through the binding of biotin and streptavidin, and it is anchored on the glass surface by interaction of digoxigenin and anti-digoxigenin.

4.2.5 Data analysis

The lifetime of ITC on N25 promoter was obtained by determining the time window between the two transitions of l_{obs} which correspond to the formation of ITC and collapse of ITC. Each lifetime of ITC on the N25 promoter corresponding to a particular subset of NTPs was obtained from the average of data for each particular subset of NTPs.

4.3 Results

The table 4.1 shows the lifetime of abortive initiation state of RNAP on N25 promoter with the various subsets of NTPs.

nucleotide	**Non**	**ATP**	**ATP/UTP**
Lifetime (second)	38±4	48±3	250±30

Table 4-1
The lifetime of ITC on the N25 promoter with various sets of nucleotides

With low concentration of UTP (2 μM) and high concentration of ATP (500μM) in the system, the ITC on the N25 promoter changes the DNA extension (l_{obs}) with time.

The time trace of DNA extension has a tooth-like pattern (Figure 4-5). The number of nucleotides incorporated into the nascent RNA is proportional to the change in DNA extension (Δl_{obs}). As RNAP synthesized RNA after the first two nucleotides, the bubble began to extend downstream of the N25 promoter, according to the model proposed by Hsu [2]. The largest size nascent RNA with a supply of ATP/UTP in the system with the N25 promoter is 8 nt. At the beginning of initiation, the size of ITC bubble is equal to the size of open complex bubble. The RNAP keeps incorporating nucleotides into the RNA until the size of the DNA: RNA hybrid reached a length of 8 bp. Thus, with a supply of ATP/UTP to the system with the N25 promoter, the change in DNA extension (Δl_{obs}) caused by formation of the ITC is larger than the one corresponding to an open complex (Figure 4-4). In the presence of low concentration of UTP (2 μM) and high concentration of ATP (500μM), the DNA extension (l_{obs}) changes with time. Because UTP was limited relative to ATP, we were able to observe the transitions of sub-states of abortive initiation (Figure 4-5).

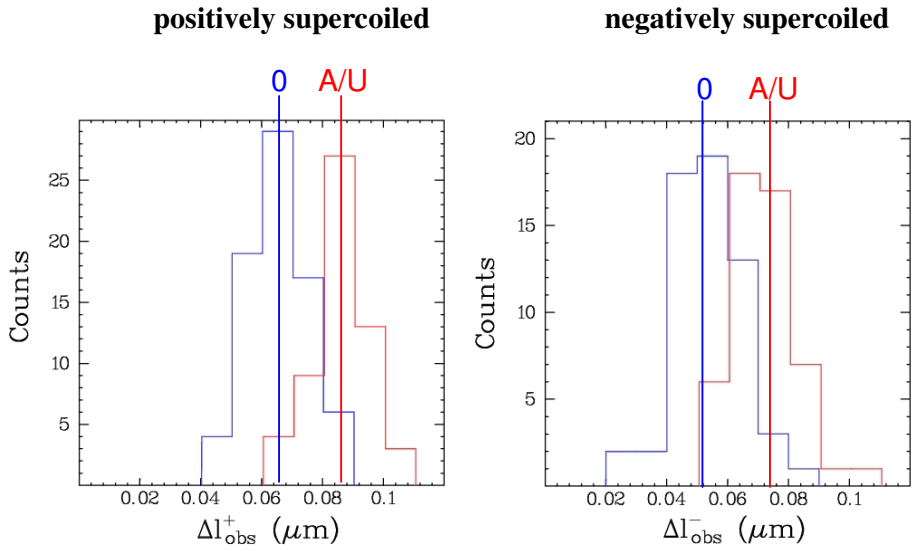


Figure 4-4 The transitionamplitude histograms for RPo (0 NTPs) and RPitc,≤8 (ATP+UTP) at the N25 promoter; Δl^{+}_{obs}, transition amplitude with positively supercoiled DNA; Δl^{-}_{obs}, transition amplitude with negatively supercoiled DNA.

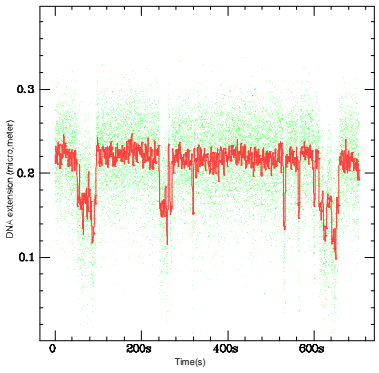


Figure 4-5
The time trace of DNA extension in abortive initiation on the N25 promoter with ATP (500μM) and UTP (2 μM)

4.4 Conclusions

The lifetime of ITC increases when it is provided with the first two initial nucleotides. This indicates that the abortive initiation state is more stable in the

presence of the first two initial nucleotides. The more NTPs are added to the system, the longer is the nascent RNA transcripts. This result also implied that the rate limiting step of the promoter escape is at abortive initiation RNA synthesis rather than at the stage of promoter rewinding because the lifetime of ITC becomes longer upon adding a more complete subsets of NTPs.

Figure 4-5 represents the DNA extension during abortive initiation, which reflects the changes in transcription bubble size during abortive initiation. It has been easily observed that the ITC with a 8 nt DNA:RNA hybrid caused a larger change in DNA extension (l_{obs}) than of open complex (Figure 4-4). Since the different transcription bubble size can represent the length of nascent RNA products, from Figure 4-4, it can be interpreted as the change in length of nascent RNA during abortive initiation. RNAP repeated cycles of the RNA synthesis and release of the nascent RNA product in different lengths. These cycles were repeated until the RNAP dissociated from the promoter.

References

[1] Revyakin,A., Ebright,R.H., & Strick,T.R. (2004) Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. *Proc. Natl. Acad. Sci. U. S. A* **101,** 4776-4780.

[2] Revyakin,A., Liu,C., Ebright,R.H., & Strick,T.R. (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314,** 1139-1143.

Chapter 5

Detection of promoter escape by single-molecule DNA nanomanipulation

5.1 Background and Objectives

Promoter escape is the last step of transcription initiation, following the abortive synthesis of nascent RNA. As RNAP escapes from a promoter, it begins elongation along a template DNA molecule. We have discussed the abortive synthesis of nascent RNA on the N25 promoter with a supply of NTPs subsets in the previous chapter. RNAP is in a cycle of synthesizing and aborting nascent RNA, restarting the synthesis of nascent RNA. Without a full set of NTPs, the RNAP can not escape from a promoter and stays in the cycle of synthesizing and aborting RNA until the RNAP dissociates from the promoter. Therefore, we used a full set of NTPs to determine promoter escape on the N25 promoter.

5.2 Experimental design and methods

With a full set of NTPs, RNAP can start abortive initiation on the N25 DNA

templates and escape from the promoter. Through the changes of the DNA extension, the ITC can be easily identified and distinguished from the open complex and the TEC.

We use a DNA template N25-100-tr2^{-} (Figure 5-1). The DNA template has the N25 promoter, a 100 bp transcribed region and a modified *tR2* terminator. The terminator on the N25-100-tr2^{-} has been modified by a change of $T \leftrightarrow G$ on the T-tract sequence of the *tR2* terminator, which can abort the transcription termination on this modified terminator.

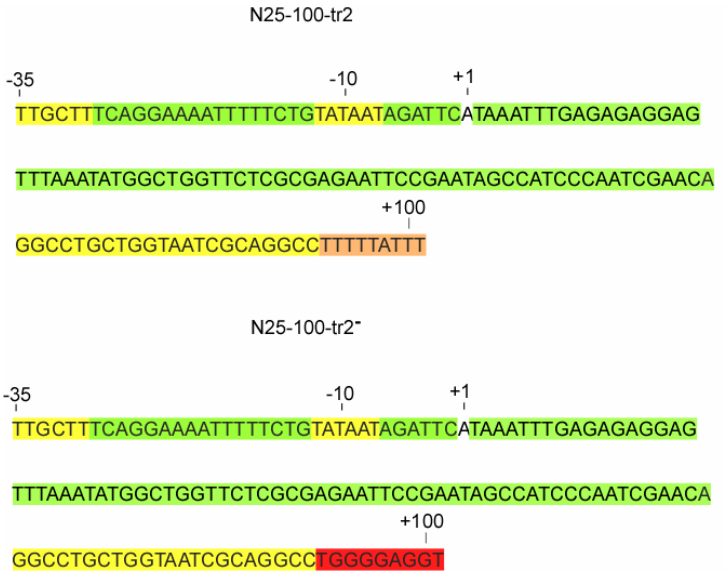


Figure 5-1
The N25-100-tr2 and N25-100-tr2^{-} DNA templates (top strand): An invert repeat sequence (in yellow) is next to the ploy G region (red) or the poly T region (in orange).

5.3 Results

In a transcription assay experiment, with a full set of NTPs, most RNAP read through the *tR2*$^{-}$ terminator while transcribing on the N25-100-tr2^{-} DNA template (Figure 5-2).

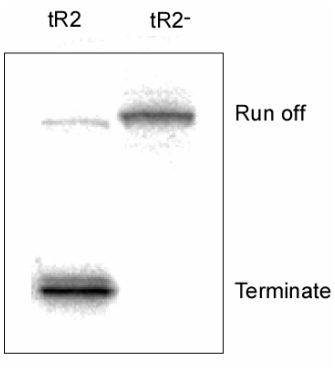


Figure 5-2
In the transcription assay experiment, most RNAP read through the *tR2*$^{-}$ terminator while some RNAP terminated at *tR2* terminator.

With a full set of NTPs, after abortive initiation, the RNAP escaped from the promoter and entered elongation (Figure 5-3). Base on the observation in chapter4 (figure 4-x), it is clear that the ITC corresponds to a larger size of DNA extension than an open complex in N25 promoter. When RNAP entered into the elongation state, the DNA extension reduced from one corresponding to the ITC. The RNAP continues elongation and reads through the N25-100-tr2^{-}.

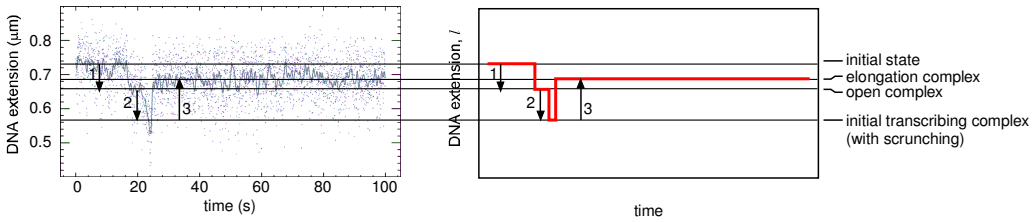


Figure 5-3
The time trace of a transcription event on the N25-100-tr2⁻ construct: the transition 1 corresponds to formation of open complex, transition 2 corresponds to formation of ITC, transition 3 is formation of TEC.

5.4 Conclusions

It has also been found in footprint experiments that the RNAP upstream promoter contacts are fixed during abortive initiation, but the active center of RNAP can move relative to the template DNA [1]. In order to explain this observation, three models have been proposed for abortive initiation: the inchworm model, the transient excursion model and the scrunching model. The inchworm model and the transient excursion model predict that the size of the ITC bubble is fixed during transcription and is almost equal to the size of the open complex bubble. However, the scrunching model proposes an increase of bubble size when RNAP is in the abortive initiation stage. Through the observation of the change in the DNA extension, it shows that open complex bubble extends to the ITC bubble during abortive initiation (Figure 5-3). Therefore, we conclude that abortive initiation occurs via the scrunching mechanism in *Escherichia coli* RNAP.

5.4.1 Initiation-elongation transition is within the time resolution of nanomanipulation

One can assume that the more stable an ITC is, the less abortive RNA transcripts are made. From the abortive transcript probability profile, most transcription abortive events occur at 8 nt for N25 promoter [2]. This means that RNAP confronts a rate determining step for promoter escape when the length of nascent RNA reaches 8 nt. With a supply of 100μM NTPs, this rate determinig event of promoter escape must occur within the time resolution of nanomanipulation. If this was not the case, we would be able to observe a DNA extension- l_{obs} which corresponds to a transcription bubble with a size of 8 nt. Because the rate determining step of promoter escape on the N25 promoter occurs at 8nt from transcription start site, when the ITC reaches the rate determining step, the ITC has 8 nt DNA:RNA hybrid. Then one is able to observe a transcription bubble with a size of 8nt on the N25 promoter during initiation-elongation transition like the transcription bubble in Figure 4-5. It is clear, in Figure 5-3, that there is no other sub-state transition during abortive initiation being observed like those in Figure 4-5. Therefore, the time resolution of nanomanipultaion does not allow us to determine the time for the N25 promoter escape with a supply of 100μM NTPs.

References

[1] Straney,D.C. & Crothers,D.M. (1987) A stressed intermediate in the formation of stably initiated RNA chains at the Escherichia coli lac UV5 promoter. *J. Mol. Biol.* **193,** 267-278.

[2] Hsu,L.M., Vo,N.V., Kane,C.M., & Chamberlin,M.J. (2003) In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry* **42,** 3777-3786.

Chapter 6

Detection of elongation and termination by single-molecule DNA nanomanipulation

6.1 Background and Objectives

In the previous chapter, we showed by nanomanipulation that we could observe that RNAP escaped from a promoter and entered into elongation. Following elongation, termination occurs during a normal transcription. During transcription termination, RNAP rewinds the melted DNA and releases itself from the template DNA. Each transcription stage involves different sizes of transcription bubbles. Each transcription bubble corresponds to a change in the DNA extension (Δl_{obs}), so we can assign each transcription stage to each corresponding l_{obs}. Therefore, by observing the change of Δl_{obs} with time, we were able to determine the kinetics of elongation and termination events.

6.2 Experimental design and methods

In the last chapter, we used a DNA template- N25-100-tr2^{-} which has a modified *tR2* terminator, and most RNAPs read through the terminator without terminating RNA synthesis. In order to detect the termination events, we construct a DNA template N25-100-tr2, which has N25 promoter, a 100 bp transcribed region and a *tR2* terminator (Figure 6-1). When RNAP transcribes on the N25-100-tr2 and reaches the terminator, transcription termination should occur.

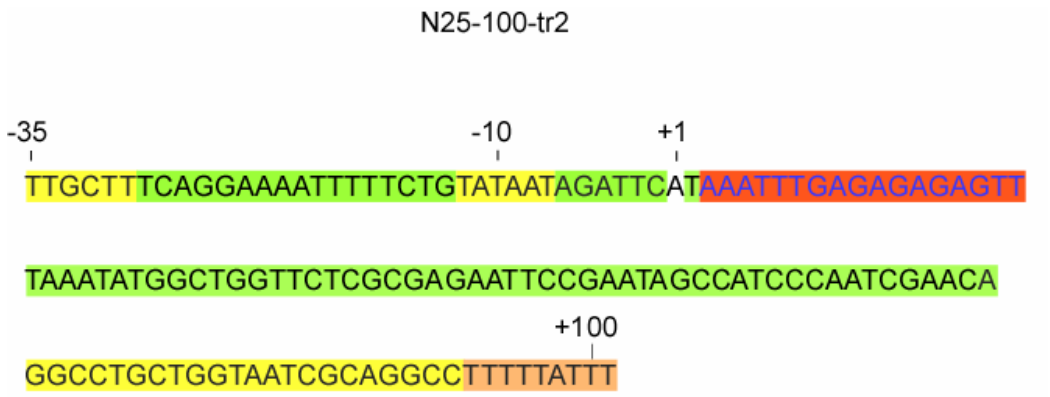


Figure 6-1
N25-100-tr2 DNA templates (top strand): The promoter element -10 and -35 are in yellow. The inverse sequence of *tR2* terminator is in yellow and the poly U corresponded sequence is in orange. The N25-100-tr2 DNA template has a terminator (in orange).

6.3 Results

The time trace for l_{obs} during the transcription carried out on the DNA template

N25-100-tr2 by the single-molecule nanomanipulation is shown in Figure 6-2. This time trace corresponds to a complete transcription event with different transcription stages. These transcription stages and their corresponding Δl_{obs} have been discussed in the previous chapters. We can assign transition 1 as the formation of the open complex, transition 2 as the formation of the initial transcribing complex, and transition 3 as the formation of the elongation complex.

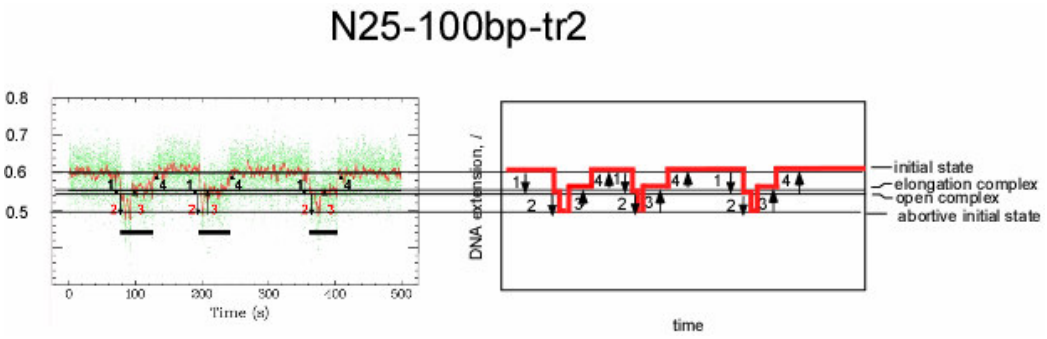


Figure 6-2
The time trace of RNAP transcribing on the N25-100-tr2 DNA template: the transition 1 corresponds to formation of open complex, transition 2 corresponds to formation of ITC, transition 3 is formation of TEC and transition 4 corresponds to termination. The time frame between transition 2 and transition 3 is the abortive initiation state, and the elongation state is between transition 3 and transition 4. Therefore transition 1 to transition 4 is a complete transcription event.

6.4 Conclusions

In order to verify that transition 4 of l_{obs} corresponds to a termination event in Figure 6-2, we compare the time traces of l_{obs} for the two DNA templates. It is clear

that for N25-100-tr2^{-} which has the modified tr2^{-} terminator, the time trace of the l_{obs} after transition 3 could not return to the initial state. From the result of transcription assay experiment, it was observed that most RNAP read through the modified tr2^{-} terminator and continued elongation (Figure 5-3). Thus, the time trace of the l_{obs} after transition 3 corresponds to the elongation state. On the contrary, for the N25-100-tr2 which has the *tR2* terminator, in the transcription essay experiment, the RNAP terminated the synthesis of RNA and released the RNA product when RNAP arrived at the terminator. The trace of l_{obs} for RNAP transcribing on the N25-100-tr2 returns to the initial state. When this occurs, this implies that RNAP completed a transcription cycle on the DNA template and rewound the melted DNA template. Therefore, the transition 4 in Figure 6-2 corresponds to a termination event. The time between transition 3 and transition 4 is the elongation time.

Chapter 7

Determination of kinetics of elongation and termination

7.1 Background and Objectives

In the previous chapters, we have detected transcription elongation and terminator rewinding. In this chapter, I will discuss how to determine the kinetics of these events. We also used single-molecule nanomanipulation to study termination kinetics of DNA templates with various transcribed region lengths.

T is the time for transcription during elongation and terminator rewinding between transition 3 and transition 4 in Figure 6-2. In order to determine the kinetics, we have to know long an elongation stage take. Under the assumption that the elongation time is proportional to the transcribed region length, the distribution of *T* for various transcribed region lengths is Gaussian. For a particular DNA template, *T* can be expressed as:

$T = t_{elongation} + t_{terminator\ \ rewinding} + t_{unscruncing}$ or

$$T = \frac{l}{v_{elongation}} + \frac{1}{v_{terminator\ rewinding}} + \frac{1}{v_{unscruncing}}$$

where $t_{elongation}$, $t_{terminator\ rewinding}$ and $t_{unscruncing}$ are the times corresponding to elongation, terminator rewinding, and promoter escape, respectively. l is the length of the transcribed region. $v_{elongation}$, $v_{terminator\ rewinding}$ and $v_{promoter\ rewinding}$ are the velocities for elongation, terminator rewinding, and promoter escape, respectively. In Chapter 5.4.1, it has been shown that $v_{promoter\ rewinding}$ here is much larger than the $v_{elongation}$ and $v_{terminator\ rewinding}$. Thus, the promoter rewinding event can occur within the time resolution of the nanomanipulation. And in Chapter 8.3.3, it will show that the different promoters will not affect the elongation plus terminator rewinding time, which implies the time of promoter rewinding is very short comparing to the elongation time plus terminator rewinding time. Therefore, $t_{promoter\ rewinding}$ can be ignored in the measurement. By increasing the length of the transcribed region *l,* it will prolong the time for elongation $t_{elongation}$. We are trying to demonstrate the relationship between the transcribed region length of the DNA template and the time *T,* as described by $T = \frac{l}{v_{elongation}} + \frac{1}{v_{terminator\ rewinding}} + \frac{1}{v_{promoter\ rewinding}}$. (equation 7-1)

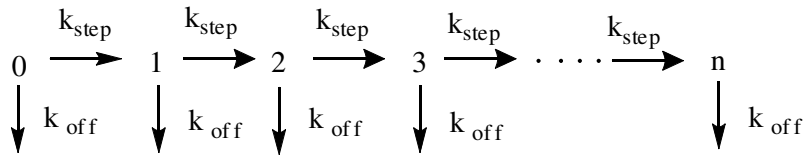
To prove this relationship, we constructed DNA templates with the N25 promoter, *tR2* terminator, and various transcribed region lengths from 160 bp-340 bp

(Figure 7-1). The transcribed regions of the DNA templates had identical repeat sequence, but various repeat numbers. Since each template had different numbers of tandem repeat sequences in the transcribed region, $t_{elongation}$ should increase proportionately to the transcribe region length.

We measure *T* and calculate the average of *T* for RNAP transcription over various transcribed region lengths of the DNA templates shown in table 7-1. Figure 7-2 shows that RNAP elongated on a template DNA at a heterogeneous velocity, which agrees with the study in [1]. However, the width of the distribution increased proportionately to the length of transcribed region (Figure 7-2). According to Eq 7-1, the $t_{terminator\ rewinding}$ for the *tR2* terminator can be extracted by making a linear regression plot for each *T* over various transcribed region lengths.

7.1.1 Statistical model

Another way to study the elongation of RNAP on the DNA template can be described by the following statistical model (Scheme 7-1) [2].



Scheme 7-1

In the model, the RNAP begins translocation at 0 and moves one step forward each time and can not move back. Each step normally corresponds to the incorporation of a single nucleotide. The k_{step} is the rate constant for each step on the DNA template. Here, we assume the k_{step} is identical for each step, and that each step is statistically independent. The k_{off} is the rate constant for other reaction pathways like a transcription pausing or arrest. Then the probability of RNAP moving along the DNA template can be expressed by the equations derived as follows:

For n=0:

$$\frac{dP_0(t)}{dt} = -(k_{step} + k_{off}) \cdot P_0 \quad \text{(equation 7-2)}$$

$$\frac{dP_0}{P_0} = -(k_{step} + k_{off})dt \quad \text{(equation 7-3)}$$

Integrate equation 7-3, one can obtain:

$$P_0 = \exp[-(k_{step} + k_{off}) \cdot t] \quad \text{(equation 7-4)}$$

For n=1:

$$\frac{dP_1(t)}{dt} = (k_{step} \cdot P_0) - (k_{step} + k_{off}) \cdot P_1 \quad \text{(equation 7-5)}$$

$$\frac{dP_1(t)}{dt} + (k_{step} + k_{off}) \cdot P_1 = k_{step} \cdot P_0 \quad \text{(equation 7-6)}$$

Substitute equation 7-6 with $P_0 = \exp[-(k_{step} + k_{off}) \cdot t]$ which is obtained by integration of equation 7-3. One can have:

$$\frac{dP_1(t)}{dt} + (k_{step} + k_{off}) \cdot P_1 = k_{step} \cdot \exp[-(k_{step} + k_{off}) \cdot t] \quad \text{(equation 7-7)}$$

Divide both sides of (equation 7-7) with $\exp[-(k_{step} + k_{off}) \cdot t]$

Then it gives (equation 7-8)

$$\exp[(k_{step} + k_{off}) \cdot t]\frac{dP_1(t)}{dt} + [\exp(k_{step} + k_{off}) \cdot t](k_{step} + k_{off}) \cdot P_1 = k_{step} \quad \text{(equation 7-8)}$$

Again, integrate (equation 7-8)

$$\frac{d[[(\exp(k_{step} + k_{off}) \cdot t] \cdot P_1(t)]}{dt} dt = k_{step} \cdot dt \quad \text{(equation 7-9)}$$

One can have

$$P_1(t) = [\exp-(k_{step} + k_{off}) \cdot t] \cdot k_{step} t \quad \text{(equation 7-10)}$$

By extension, one can obtain

$$P_n(t) = (k_{step} t)^n \frac{(-\exp(k_{step} + k_{off}) \cdot t)}{n!} \text{(equation 7-11)}$$

During elongation, the rate constant for reaction leading to a transcription pause or arrested state is small, k_{off} can be ignored:

$$P_n(t) = (k_{step} t)^n \frac{-\exp(k_{step} t)}{n!}$$

Thus, in this model, after transcribing for a time t, the probability of the RNAP being on the n^{th} nucleotide position of the DNA template follows a Poisson distribution.

7.2 Experimental design and methods

We prepared DNA constructs: N25-160-tr2, N25-220-tr2, N25-280-tr2, and N25-340-tr2 by increasing the number of nucleotides in the transcribed region at 60 bp increments (Figure 7-1). By carrying out transcription on these DNA templates, the termination and elongation kinetics were determined, which will be discussed in the following section.

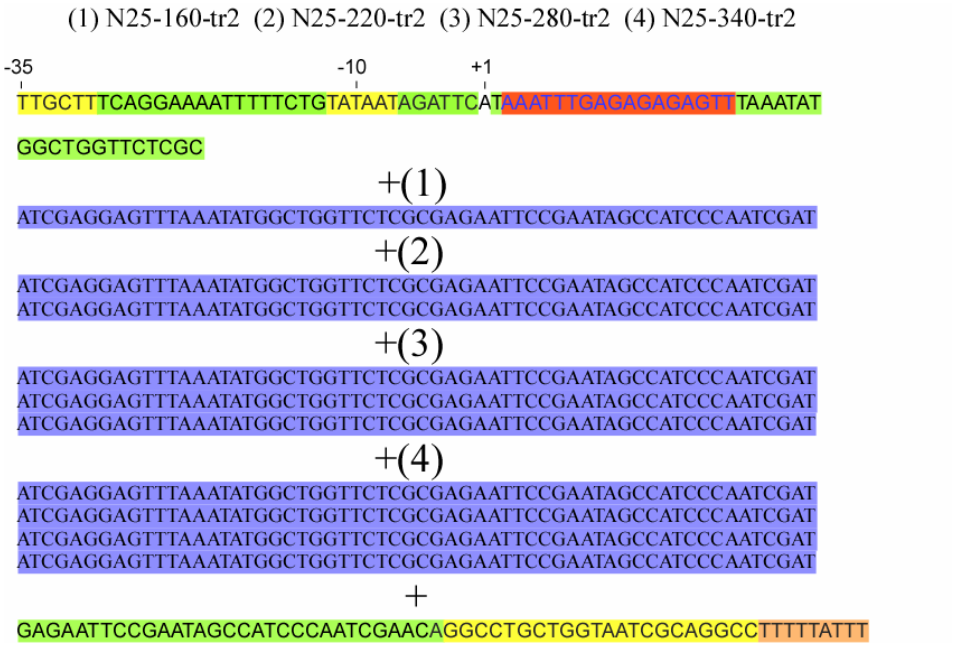


Figure 7-1

The constructs of the N25-160~340-tr2 has the elements in order(top strand):

The promoter elements: -10 and -35 are in yellow, the ITS is in the red,

(1) The N25-160-tr2 has one 60 bp (+13 to +66 in N25-100-tr2) insert,

(2) The N25-220-tr2 has two 60 bp (+13 to +66 in N25-100-tr2) inserts,

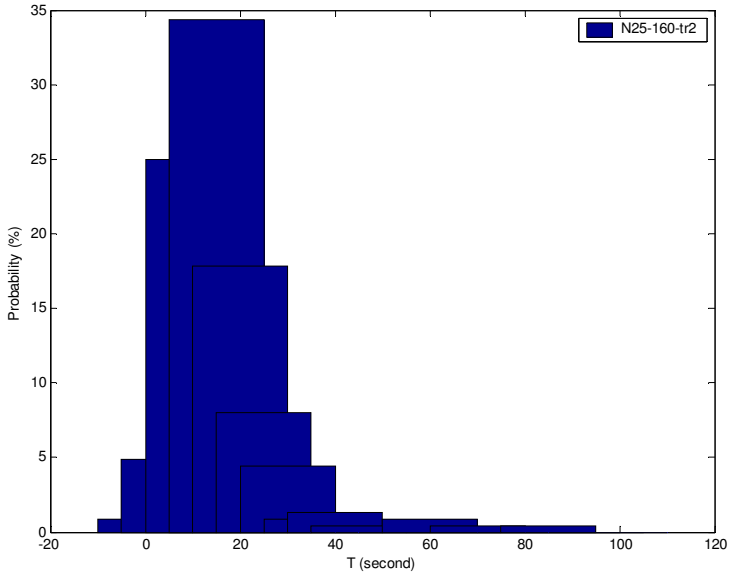
(3) The N25-280-tr2 has three 60 bp (+13 to +66 in N25-100-tr2) inserts,

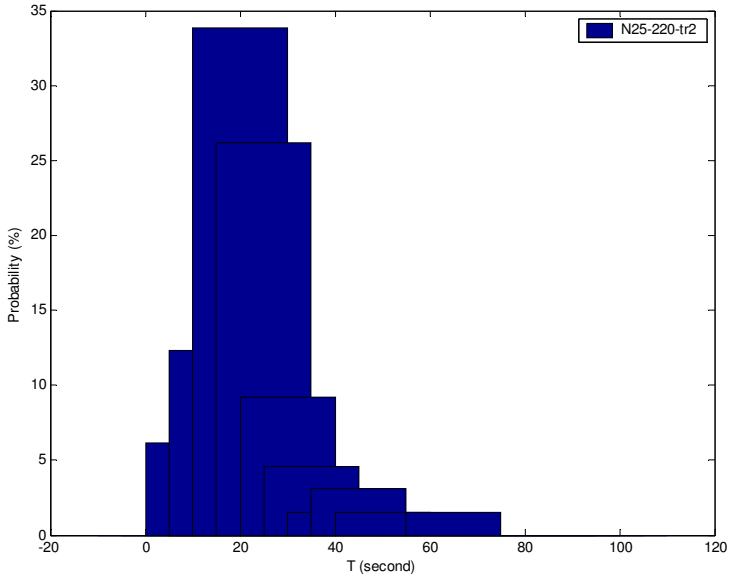
(4) The N25-340-tr2 has four 60 bp (+13 to +66 in N25-100-tr2) inserts,

the inverse sequence of *tR2* terminator is in yellow and the T-tract sequence is in orange.

7.3 Results

The distribution of T for each transcribed region length is shown in Figure 7-2.


N25-160-tr2
Probability (%)
T (second)
35
30
25
20
15
10
5
0
-20
0
20
40
60
80
100
120


N25-220-tr2
Probability (%)
T (second)
35
30
25
20
15
10
5
0
-20
0
20
40
60
80
100
120

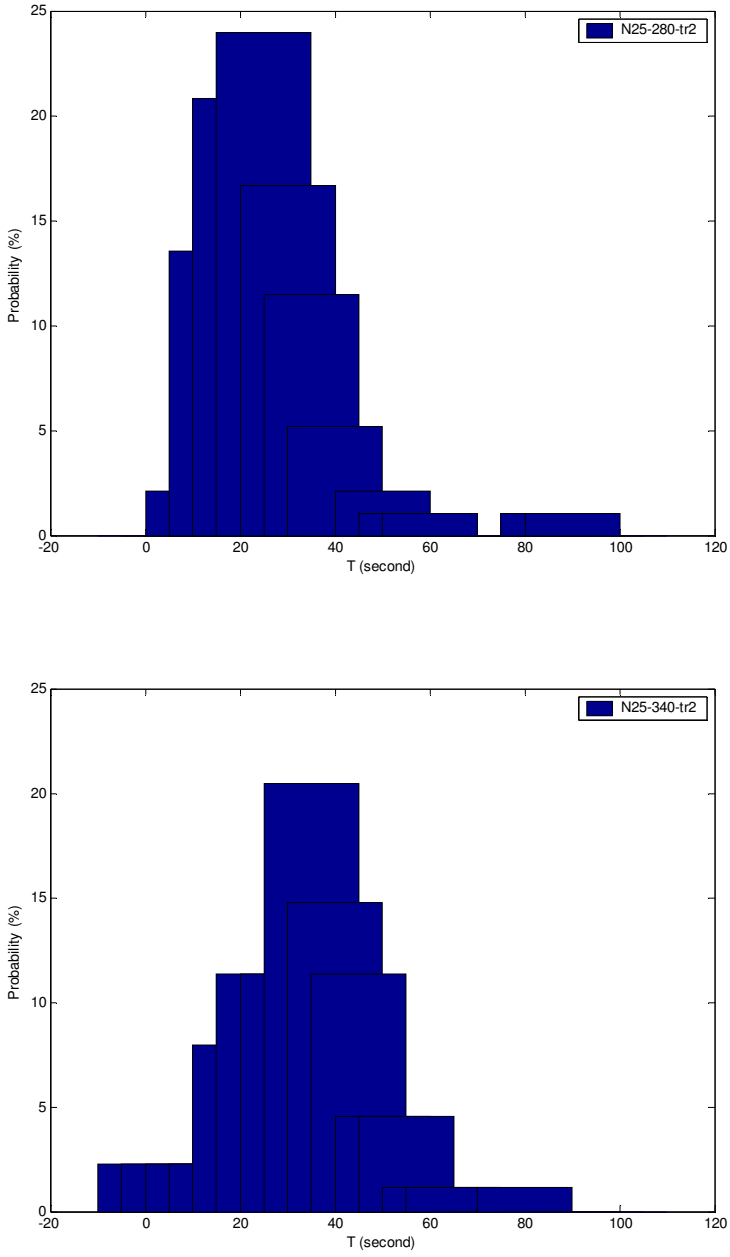


Figure 7-2

The distributions of T with each different length of transcribed region of DNA templates: the distributions were normalized by the number of data for each particular length of transcribed region of template DNA.

7.3.1 Data analysis

7.3.1.1 The analysis based on linear regression

Each data from the same transcribed region length was treated as a Gaussian distribution. After fitting the data with a Gaussian distribution, the outliers were defined as data beyond 3 times the standard deviation from mean. After removing these outliers, the average of each set of data is shown in Table 7-1 (B). Even with this adjustment, the average of data still cannot be fitted well into a linear function. In order to further reduce the deviation from a linear function, the means of the data were obtained by fitting the rest of data into a Gaussian distribution. In this Gaussian distribution fitting, each mean value corresponds to the peak of each Gaussian distribution and is shown in Table 7-1 (A). The reason for us to treat the data in the two different ways (the results are shown in Table 7-1 (A) and (B)) is that the mean values obtained from Gaussian distributions fitting are different with the average of the data for which the outliers have been removed. The mean value for each transcribed region length of DNA template was used in the linear regression, from which the *tR2* terminator rewinding time and elongation velocity were determined

	N25-100-tr2	N25-160-tr2	N25-220-tr2	N25-280-tr2	N25-340-tr2	N25-400-tr2
A	12.5s	19.2s	26.0s	29.4s	37.0s	42.0s
B		21.4s	26.8s	32.3s	42.0s	
C		21.6s	27.2s	32.7s	42.5s	

Table 7-1 The *T* over various transcribed region lengths ((A) The T is the mean of the data and is obtained by fitting the data into a Gaussian distribution after removing outliers. (B) The *T* is the average of data after removing the outliers. (C) The *T* is the average of data and obtained without removing any raw data.)

T is linearly correlated with the nucleotide number of the transcribed region. From these results, the elongation velocity of RNAP on the N25-100~400-tr2 DNA templates is around 10± 1 bp/s , and the terminator rewinding time of *tR2* terminator is 3.4 ± 2.9 seconds (Figure 7-3).

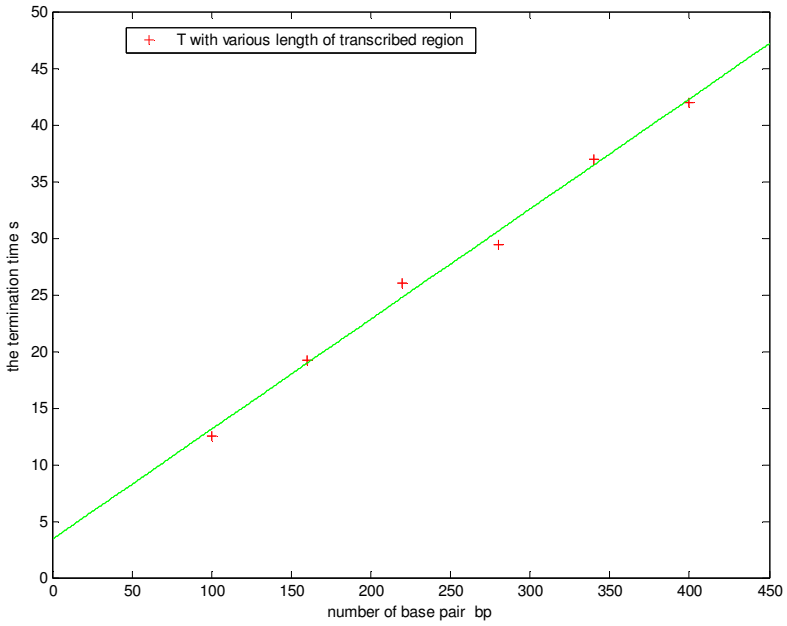


Figure 7-3
The linear regression of the data which correspond to each particular length of transcribed region of template DNA: N25-100~340-tr2.

7.3.1.2 The analysis based on Poisson Statistics

Each data set from the same transcribed region lengths was treated as a Poisson distribution. In this case, the fitting using a general linear model has a link function –log. The distributions of T over different transcribed region lengths are shown in Figure 7-2. The result of fitting of General linear model with Poisson distribution is: $\log(T) = 2.48 + 0.0037 \cdot l$ or $T = e^{2.48+0.0037 \cdot l}$ Here, the l is the length of transcribed region. The elongation velocity from the fitting of the data can be expressed as $v = \frac{\partial l}{\partial T} = \frac{1}{0.0037} \cdot e^{-(0.0037 \cdot l + 2.48)}$. Thus, the elongation velocity is not

constant along the DNA template. Instead, the elongation velocity decreases with an increase in the nucleotide number of the transcribed region. In the transcribed region 160 bp ~340 bp, the velocity varies from 12.5 bp/s to 6.4 bp/s. The time for terminator unwinding is 11.94±0.35 second.

7.5 Conclusions

The elongation velocity and the *tR2* terminator rewinding time determined by single-molecule nanomanipulation are 3.4 seconds and 10 bp/s, respectively, using linear regression. The elongation velocity here is close to the elongation velocity previously determined by single molecule technique [3]. The termination rewinding time measured here is lower than termination time determined from other terminators (14~8 seconds) by conventional method [4]. A possible explanation for the termination rewinding time being lower than the termination time could be the involvement of other termination processes and the other processes take a comparable order of time. A complete termination event can involve cessation of RNA synthesis, release of RNA transcript, and release of the RNAP from the DNA template. The release of the RNA transcript can correspond to the collapse of elongation bubble or DNA rewinding, which can reflect on the Δl_{obs}. If the termination time for *tR2* terminator is between 14~8 seconds, it implies the release of the RNAP from the DNA

template plays an essential role in the kinetics of termination.

Using the Poisson regression, the terminator rewinding time for *tR2* is 11.9 seconds, which is much closer to the observed termination time for other terminators. However, the regression was based on an assumption that the kinetics of elongation is a Poisson process, such that the elongation velocity of RNAP is not constant. Instead, RNAP moves more quickly at the beginning of transcription than at the end of transcription. It might be able to explain why lower elongation velocity was obtained from most single-molecule experiments while higher elongation velocity was determined from most of conventional experiments. Because, in these single-molecules experiments, the elongation velocity was obtained by providing DNA templates with a length of a couple kb in transcribed region while in most of conventional experiments the elongation velocity was obtained by using the DNA templates with shorter transcribed region length. A Poisson process has also been observed with DNA polymerase synthesizing DNA fragments [5]. Thus, a similar model may apply to transcription as well.

References

[1] Adelman,K., La,P.A., Santangelo,T.J., Lis,J.T., Roberts,J.W., & Wang,M.D. (2002) Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc. Natl. Acad. Sci. U. S. A* **99,** 13538-13543.

[2] Szczelkun,M.D. (2002) Kinetic models of translocation, head-on collision, and DNA cleavage by type I restriction endonucleases. *Biochemistry* **41,** 2067-2074.

[3] Bai,L., Santangelo,T.J., & Wang,M.D. (2006) Single-molecule analysis of RNA polymerase transcription. *Annu. Rev. Biophys. Biomol. Struct.* **35,** 343-360.

[4] Arndt,K.M. & Chamberlin,M.J. (1988) Transcription termination in Escherichia coli. Measurement of the rate of enzyme release from Rho-independent terminators. *J. Mol. Biol.* **202,** 271-285.

[5] Peller,L. (1977) Thermodynamic limits on the size and size distribution of nucleic acids synthesized in vitro: the role of pyrophosphate hydrolysis. *Biochemistry* **16,** 387-395.

Chapter 8

Determination of effects of promoter sequence on initial transcription, promoter escape, elongation, and termination

8.1 Background and Objectives

The N25 DNA construct and the $N25_{antiDSR}$ construct are identical in the promoter recognition region but the downstream region of the N25 promoter from +3 to +20 has been replaced with $A \leftrightarrow C / G \leftrightarrow T$. RNAP 's initial transcription on the N25 and the $N25_{antiDSR}$ promoters has been well characterized by conventional experiments [1-3]. RNAP on the $N25_{antiDSR}$ construct has a different kinetic behavior than the N25 construct during initial transcription. For instance, transcription has a lower abortive to productive ratio (APR) on the N25 promoter than on the $N25_{antiDSR}$ promoter. The maximum size of the initial transcript for the N25 promoter is 10 nt, while it is 15 nt for the $N25_{antiDSR}$ promoter. The abortive probability is also different between the two promoters (Figure 8-1).

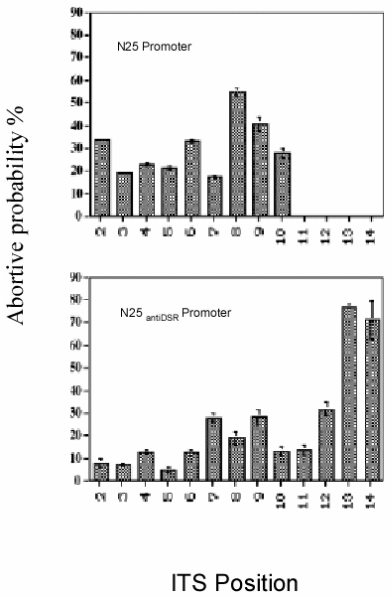


Figure 8-1
The profile of abortive probability over different ITS position for the N25 promoter and the N25 $_{antiDSR}$ promoter. Figure from ref [4].

These kinetic differences can be studied by nanomanipulation. It is also unknown whether the difference in promoters affects the kinetics of transcription elongation and termination steps. Thus, in this chapter, we investigate the kinetics of transcription with the N25 and the N25$_{antiDSR}$ promoter.

8.2 Experimental design and methods

We used the N25-160-tr2, N25-220-tr2, N25-280-tr2 and N25-340-tr2 constructs as the DNA templates in order to characterize the transcription kinetics of RNAP on the N25 promoter. These DNA templates have identical promoters, but in the length of the transcribed region. In order to characterize the kinetics of RNAP

on the $N25_{antiDSR}$ construct, we used $N25_{antiDSR}$-160-tr2 as the DNA template. The construction of N25-160~340-tr2 and $N25_{antiDSR}$-160-tr2 DNA template are shown in Figure 7-1 and Figure 8-2, respectively.

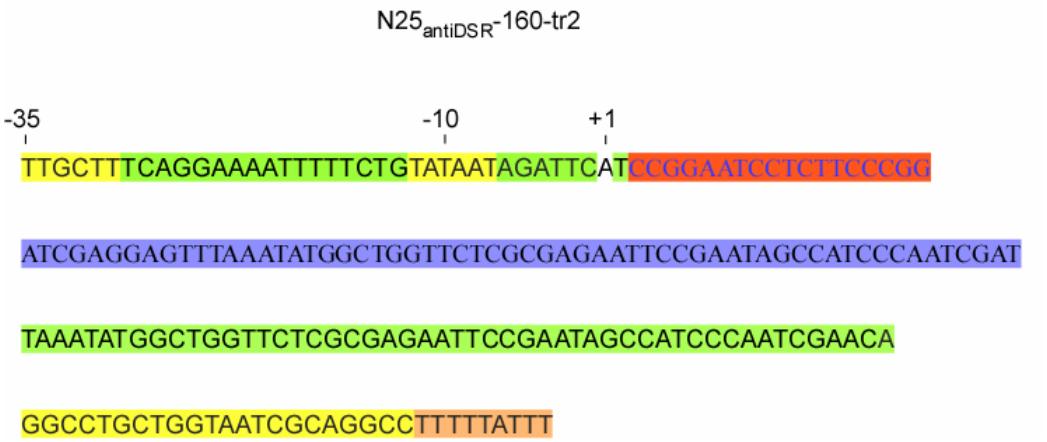


Figure 8-2
The constructs of the $N25_{antiDSR}$-160-tr2 (Top strand)
The promoter element -10 and -35 are in yellow. The ITS is in the red. A 60 bp (+13 to +66 in N25-100-tr2) insert is blue. The inverse sequence of *tR2* terminator is in yellow and the poly U corresponded sequence is in orange.

The kinetics of transcription can be extracted from the time trace of l_{obs} while a transcription event occurs. For instance, the lifetime of the abortive initiation state was determined by the distribution of time between transition 2 and transition 3 of the RNAP transcription time trace (Figure 6-2). The outliers in the collected data, again defined as data beyond 3 times the standard deviation, were removed before further data processing. The lifetime of initial transcription for DNA templates with the

N25 or $N25_{antiDSR}$ promoters was obtained by fitting the remaining data with a Gaussian distribution.

The DNA extension caused by abortive initiation was obtained by adding the Δl_{obs} of transition 1 and transition 2 from the time trace (Figure 6-2). However, due to the spatial resolution or other unknown reasons, the l_{obs} did not always return to the initial level after a complete transcription event, making it difficult to precisely determine the DNA extension caused by abortive initiation. In order to have a more precise determination for the DNA extension caused by abortive initiation, the l_{obs} time trace is treated as a datum only when the difference between l_{obs} immediately before transcription initiation and immediately after termination is less than 20 nm. As before, the outliers in the collected data were defined as data beyond 3 times the standard deviation and were removed before further data processing.

8.3 Results

8.3.1 The open complex on the N25 promoter vs. the N25 $_{antiDSR}$ promoter

The result of open complex formation shows that it caused a larger DNA extension on the N25$_{antiDSR}$-160-tr2 construct than on the N25-160~340-tr2 constructs (Table 8-1, Figure 8-2). This could be due to the larger size of the transcription bubble on the N25 $_{antiDSR}$ promoter than on the N25 promoter. Alternatively, the N25 $_{antiDSR}$ promoter could cause larger DNA compaction in the open complex.

	N25-160~340-tr2	**N25$_{antiDSR}$-160-tr2**
DNA Extension(nm)	82 ± 24	110 ± 29

Table 8-1
The change in DNA extension caused by the formation of open complex on the N25 and N25$_{antiDSR}$ promoter

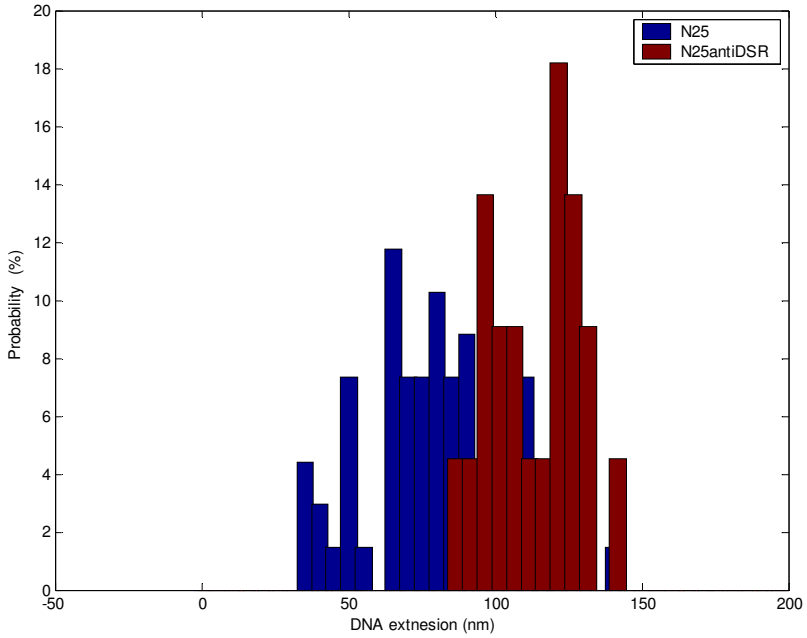


Figure 8-3
The distribution of DNA extensions caused by the formation of open complex on the N25 promoter and the $N25_{antiDSR}$ promoter. (See Appendices A, B, C, and D)

8.3.2 Transcription initiation on the N25 promoter and the $N25_{antiDSR}$ promoter

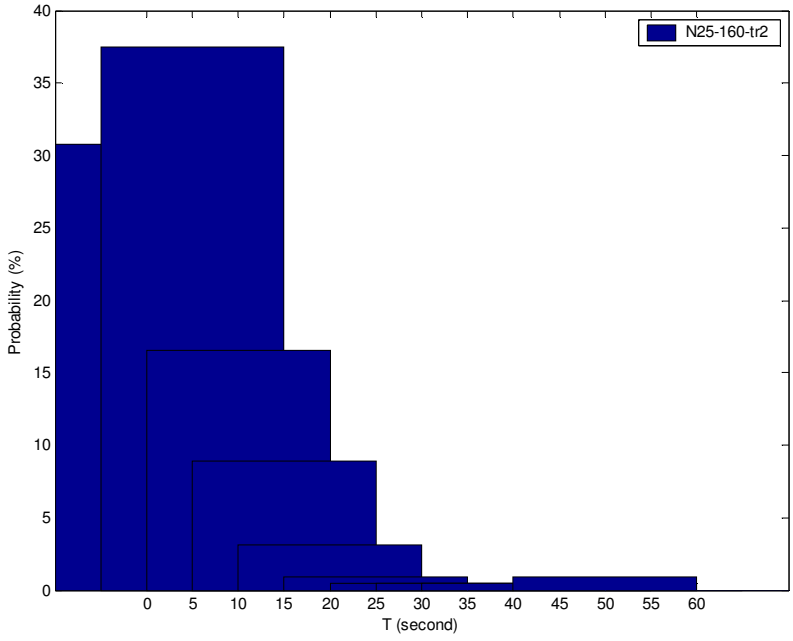
RNAP on the $N25_{antiDSR}$ promoter took 6 times as long to escape from the promoter than on the N25 promoter (table 8-2). The distributions of initial transcription lifetime on the DNA templates of N25-160~340-tr2 are very similar (Figure 8-4). Figure 8-5 shows the similarity among the DNA extension distributions. Again, this indicates that the initial transcription was not affected by the DNA sequence downstream of the N25 promoter on the N25-160~340-tr2 DNA

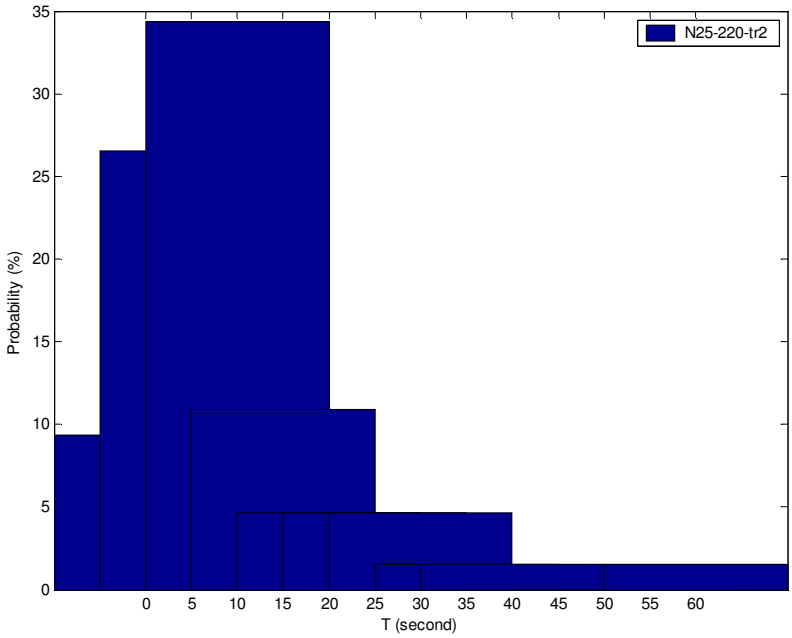
templates.

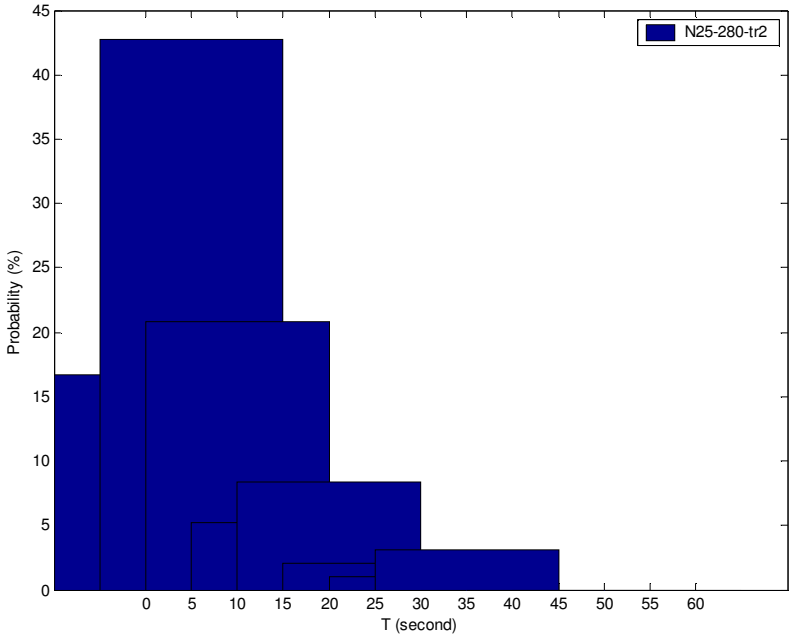
However, these distributions are quite different from the $N25_{antiDSR}$ promoter (Figure 8-6). The lifetime distribution of initial transcription state on the $N25_{antiDSR}$ promoter has more populations with a longer lifetime than the N25 promoter.

DNA template	N25-160-tr2	N25-220-tr2	N25-280-tr2	N25-340-tr2	N25 *	$N25_{antiDSR}$-160-tr2
T (second)	9 ± 8	14 ± 11	11 ± 8	15 ± 11	11 ± 10	69 ± 64

Table 8-2
The lifetime was obtained by fitting the data from N25-160~340-tr2 with removing outliers.




N25-220-tr2
Probability (%)
T (second)
0
5
10
15
20
25
30
35
40
45
50
55
60


N25-280-tr2
Probability (%)
T (second)
0
5
10
15
20
25
30
35
40
45
50
55
60

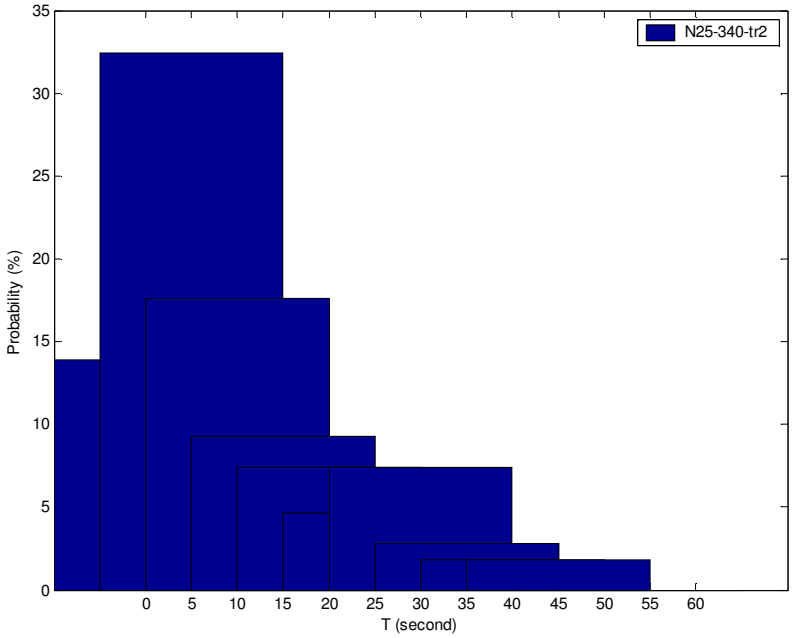


Figure 8-4

The distribution of the lifetime of abortive initiation state for each N25 promoter DNA template is very similar. The distributions shown here are normalized. (See Appendices A, B, C and D)

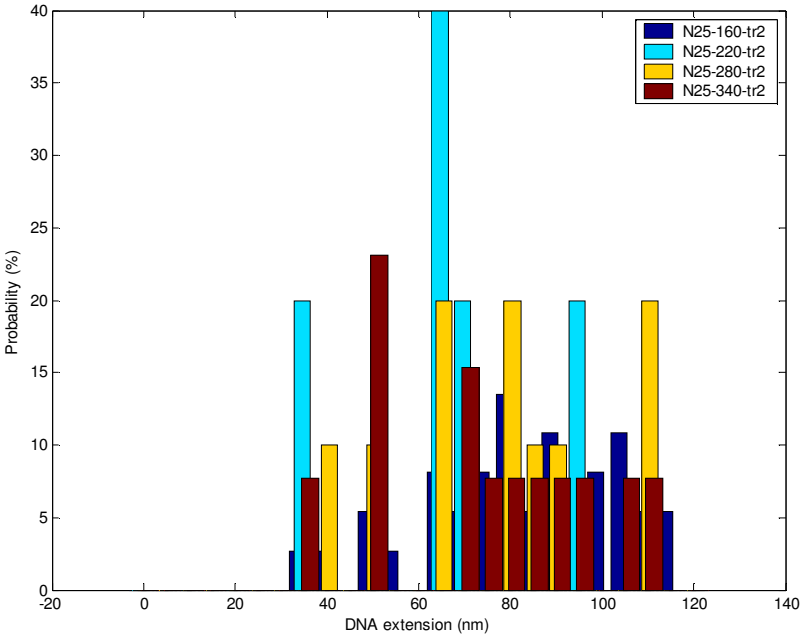


Figure 8-5

The distributions shown here indicate that DNA extension was not affected by the downstream sequence of N25 promoters. The distributions are normalized. (See Appendices A, B, C, and D).

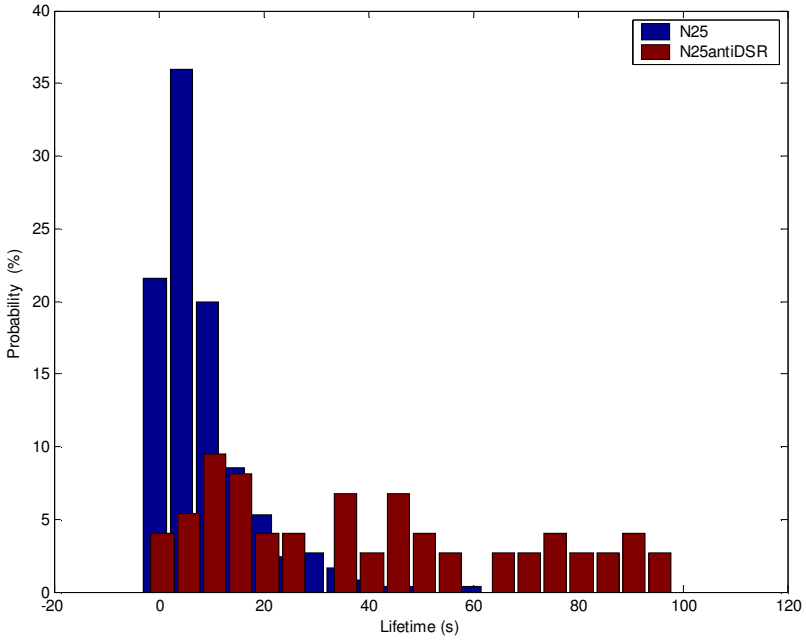


Figure 8-6
The lifetime distribution of initial transcription on N25antiDSR-160-tr is broad, On the contrary, the distribution of initial transcription is sharp on the N25-160~340-tr2 DNA templates. This distribution has been normalized. (See Appendices A, B, C, and D).

Interestingly, the distribution of DNA extension caused by ITC formation on $N25_{antiDSR}$ promoter has two peaks (Figure 8-7). If a peak can represent a population of a intermediate, this seems to imply the presence of two intermediates populations on the $N25_{antiDSR}$ promoter during abortive initiation. The distribution of ITC on the N25 promoter has only a single peak.

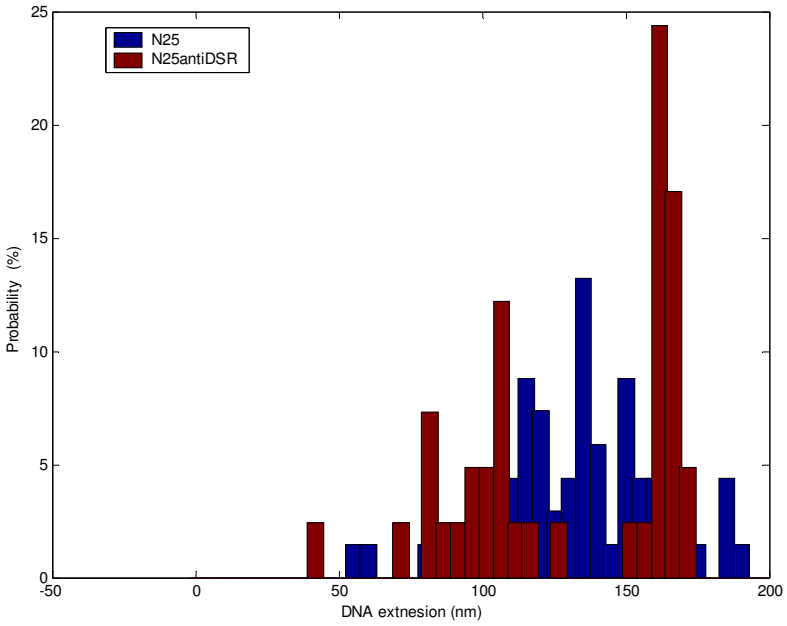


Figure 8-7
The distribution of change in DNA extension caused by the formation of ITC on the N25 promoter and the $N25_{antiDSR}$ promoter. The distributions here are normalized. (See Appendix A and F)

8.3.3 Elongation and termination on the N25 promoter and $N25_{antiDSR}$ promoter

The results of the elongation plus the termination time for the N25-160-tr2 construct and the $N25_{antiDSR}$-160-tr2 construct are shown in Figure 8-8 and Table 8-3.

	N25-160-tr2	**$N25_{antiDSR}$-160-tr2**
T(second)	19 ± 10	18 ± 6

Table 8-3
The elongation plus the termination time on N25-160-tr2 and $N25_{antiDSR}$-160-tr2

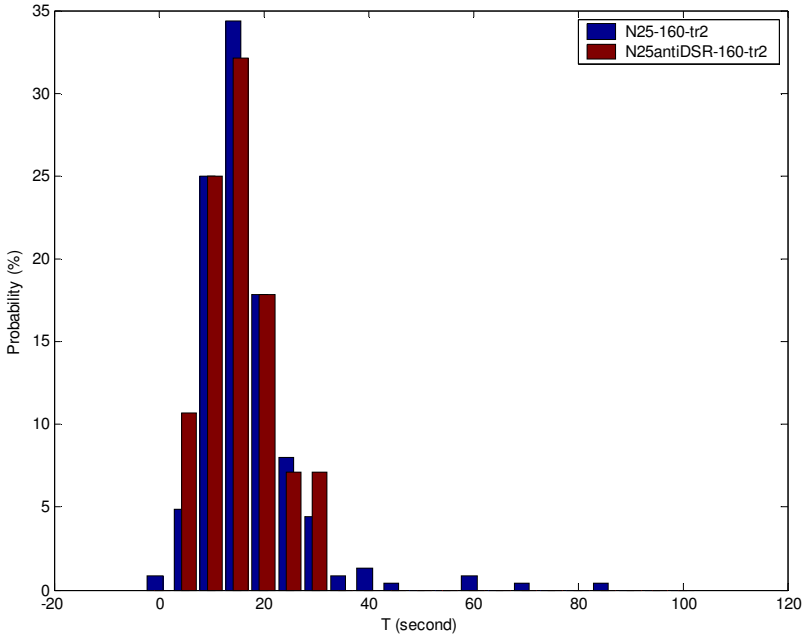


Figure 8-8
The distributions of elongation time plus termination time on N25-160-tr2 and $N25_{antiDSR}$-160-tr2. The distributions are normalized. (See Appendix A and F)

Although the N25-160-tr2 and the $N25_{antiDSR}$-160-tr2 have different promoters, it is clear that the elongation time plus termination time for the two DNA templates is very close, their distributions are almost identical for the two DNA templates. In addition, the DNA extension change caused by the formation of TEC were the same for N25-160-tr2 and $N25_{antiDSR}$-160-tr2 (Figure 8-9) (Table 8-4).

	N25-160~340-tr2	**N25antiDSR-160-tr2**
DNA Extension (nm)	**58 ± 12**	**65 ± 10**

Table 8-4
The DNA extension caused by the formation of TEC on the N25 promoter and the $N25_{antiDSR}$ promoter

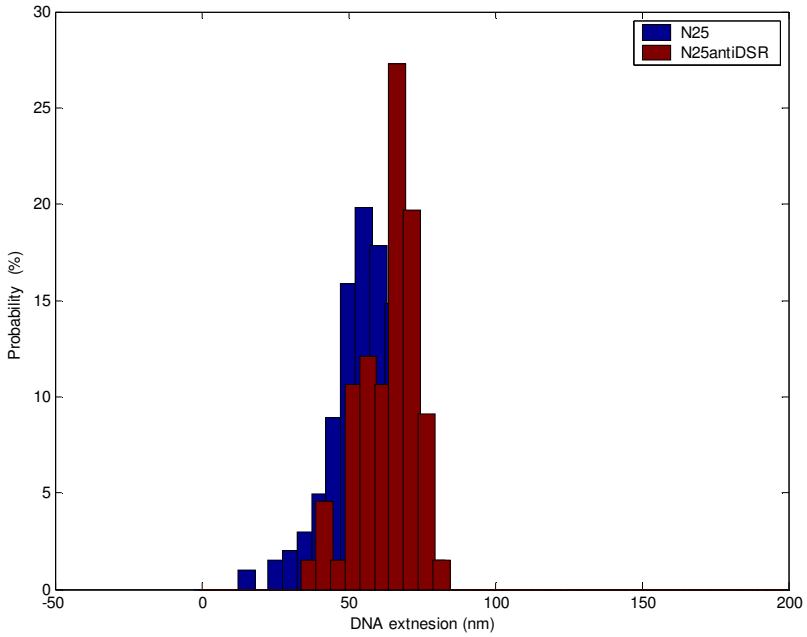


Figure 8-9
The distributions of DNA extension was caused by the formation of TEC on the N25 promoter and the $N25_{antiDSR}$ promoter

8.4 Conclusions

The difference between the $N25_{antiDSR}$ construct and the N25 construct is the ITS, which is involved with DNA scrunching. The result here implies that ITS alters the kinetic energy barrier for promoter escape. The distributions of bubble size, DNA compaction are different, reflecting difference in stressed intermediates during initial transcription. The difference in promoter, however, does not affect the rate of transcription elongation once RNAP has escaped the promoter. In addition, the result implies that the time of a promoter rewinding event is much shorter than the

elongation and termination time because difference ITS des not affect the kinetics of elongation and termination.

References

1. Vo N.V., Hsu L.M., Kane C.M., and Chamberlin M.J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. Biochemistry *42*: 3798-3811.

2. Vo N.V., Hsu L.M., Kane C.M., and Chamberlin M.J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 2. Formation and characterization of two distinct classes of initial transcribing complexes. Biochemistry *42*: 3787-3797.

3. Hsu L.M., Vo N.V., Kane C.M., and Chamberlin M.J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. Biochemistry *42*: 3777-3786.

4. Hsu L.M., Vo N.V., Kane C.M., and Chamberlin M.J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. Biochemistry *42*: 3777-3786.

Chapter 9

Determination of the effects of elongation factor GreB on initial transcription and promoter escape

9.1 Background and Objectives

A moribund complex is an ITC which cannot generate full length of RNA products [3]. Some of ITCs can only generate nascent RNA. Such ITCs are called moribund complex. It has been shown that GreB can reduce the number of moribund complexes at transcription initiation. Instead the moribund complex will enter an arrested state with short RNA transcripts [3].

It has also been shown that GreB can reduce the lifetime of transcription initiation by 2~5 fold on the N25 promoter [2]. In this chapter, we studied the kinetics of transcription initiation and elongation on the N25 promoter in the presence and absence of GreB.

9.2 Experimental design and methods

We used the N25-160~340-tr2 DNA templates to determine the kinetics of transcription with GreB by including GreB in the transcription assay. The N25-100-tr2 DNA template was used to determine transcription kinetics in the absence of GreB. The design of N25-160~340-tr2 is shown in Figure 7-1 and N25-100-tr2 is shown in Figure 4-1.

The methods to determine the lifetime of the abortive initiation state and the lifetime of elongation were the same as in Chapters 8. We also used the same method to measure the changes of DNA extension on these DNA templates. All of the data were processed in the same manner as in Chapter 8.

9.3 Results

9.3.1 Abortive initiation with/without GreB

In our study, upon addition of GreB into the transcription assay, the lifetime of ITC was reduced by one half (Table 9-1). The populations of ITC with longer lifetimes for the N25 promoter in the presence of GreB decay quickly (Figure 9-1).

	N25-160~340-tr2 with GreB	N25-100-tr2 without GreB
T(second)	11±9	21±26

Table 9-1
The lifetime of abortive initiation on the N25 promoter with and without GreB

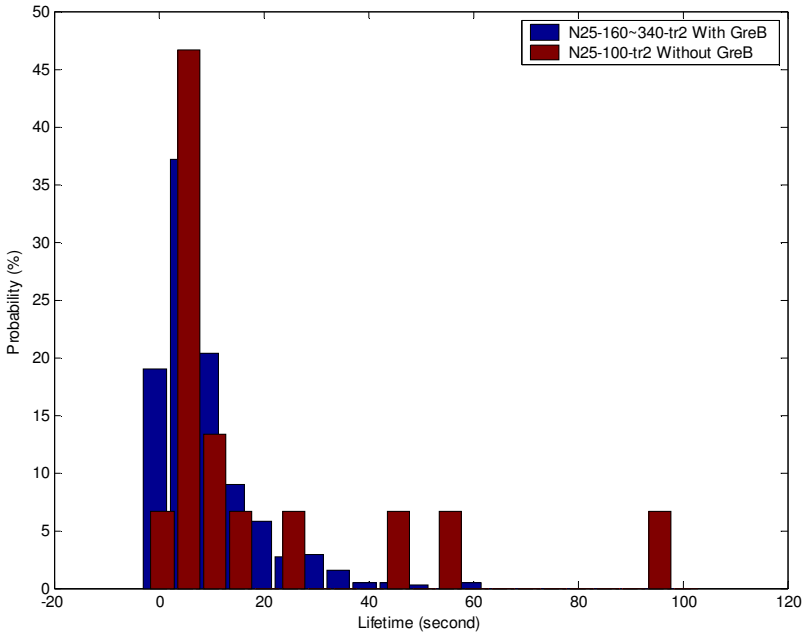


Figure 9-1
The lifetime distribution of the abortive initiation state on the N25 promoter with and without GreB. (See Appendices A – E)

Changes of DNA extension are proportional to the transcription bubble size or DNA compaction. The transcription bubble size or DNA compaction was not significantly affected by the presence or absence of GreB (Table 9-2). The amount of data used to determine the kinetics in the presence of GreB is much more than that

used in the absence of GreB. In the both cases, however, the change of DNA extension caused by ITC formation has similar values, and one is still able to observe similarity between the two distributions. Figure 9-2 shows the l_{obs} does not change much with the N25 promoter during transcription in the absence of GreB in the presence and absence of GreB.

	N25-160~340-tr2 with GreB	**N25-100-tr2 without GreB**
DNA Extension(nm)	118 ± 26	109± 26

Table 9-2
The DNA extension caused by the formation of ITC on N25 promoter with and without GreB

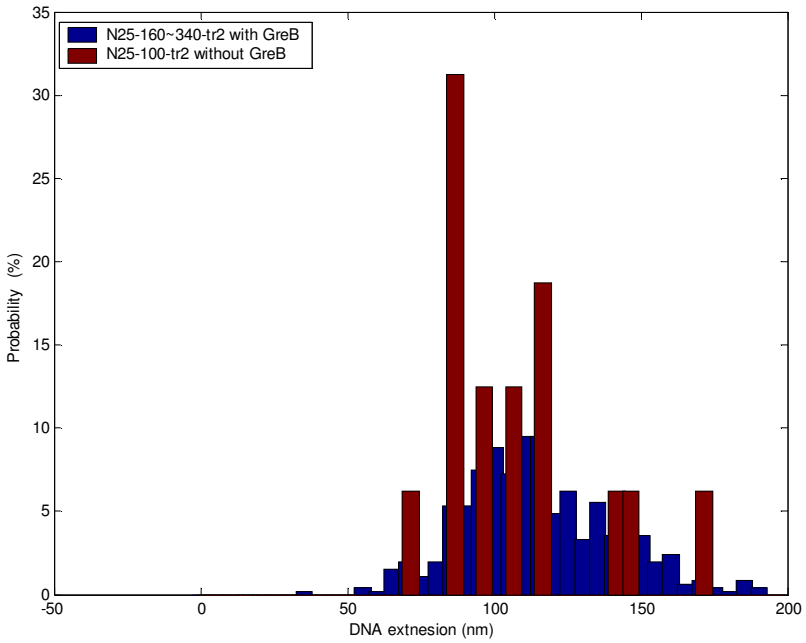


Figure 9-2
The distributions of DNA extension caused by the formation of ITC on the N25 promoter with and without GreB. (See Appendices A – E)

9.3.2 Elongation and termination with and without GreB

The DNA extension caused by the formation of elongation complex is almost the same in the presence or absence of GreB (Table 9-3) (Figure 9-3). This implies that the bubble size or DNA compaction of TEC are not be affected by GreB.

	N25-160~340-tr2 with GreB	N25-100-tr2 without GreB
DNA Extension(nm)	58 ± 11	54 ± 18

Table 9-3
The DNA extension caused by the formation of TEC on the N25 promoter with and without GreB

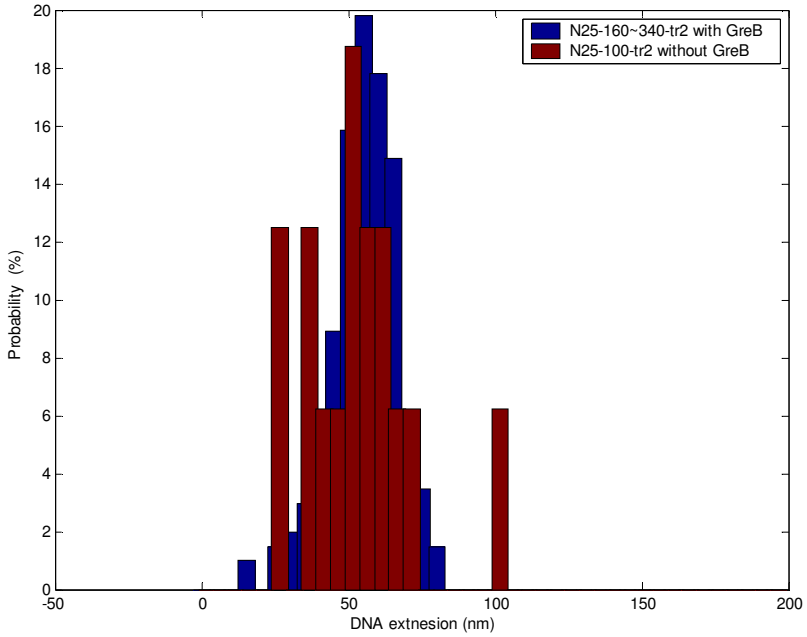


Figure 9-3
The distributions of DNA extension caused by the formation of ITC on the N25 promoter with and without GreB. (See Appendices A – E)

9.4 Conclusions

These results clearly indicate that GreB enhances promoter escape during the abortive initiation step, and the lifetime of abortive initiation state is sensitive to the presence of GreB. Since the lifetime of abortive initiation can correspond to the

number of abortive initiation cycles, these results show that GreB can reduce the number of abortive initiation cycles. Alternatively, the long lifetime state may correspond to a moribund complex, with GreB restoring the complex to a productive complex. The results also indicate that GreB does not affect the size or DNA compaction of ITC and TEC. It has been shown that GreB reduces the abortive probability of nascent RNA products [1]. Therefore, GreB may play the same role during abortive initiation as in transcription elongation- it helps to activate the backtracking complex and regenerate the active site of RNAP.

References

[1] Hsu,L.M., Cobb,I.M., Ozmore,J.R., Khoo,M., Nahm,G., Xia,L., Bao,Y., & Ahn,C. (2006) Initial transcribed sequence mutations specifically affect promoter escape properties. *Biochemistry* **45,** 8841-8854.

Chapter 10

Materials and Methods

10.1 The construction of plasmid DNA templates

pARn25-100-tr2 (pARn25$_{antiDSR}$-100-tr2) and its multiple 60 base pair insert derivatives were used as transcription templates in this study. All of these templates had a *tR2* terminator and T5 N25 or N25$_{antiDSR}$ promoter. They varied in the number of copies of a 60 bp DNA fragment, which contained +13 to +66 of the transcribed region of N25-100-tr2.

These pARn25-100-tr2 (pARn25$_{antiDSR}$-100-tr2) derivates plasmids were prepared by digesting pARn25-100-tr2(pARn25$_{antiDSR}$-100-tr2) with Nru I restriction enzyme, and inserting varying numbers of a 60 bp DNA fragment. The 60 bp DNA fragment was obtained from the PCR product of the N25-100-tr2 (pARn25$_{antiDSR}$-100-tr2) construct, which was treated with EcoRV at both ends (Figure 5-1). The two DNA primers – pARn25for and pARn25rev - were used in the PCR reaction. The 60 bp piece of DNA was ligated into pARn25-100-tr2

($pARn25_{antiDSR}$-100-tr2) in a molar ratio of 150:1, resulting in multiple 60 bp inserts. The plasmid were transformed into Stbl2 competent cells, and plated on ampicillin. The clones were selected and sequenced to determine the exact number of inserts.

10.2 The construction of DNA fragment for nanomanipulation

In order to construct the DNA fragment for the nanomanipulation experiment, the segments in (Figure 4-1) were obtained from the PCR products of the pARn25-100-tr2 ($pARn25_{antiDSR}$-100-tr2) plasmid or its multiple 60 bp insert derivatives add-on primers – MluIPROC4050 and NotPROC2050. The length of the resulting DNA construct was around 2kb.

After treating the 2kb piece of DNA with the restriction enzymes - MluI and NotI, the DNA fragment was ligated with 1kb biotin labeled DNA on one extremity and 1kb digoxigenin labeled DNA on the other extremity. The biotin-labeled DNA fragments were prepared via PCR amplification of pARn25-100-tr2 with the primers MluI RPOC4050 and RPOC 3140, along with Biotin-16-dUTP, and were subsequently digested with MluI. Similarly, the digoxigenin-labeled DNA was

prepared by amplification of pARn25-100-tr2 with Not PROC50, RPOC820 and Digoxigenin-11-dUTP, and was digested with NotI.

The sequences of the primers used are below:

pARn25for:

5'-GAGAGAGATATCGAGGAGTTTAAATATGGCTGGTTCTCGCG-3'

pARn25rev:

5'-GAGAGAGATATCGATTGGGATGGCTATTCGGAATTC-3'

MluI PROC4050:

5'-GAGAGAACGCGTGACCTTCTGGATCTCGTCCACCAGG-3'

Not PROC2050:

5'-AGAGAGCGGCCGCGAGAAGATCCGCTCCTGGAGCTACG. -3'

RPOC50:

5'-GAGAGAGCGGCCGCGAGAAGATCCGCTCCTGGAGCTACG-3'

RPOC820:

5'-TCCTGGCGCAGGTAGATGAG-3'

PROC3140:

5'-CTGATGCAAAAGCCCTCGGG-3'

10.3 Preparation of the Capillary

A square glass capillary, (1×1×50 mm) was immersed in nitric acid for 4 hours and rinsed with an excess amount of water. After drying, the inside of the capillary was coated with 0.1% polystyrene (w/v) in toluene, and dried with nitrogen. The dried polystyrene coated surface of the capillary was injected with 100 μl PBS containing 0.1 mg/ml polyclonal anti-digoxigenin. The capillary was incubated with the polyclonal anti-digoxigenin solution for 24 hours at 37 ℃. A blocking solution containing 10 mg/ml BSA and 3.3 mg/ml of polyglutamic acid, 25 mM Hepes NaOH, pH 7.9, 75mM NaCl, 10 mM $MgCl_2$,0.1% Tween 20, and 10 mM 2-mercaptoethanol was injected into the capillary and incubated for another 24 hours. The capillary with blocking solution was washed with 1×PBS. Polystyrene beads was injected into the capillary and allowed to settle. After sedimentation of these beads, the capillary was washed by SB buffer until nonbinding beads were removed. A 15 μl aliquot of DNA-anchored beads was injected into the capillary. The DNA anchored beads attached to the capillary surface through interaction of digoxigenin and anti-digoxigenin. After washing the capillary surface with a sufficient amount of SB buffer, most of the nonspecifically attached beads were removed.

10.4 Calibration and Data Acquisition

The magnets were loaded to a z axis position where the stretching force did not denature the DNA. The magnets were rotated causing a known number of superhelical turns to be introduced into the DNA molecule, with which we were interested. In order to determine the position of the beads, a calibration was made for the position of an object, as described in Figure 3-3. This was used to determine the l_{obs}. For each particular superhelical turn, the l_{obs} was collected from averaging every l_{obs} over a period of time. The length of acquisition time depended on the signal to noise ratio. With various superhelical turns, the curve of the superhelical turn vs. the l_{obs} of the DNA molecule was obtained (Figure 2-9). Superhelical turns were introduced into the DNA and made the corresponding $\sigma = 0$. At this supercoil-free state, by adjusting the z positions of the magnets, the fluctuations of the magnetic bead and l_{obs} were determined. The magnitude of the stretching force and the persistent length of the DNA were extracted from the fluctuation and l_{obs} measurement.

The DNA molecule was rotated to a region of constant torque (usually superhelical density=±0.021), and the z position of magnets was adjusted to a position

at which the applied magnetic force was around 0.3±0.03 pN. The capillary was washed with a solution containing: 25 mM Hepes NaOH, pH 7.9, 150 mM (75mM) NaCl, 10 mM $MgCl_2$,0.1% Tween 20, 0.1 mg/ml BSA and 10 mM 2-mercaptoethano. In the buffer, a curve for superhelical turns vs. l_{obs} was taken, which was used for the determination of the constant torque region of the DNA molecule. The sample was injected into the capillary, and data acquisition was begun.

10.5 Transcription Assay

The transcription assay uses 25 mM Hepes NaOH, pH 7.9, 75mM NaCl, 10 mM $MgCl_2$,0.1% Tween 20, 0.1 mg/ml BSA, 10 mM 2-mercaptoethanol and contains 10~1000 pM RNAP, 1μM GreB and 15nM sigma 70 with various concentration of NTPs. The transcription assay was injected into the capillary. The concentrations of RNAP used in the experiments were determined by the frequency of transcription events as the sample was injected.

In order to study the role of GreB, in some control experiments we did not add GreB into the sample solution. The role of sigma subunit here was to prevent the sigma 70 of the RNAP Holoenzyme from dissociating.

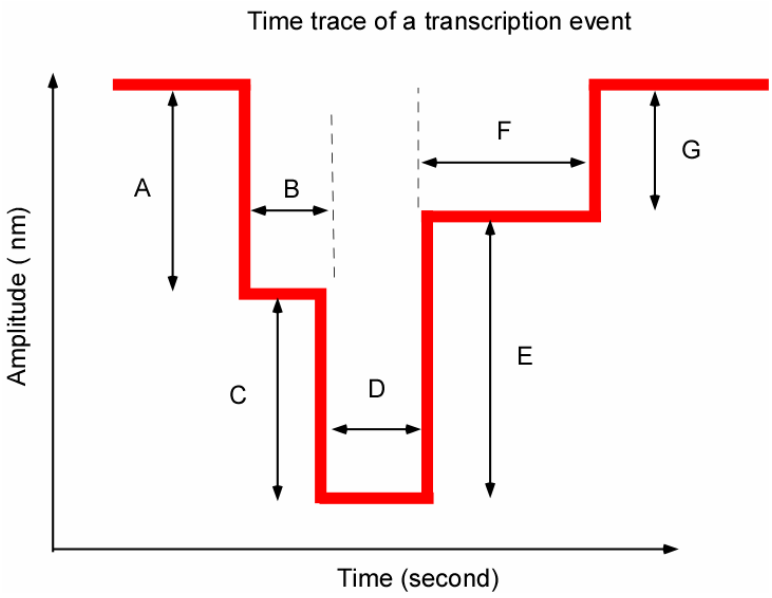
10.6 Buffer used in the experiments

PBS buffer:
Phosphate-buffered saline

SB buffer:
10 mM potassium–phosphate buffer, pH 8,
0.1 mg/ml BSA,
0.1% Tween 20

Appendix:

Appendix A: N25-160-tr2
Appendix B: N25-220-tr2
Appendix C: N25-280-tr2
Appendix D: N25-340-tr2
Appendix E: N25-100-tr2 without GreB
Appendix F: $N25_{antiDSR}$-160-tr2



The time trace of a complete transcription event can be observed in (Figure 6-2)

Appendix A: N25-160-tr2 time trace data for a complete transcription event

N25-160-tr2

date	August 17/2006							
		A	B	C	D	E	F	G
file name	800mpd							
time of event	131	78	18	100	9	118	21	55
	216			129	10	72	14	59
	497	96	5	64	2	37	11	62

	551	110	21	44	5	50	20	59
	668			116	20	42	46	70
	748	103	8	51	3	70	70	79
	1239			146	4	60	31	85
date	August 18/2006							
	100pmd/mol1							
	1866	107	8	60	6	109	8	51
	100pmd							
	59			146	2	88	33	65
	138			93	12	36	16	51
date	August 21/2006							
	250pmd							
	103	68	10	74	12	100	10	62
	tracd000							
	1265			116	9	55	10	72
	1497	61	2	106	4	116	12	86
	1614			208	3	147	16	64
	1697	72	10	87	13	94	21	62
	1776			140	3	84	19	53
	1974	103	4	29	4	65	14	50
	2019			155	4	110	8	74
	2045	74	17	76	13	94	13	81
	2159			89	2	30	16	50
	2346			143	6	80	12	71
	2495			199	3	129	15	49
	2542	113	8	79	3	117	10	56
	2756			128	5	32	14	62
	4026	74	9	48	9	73	21	44
	tracd001d							
	1286			90	6	34	14	70
	3011			148	15	93	11	67
date	August 22/2006							
	100pmd/mol1							
	957	106	20	58	5	102	28	51
	1536			102	37	34	61	68
	50pmd/mol2							
	1667			98	5	58	23	36

	100pmd							
	750			97	6	61	25	35
	825	44	10	37	11	48	14	30
	1905			130	13	89	10	32
date	August 23/2006							
	100pmd2d							
	857			141	5	81	27	60
	1117			146	4	88	8	59
	100pm3d							
	376			134	1	98	15	42
	431			87	12	48	13	46
	680			98	13	43	20	40
	875	52	2	108	4	57	14	44
	948			87	6	39	14	49
	992	69	4	65	3	46	9	64
	1037	74	5	136	1	59	2	64
	1133			108	13	51	34	62
	1303			94	1	37	1	39
	100pm4d							
	193			121	5	60	17	78
	312	90	14	72	3	83	34	73
	577	87	21	160	3	87	41	57
	798			115	5	53	20	58
	848			92	11	45	18	45
	1122			86	8	40	16	50
	100pm5d							
	101			85	19	29	17	48
	tracd001							
	225			86	12	31	22	50
	460			88	20	28	28	55
	713			161	7	94	8	68
	tracd002							
	803			87	52	37	17	46
	1661			117	15	51	14	58
	5852			85	16	34	12	53
date	August 24/2006							
	75pm/mol1							

	1790	111	10	94	3	30	13	78
	1883	94	4	39	4	26	7	67
	75pmd01							
	1145			111	6	55	14	76
	1269			103	11	27	31	69
	tracd01							
	15			115	9	48	19	71
	149			101	13	31	18	83
	339			117	15	71	7	48
	1963			114	12	43	17	68
	5465			109	14	83	20	51
	7056			127	4	58	10	58
	7330	91	10	47	5	88	15	58
	7677			165	6	85	29	66
	8097			159	7	90	15	80
	tracd02							
	3916			139	11	53	15	48
	6926			102	3	53	15	52
	tracd03							
	5746			71	12	30	16	44
date	August 28/2006							
	75pmd							
	177	115	11	57	4	97	13	64
	324			105	11	44	9	66
	1970			116	9	62	18	61
	2149	86	10	33	5	56	13	68
	75pmd01							
	347			186	2	120	35	68
	607			132	5	67	11	71
	787			115	5	45	28	72
	846			182	4	112	32	66
	75pmd03							
	221			126	4	52	13	80
	258			171	3	115	14	91
	378			150	2	85	8	54
	596			160	2	102	15	52
	797	80	3	51	7	58	16	73

	855	134	3	66	3	102	9	71
	903	108	4	82	3	124	14	54
	75pmd04							
	427	91	4	17	4	74	13	43
	456			96	3	47	13	60
	505	49	5	41	5	27	19	23
	688			73	6	34	63	39
	771			110	4	67	12	26
	810			66	8	45	13	16
	992			69	5	38	12	33
	1107			87	3	80	19	16
	1150	58	9	43	3	33	15	23
	1647	50	6	70	3	87	14	27
	1691	38	5	80	10	90	13	31
	2252			74	3	33	16	42
	2846			114	4	34	16	19
	tracd000							
	108			110	3	74	11	45
	127			92	4	34	13	81
date	August 30/2006							
	20pmd							
	1071			116	5	54	23	64
	50pmd							
	265			116	12	40	39	66
	510	96	18	48	2	75	13	65
	558	119	4	67	1	113	15	71
	626			174	2	107	25	64
	1446	91	5	113	5	135	29	64
	2026			125	5	55	17	48
	50pmd1							
	158	101	5	153	3	57	21	61
	814			128	3	68	23	53
	1132	141	6	86	2	164	19	77
	1162			164	2	88	17	62
	1193	68	17	72	3	120	21	40
	1765			130	4	68	20	68
	1818			137	6	71	14	61

	50pmd02							
	131			119	10	53	22	57
	369			130	6	55	11	65
	413	82	33	115	4	135	20	59
	758			178	7	101	16	70
	804			92	13	30	20	72
	2048			107	22	37	21	61
	2137			115	5	56	15	53
	80pmd							
	1005	113	21	20	18	65	15	57
	1177			159	3	89	18	58
	1227			100	16	42	24	63
	1813			114	4	64	14	56
	tracd0							
	119			130	8	87	14	42
	510			111	21	55	27	58
	577	104	5	57	13	115	18	58
	757			104	5	35	12	52
date	August 31/2006							
	50pmd							
	435			120	3	52	44	48
	50pm1d							
	59			120	7	73	24	47
	113			90	4	45	26	45
	333			104	13	52	22	48
	476			113	14	54	40	47
	907	86	5	52	6	84	18	52
	1032			124	8	82	31	48
	1154			104	2	45	23	52
	50pmd/mol2							
	29			102	24	48	19	57
	866			76	19	23	30	55
	1158			120	2	56	16	54
	1356			84	21	33	21	58
	1788			102	10	39	17	63
	2022			118	14	55	10	57
	tracd000							

	627			109	11	56	16	62
	679			104	17	42	26	61
	3512			130	7	69	16	68
	3760			138	12	81	18	64
	4586			152	7	103	19	51
	4671			121	6	57	13	63
	5047			129	5	66	18	57
	5103			95	7	43	13	53
	5218			105	5	37	13	64
	5490			136	5	68	14	65
	5668			143	9	80	20	66
	5972			147	5	85	15	65
	6780			113	6	56	15	55
	7105			96	7	36	17	61
	7504			133	11	46	16	81
	7563			115	9	65	17	56
	7654			70	5	18	18	58
	7960			93	14	39	24	65
	8278			97	15	29	14	63
	8572			130	25	70	15	56
	tracd01							
	62			90	7	62	14	40
	1420			151	12	81	16	65
	1878	84	14	41	5	69	15	58
	2944			153	15	96	16	55
	3485			139	6	78	27	68
	3966			121	8	65	14	58
	4101			111	32	42	15	63
	4653			101	11	46	23	60
	5219	57	28	57	9	55	15	72
	5378			148	19	84	17	53
	6667			97	8	37	17	57
	6936			128	5	60	30	73
	7026			85	124	24	87	57
	7932			114	10	45	12	70
	8691			148	6	90	18	51
	tracd02							

	1552			149	16	97	19	49
	1645			162	12	103	21	53
	7052			91	9	31	25	52
	7501			107	5	61	23	53
	7825			174	5	107	20	67
	8631			118	5	62	20	57
	tracd03							
	479			105	8	46	19	59
	1170			132	7	86	17	52
	4509			102	3	40	19	64
	4607	67	15	197	88	138	148	58
	5401	77	11	77	3	105	23	52
	5537			101	3	39	16	57
	6138			92	16	40	15	45
	6208			94	17	48	16	51
	6324			163	5	116	18	50
	6503			142	7	82	20	56
	6650	62	4	36	52	64	24	63
	7785			100	3	37	14	60
	7940			122	3	51	29	66
	8324			87	24	34	21	48
	tracd04							
	216			108	5	49	21	55
	772			138	6	72	23	68
	964			124	5	53	23	69
	1211			94	18	23	15	59
	1436			107	29	47	14	59
	3591			107	8	46	18	62
	4036			112	17	38	9	77
	4636			157	16	95	19	60
	5097			130	4	71	14	68
	5216	84	3	54	4	80	31	56
date	August 31/2006							
	tracd004							
	5893	69	2	45	4	58	223	59
	8046			105	7	52	19	51
	8230			151	2	87	21	65

	8504			152	2	99	19	64
	8637	84	2	33	9	66	23	67
	tracd005							
	2768			100	19	54	16	48
	2922	54	10	91	7	88	17	65
	2961			153	6	82	21	70
	3396	108	8	79	6	124	26	67
	3799			136	5	71	12	68
	4735			98	8	44	25	60
	4876			110	40	48	26	64
	5327	78	20	44	236	66	245	58
	5957			127	10	59	18	67
	6143			191	15	126	28	66
	6217			197	5	129	18	63

Appendix B: N25-220-tr2 time trace data for a complete transcription event

N25-220-tr2

date	July 11/2006							
		A	B	C	D	E	F	G
file name	elonga.gr							
time of event	733	74	5	65	19	91	24	58
	891			151	8	78	24	136
	1271			172	7	115	26	62
	trac0a.gr							
	243			106	18	51	23	52
	296			92	13	46	33	69
	550			105	34	32	47	72
	1161			116	10	42	17	59
	1381			166	7	98	13	65
	1630			125	8	45	24	70

	7339			123	5	62	27	58
date	July 12/2006							
	100pmd							
	50			116	3	56	26	50
	429			132	10	59	25	74
	50pma.gr							
	547			107	12	43	23	59
	635			111	11	52	31	62
	826			163	11	100	28	60
	trac0a.gr							
	106			117	12	49	45	62
	trac2a.gr							
	169			93	27	39	32	52
	373			90	20	36	26	58
	591	68	16	87	25	105	19	45
	952			128	10	69	19	49
	1158			97	19	46	24	47
	1397			119	11	56	18	56
	1545			74	30	24	16	59
	1911	63	23	67	13	55	28	28
	2374			112	9	62	12	46
	2494			130	11	77	32	57
	3436			98	11	57	35	50
	3721			100	12	35	35	53
	4017			91	22	52	22	40
	4100			85	128	48	42	60
	4814			58	60	21	22	49
	4986			85	44	36	29	76
	tracd02.gr							
	5342			130	11	84	22	52
	5420			125	12	76	22	48
	5832			88	39	63	106	57
	6766			80	11	18	22	59
	7616			108	34	46	29	58
	8060			144	12	77	16	62
	tracd03							
	857	69	12	79	3	95	31	56

	1960			100	7	50	28	42
	3063			89	10	39	31	49
	5613			85	16	25	50	68
	6551			128	3	68	19	63
	7912			111	7	57	25	58
date	July 13/2006							
	100pmd							
	457			114	15	53	28	62
	1310			107	10	50	24	54
	1784			102	5	57	20	59
	200pmd							
	591			96	10	36	28	67
	226			107	10	44	18	56
	274			165	3	98	26	48
	tracd001							
	2706			136	7	86	14	39
	tracd001							
	2770			93	9	32	25	38
	2867			99	5	65	23	35
	3688			67	19	25	21	43
	3761			98	23	48	22	40
	3835			72	12	58	22	27
	4084			141	6	63	24	75
	4304	37	8	50	18	38	23	60
	4704			106	8	45	21	66
	4771			100	8	61	25	38
date	July 17/2006							
	100pmd							
	160			125	6	69	22	45
	277			114	26	48	11	60
	327	98	2	38	3	67	22	64
	393			97	6	36	25	64
	501			143	3	75	38	54
	1703			105	76	26	67	73

Appendix C: N25-280-tr2 time trace data for a complete transcription event

N25-280-tr2

date	July/25/2006							
		A	B	C	D	E	F	G
file name	200pmd							
time of the event	1137	52	15	62	18	33	18	80
	tracd010							
	45			103	5	65	26	66
	456			117	10	49	18	54
	673			127	9	83	11	36
	702			96	2	26	36	43
	758			103	14	53	30	28
	817			92	4	57	18	66
	1449			120	6	81	38	39
	1681			89	19	52	29	37
	4764			103	6	69	22	34
	4958	83	5	86	3	96	21	69
	5257			75	25	35	38	40
	5332			95	7	35	22	54
	5411	68	7	69	3	112	16	63
	5561			97	20	44	31	54
	6052			88	12	27	19	57
	6455			96	20	49	27	35
	6998	67	10	85	7	76	35	66
	tracd010							
	7567			81	23	32	28	51
	8623			77	6	24	28	45
	tracd011							
	837			98	6	50	32	41
	1719			79	36	34	54	30
	1965			103	16	68	33	42
	2524			112	20	59	57	58

	3052			98	7	57	27	48
	3692			147	7	101	44	32
	tracd013							
	437			40	38	24	92	23
	1928			92	21	32	37	50
	tracd014							
	1453			103	19	43	20	55
	tracd015							
	3011	41	25	24	9	18	35	38
	150pmd							
	173	110	4	91	2	132	40	56
	415			99	6	32	28	63
	558			100	39	32	60	48
	1407			139	2	75	34	59
date	July/27/2006							
	50pmd							
	1021			123	12	82	29	32
	50pmd1							
	264	93	3	25	3	47	52	63
	tracd00							
	246			136	5	60	38	60
	803			147	5	89	16	58
	1841			148	9	67	85	58
	2131			105	10	42	32	62
	2792			137	13	65	40	79
	2872			105	11	26	29	69
	3491			105	10	48	14	54
	3612			150	7	89	33	57
	4156			139	12	69	21	68
	5139			107	6	47	23	62
	5869			115	20	45	31	54
	5929			156	4	100	27	62
	5966	66	6	35	4	32	20	67
	6138			93	32	41	34	59
	6228			159	3	97	19	52
	6378			99	3	97	19	52
	6423			142	4	122	38	64

	6548			134	2	69	22	67
	tracd00							
	6768			128	6	61	26	58
	8573	81	26	28	12	50	31	61
	8672			151	6	82	40	64
	tracd01							
	2546	87	24	49	6	52	31	78
	2625			114	5	38	42	67
	3165			131	20	80	24	57
	3637			120	9	64	32	67
	3990			116	6	55	32	63
	4213			107	26	48	26	54
	4339			129	6	69	35	66
	4568			80	4	18	26	60
	4689			137	4	71	25	69
	4777			115	8	46	33	68
	4902			96	5	36	28	60
	5041			129	7	70	24	55
	5101			124	5	63	28	99
	5419			144	8	85	20	55
	tracd01							
	5500			154	5	85	27	52
	5606			98	12	40	35	60
	5739			118	5	67	16	54
	6128			120	5	63	27	54
	6244			105	7	44	21	55
	8179			122	10	62	22	56
	8507			116	12	61	32	56
	8590			122	14	53	26	71
	tracd02							
	942			102	5	34	16	67
	1358			112	4	41	23	69
	1536			124	7	52	26	68
	2508			121	22	49	25	61
	3027			112	14	42	21	73
	7187			111	10	50	16	64
	7422	114	10	46	5	106	24	57

	8032			116	14	51	20	65
	8238			126	7	59	28	67
	tracd03							
	4715			107	9	46	20	66
	7040			106	14	33	25	72
	tracd04							
	1061			114	7	57	36	64
	4975			118	6	60	20	48
	5096			111	137	46	342	60
	7368			118	16	67	17	55
	tracd05							
	2967			112	14	40	23	67
	3807			117	6	46	19	64
	4891			114	10	42	32	68

Appendix D: N25-340-tr2 time trace data for a complete transcription event

N25-340-tr2

date	July 19/2006							
		A	B	C	D	E	F	G
file name	150pmd							
time of event	152			108	9	53	36	65
	793			106	11	47	40	56
	862			107	32	40	37	59
	1219			137	6	66	4	66
	1811			116	5	48	37	60
	200pmd							
	918			108	12	45	43	57
	1250			143	7	87	27	52

	1505			111	21	45	28	49
	1770			111	14	50	81	57
	1899			102	14	40	39	54
	200pmd1							
	7			81	18	40	48	38
	300pmd							
	146			103	9	48	51	51
	484	53	23	52	18	74	21	26
	698	112	20	40	38	44	26	106
	902			106	27	45	38	65
	1092			124	63	58	11	59
	150pmd1							
	452			93	11	39	33	64
	1101			92	13	35	28	58
	tracd00							
	849			75	21	29	39	45
	938			74	6	37	25	46
	7516			67	99	22	57	51
	8033			67	11	35	19	24
	tracd02							
	960			92	5	39	38	56
	tracd03							
	244	37	10	23	7	20	9	46
	200pmd2							
	323	52	2	71	9	73	34	57
	826	107	2	49	1	87	38	65
	1227			104	13	36	30	61
	1405			137	4	71	34	62
	1586			69	9	54	4	21
	1917			152	2	97	48	62
date	July 20/2006							
	100pmd							
	452			86	20	27	50	55
	681			143	2	81	37	61
	761			101	10	41	50	59
	100pmd1							
	1376	87	3	54	2	84	60	55

	100pmd2							
	35			146	3	98	40	50
	159	98	5	30	1	74	39	66
	361			103	28	42	57	62
	509			133	2	84	43	61
	2511			105	47	70	40	50
	2865	92	10	38	3	85	21	44
	2920	84	9	37	5	81	59	46
	3009	93	2	42	3	38	7	39
	3024			62	44	43	382	60
	3536	78	1	62	6	66	46	71
	3612			96	18	33	43	60
	3735			84	32	28	47	43
	3899			102	5	62	45	48
	3958			93	12	30	16	43
	3990	51	7	49	31	42	112	78
	100pmd3							
	336			93	8	61	47	40
	898			97	12	24	58	57
	1073			90	14	38	67	63
	1523			71	19	31	40	46
	1720			91	8	60	28	46
	2096	73	4	29	12	45	35	46
	2348			89	11	42	41	51
	2865	70	4	47	5	64	32	56
	3037			97	34	42	47	55
	1200pmd							
	900			110	18	37	39	63
	1020			117	21	46	38	68
	1118			107	8	49	37	63
	50pmd							
	110			132	11	80	42	58
	tracd0							
	17			140	11	86	37	55
	758			133	6	73	40	58
	tracd3							
	874			90	48	47	190	44

	2143			120	5	64	40	52
	2209			86	34	32	37	50
	2290			109	27	60	21	66
	5588			90	28	32	48	55
	5855			112	6	50	50	58
	7219			109	14	60	20	55
	8017			123	15	60	34	69
	tracd04							
	2058			110	19	47	40	60
	2211			129	9	70	46	64
	2435			114	38	50	29	58
	2692			102	5	45	45	54
	4280			163	21	97	30	60
	4871			128	4	82	39	59
	5789			108	9	64	23	62
	7459			118	5	34	30	71
	tracd06							
	842			113	9	56	42	62
	6352			88	11	31	32	56
	6631			119	32	57	33	61
	7170			120	5	59	36	63
date	July 21/2006							
	200pmd							
	296			129	7	77	21	61
	543			132	3	75	27	53
	651			101	31	40	28	66
	843			113	9	47	25	66
	1136			96	8	47	21	61
	150			117	6	66	13	60
date	December1/2006							
		A	B	C	D	E	F	G
file name	100pmd							
time of event	832			96	23	43	31	56
	1145			131	4	74	30	55
	1239			145	5	86	43	57
	1289			155	2	102	29	50
	1423			111	2	67	25	54

	100pm1d							
	14			141	25	55	44	85
	tracd000							
	211	75	36	16	43	36	53	54
	7564	93	13	47	5	87	26	58
	7766			102	15	37	20	99
	8448			105	31	51	31	50
	tracd004							
	734	63	15	68	7	100	23	53
	861			120	5	56	22	60
	2179			112	15	56	25	54
	2906			120	23	54	36	69
	6606			105	12	46	192	94
	7825			112	6	55	42	54
	tracd005							
	3587			93	35	28	35	61
	4608	103	9	49	16	87	42	67
	6424			113	21	53	60	64
	2958			95	79	29	22	61

Appendix E: N25-100-tr2 without GreB time trace data for a complete transcription event

N25-100-tr2 without GreB

date	July/09//2006							
file name	tracd000	A	B	C	D	E	F	G
time of the event	290			90	6	54	6	30
	370			104	15	38	9	97
	720			146	11	94	16	41
	938			118	9	59	16	64
	2040			75	97	55	11	30

	2984			97	48	65	91	40
	3174			172	22	112	10	72
	5215			90	6	34	12	52
	8635			119	8	91	11	38
	tracd001							
	4578			142	7	41	10	102
	5785			90	19	36	13	53
	7088			88	164	28	12	57
	tracd002							
	442			107	26	43	147	56
	797			98	13	48	9	49
	1145			117	6	56	17	67
	tracd002							
	5107			90	7	32	845	51
	6940			110	55	45	58	62

Appendix F: $N25_{antiDSR}$-160-tr2 time trace data for a complete transcription event

$N25_{antiDSR}$-160-tr2

date	April /27/2005							
file name	tracd025	A	B	C	D	E	F	G
time of event	2856			154	135	91	15	75
	3270	143	3	20	89	79	21	82
	3876			160	45	97	11	71
	tracd026							
	206			165	21	91	17	71
	358	98	5	62	97	90	12	66
	570			164	38	101	13	79
	1049			161	83	160	13	64
	1395			164	113	88	21	71
	2993	89	2	72	3	91	19	65

	3029	98	2	54	48	93	15	64
	trac27							
	186			168	90	95	11	67
	1774	107	3	57	43	95	271	65
	2211	122	2	43	19	96	22	71
	2356			162	88	70	15	64
	2483			162	226	97	26	66
	2810	100	4	63	92	87	13	79
	3567	126	1	43	84	92	295	70
	4030			166	106	95	120	73
	traca28							
	525			158	13	73	21	74
	682	124	3	43	13	101	9	60
	1531	115	10	49	106	98	8	65
	1906			167	273	86	31	67
	2538	131	9	38	23	91	138	68
	traca29							
	1991	121	5	40	218	98	95	65
	2941			144	49	84	16	68
	trac029							
	321			166	35	85	13	73
	1466	133	5	31	76	93	16	68
	1598	102	3	73	72	96	15	68
	1733			167	54	89	14	67
	1840			155	10	82	10	69
	1901			155	55	94	23	71
	2026	124	5	44	72	92	15	73
	2276	117	5	51	17	96	12	79
	2345	129	3	46	186	97	22	70
	2600	127	2	41	27	90	11	70
	2712	91	4	70	25	97	12	70
	traca30							
	2660	105	5	56	50	87	11	71
	traca31							
	206	110	1	57	36	86	12	69
	1619	107	3	55	46	48	14	67
	2471			165	11	99	11	69

	traca109							
	363			168	18	91	14	77
	600			172	12	92	9	79
	tracd201							
	1939			140	214	81	16	41
	4057			127	67	85	92	44
	trac202							
	280			138	8	90	22	50
	375			152	69	107	25	38
	2369			144	183	101	33	41
date	July /19/2005							
file name	tracd066.gr	A	B	C	D	E	F	G
time of event	150			138	423	67	4	68
	tracd089.gr							
	360			125	42	83	8	73
	707			143	31	78	13	65
	1416			148	10	75	12	77
	1617			136	174	85	8	62
	1972			117	32	63	6	81
date	July /22/2005							
	rnap0d.gr							
	281			140	534	94	18	50
	901			135	262	80	19	49
	1185			137	112	120	8	89
	1308			145	262	73	18	68
	1980			132	15	56	14	56
	2066	75	1	70	33	83	9	47
	3764			149	143	83	11	82
	rnap1d.gr							
	42			173	426	44	29	64
	537			110	179	59	18	54
	1209			97	403	62	11	36
	trac91d.gr							
	2987			134	813	45	22	88
	trac92d.gr							
	333			125	16	64	17	45
	1272			137	165	68	28	58

	1895	21	2	113	153	74	21	65
	trac93d.gr							
	3105			138	187	80	27	64
	3424			150	491	80	19	66
	trac94d.gr							
	2359			153	291	83	24	63
	2890	101	13	33	290	38	6	56
	3648	46	5	97	223	62	12	83
	trac95d.gr							
	1040	96	6	33	269	80	14	50
	2191			131	147	72	24	56
	2191			122	423	87	22	45
	tracd121.gr							
	73			126	56	71	25	52
	410			118	173	83	18	44
	627			141	142	81	39	61
	828			114	26	58	211	51
	1123			138	220	83	13	55
	1522			117	227	74	29	48
	3504			103	371	58	83	48
	tracd120.gr							
	633			96	31	42	35	49
	2636			119	120	77	19	52
	2797			110	65	56	31	48
	3181			124	263	86	29	51
	3528			139	651	71	23	59
	rnap2d.gr							
	326			118	21	60	117	51
	1049			109	88	48	22	74
	3224			104	186	61	13	52
	3563			109	507	69	29	42

	rnap3d.gr							
	565			117	536	46	33	46
	1217	65	2	66	99	60	22	61
	trac90d.gr							
	2311			147	16	93	13	50
	2491	66	5	54	1294	60	12	70
date	August /16/2005							
	tracd001.gr							
	86			108	191	44	12	64
	tracd000.gr							
	3392			101	17	39	14	62
	4061			105	20	53	15	52
date	November/16/2006							
file name	400pmd	A	B	C	D	E	F	G
time of event	200			100	19	29	21	76
	387	86	4	53	1	84	19	55
	650			108	11	51	26	57
	731	87	4	94	1	123	23	83
date	November/20/2006							
	200pmd							
	170			166	14	54	31	91
	219			162	7	123	22	63
	279			166	5	104	15	39
	797			126	52	70	19	54
	400pmd							
	969			108	79	50	26	52
	1965			105	166	41	14	58
	400pmd1							
	331			118	170	53	11	63
	562			111	101	49	18	53
	1868			106	38	50	16	54
	1954			112	77	30	19	71
	2070			118	235	38	18	67
	400pmd2							
	2901			105	27	55	9	51
	3011			112	58	44	11	65
	3268			87	175	73	18	61

	3600			139	16	47	17	20
	3920			113	90	55	9	57
	400pmd3							
	664			127	40	66	7	55
	1117			127	9	60	12	60
	400pmd4							
	104			130	16	67	20	62
	273			113	95	41	14	73
	582			131	47	51	31	66
	1068			123	39	65	14	58
	1908			118	21	65	22	58
date	November/27/2006							
	485			112	750	53	14	59

Curriculum Vitae
Chen-Yu Liu

1994.10~1998.6
Undergraduate student, Chemistry Department, National Tsing Hua University

1998.6~2000.7
Graduate student, Chemistry Department, National Taiwan University

2000.10~2000.12
Research assistant, The Institute of Biological Chemistry, Academia Sinica

2000.12~2000.1
Military service

2001.2~2001.6
Research assistant, The Institute of Biological Chemistry, Academia Sinica

2001.8~2002.8
Teaching assistant, Chemistry Department Taiwan University

2002.10~2003.6
Research assistant, The Institute of Chemistry, Academia Sinica

Graduate student, Biophysics and Computational Biology , Rutgers University, New Brunswick 2003.8~2007.7
Advisor: Richard H. Ebright

Publication:

Andrey Revyakin,' Chenyu Liu, Richard H. Ebright, and Terence R. Strick, ' ABORTIVE INITIATION AND PRODUCTIVE INITIATION BY RNA POLYMERASE INVOLVE DNA SCRUNCHING, *Science*, **2006**, 314,1033