MODULATION OF TRANSCRIPTION ELONGATION VIA THE MAIN CHANNEL IN ESCHERICHIA COLI RNA POLYMERASE

Scott Robert Kennedy

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry

Chapel Hill
2007

Approved by:

Advisor: Dr. Dorothy A. Erie

Dr. Matt Redinbo

Dr. Tom Kunkel

Dr. Linda Spremulli

Dr. Lee Pedersen

ABSTRACT

SCOTT R. KENNEDY: Modulation of Transcription Elongation via the Main Channel of *Escherichia coli* RNA Polymerase
(Under the direction of Dr. Dorothy A. Erie)

Conformational changes in RNA polymerase play an important role in the regulation of transcription elongation. Previous work has demonstrated that RNA polymerase can exist in an activated state and an unactivated state which synthesizes RNA fast or slow, respectively. Additionally, the distribution of elongation complexes between these two states has been shown to be regulated by an allosteric; putatively located in the main channel of the enzyme.

In this work, I determined that the proposed NTP binding site is located in the main channel of RNA polymerase and modulates elongation by using the i+2 nucleotide. Using site directed mutagenesis, I show that the allosteric site is composed of fork loop 2. Additionally, I clearly show that there are at least two routes for NTPs to enter the catalytic site and that one of these entry routes is via the main channel. The main channel has been previously proposed; however, up to this point, there has been little experimental evidence supporting this idea. The data from my experiments, combined with structural evidence, I have put forth a model in which the allosteric NTP causes a conformational shift in RNAP that leads to rapid translocation and fast synthesis (ie. activated state).

To further bolster the result that shows that fork loop 2 comprises an allosteric site, I also performed experiments on a mutant RNA polymerase in which a totally conserved

Walker B motif that is in close proximity to fork loop 2 was removed via mutagenesis. My experimental results indicated that the Walker mutant has a reduced affinity for the allosteric NTP. These results, combined with the previous work on fork loop 2 mutants, strongly indicates that the main channel is heavily utilized in transcription elongation. Further studies will be needed to elucidated the details the role the main channel plays in transcription elongation.

To Dad, for giving me my sense of curiosity and wonder about the world.

ACKNOWLEDGMENTS

I would like to, first, thank my advisor, Dorothy Erie. One could not ask for a better boss/advisor. She had patience and faith that things would work out well even when things were not working. She's one of my scientific heroes and I appreciate all that she has done for me.

I would like to also thank Lauryn Sass and Bob Brodnick. They were always willing to have interesting and thought provoking conversations. They also listened when things were rough and kept encouraging me.

Penny Faires also deserves a special "thank you". She was always supportive of me and helped me through the whole process of graduate school. I cannot thank her enough.

Lastly, I would like to thank my parents and my friends. They made sure I was aware that there was life outside of graduate school.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

3’	three prime end
5’	five prime end
~	approximately
≈	approximately equal to
≥	greater than or equal to
%	percent
A	adenosine monophosphate
A	alanine
a.a.	amino acid
ATP	adenosine-5’-triphosphate
α	alpha
β	beta
β’	beta prime
b	base
BSA	bovine serum albumin
C	cytidine monophosphate
Ci	Curie
CIP	calf intestinal phosphatase
cm	centimeter
CTP	cytidine-5’-triphosphate

Δ	deletion
Δ	delta
D	aspartic acid
D	Dalton
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
EC	elongation complex
EDTA	ethylenediamine tetra-acetic acid
F	phenylalanine
FRET	fluorescence energy resonance transfer
g	gram
G	glycine
G	guanosine monophosphate
GTP	guanosine-5'-triphophate
HEPES	*4*-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
his	histidine
hr	hour
IPTG	isopropyl β-D-1-thiogalactopyranoside
in vitro	inside a living system
k	kilo
K^+	potassium cation
λ	lamda

L	liter
μ	micro
m	milli
M	molar
Mg^{+2}	magnesium cation
mol	mole
n	nano
N	asparagine
N	N antitermination factor
nt	nucleotide
NTP	nucleotide triphosphate
ω	omega
OH	hydroxyl group
OD	optical density
OPC	open promoter complex
ORF	open reading frame
p	plasmid designation
^{32}P	radioactive phosphorous-32
P_R	“rightward” promoter
PAGE	poly-acrylamide gel electrophoresis
PCR	polymerase chain reaction
pDE13	plasmid with the DE13 sequence
Pol II	RNA polymerase II

Q	Q antitermination factor
R	arginine
RNA	ribonucleic acid
RNAP	ribonucleic acid polymerase
rpm	revolutions per minute
rpoA	gene encoding α-subunit
rpoB	gene encoding β-subunit
rpoC	gene encoding β'-subunit
rpoD	gene encoding σ-subunit
rpoZ	gene encoding ω-subunit
σ	sigma
S	Siemans
SDS	sodium dodecyl sulphate
SECIS	selenocysteine insertion sequence
Stl	streptolydigin
T	thymidine monophosphate
T_m	DNA melting temperature
T. aquaticus	*Thermus aquaticus*
T. thermophilus	*Thermus thermophilus*
U	uridine monophosphate
UTP	uridine-5'-triphosphate
v	volume
vs	versus

w	weight
wt	wild-type

CHAPTER 1:

TRANSCRIPTION ELONGATION

1.1 Introduction

Transcription is the first step in the process of gene expression and is characterized by the synthesis of RNA in a DNA template dependent manner. Transcription is the most regulated step in gene expression and must occur at a reasonable rate and high fidelity. The enzyme that carries out transcription is called RNA polymerase. All cellular life makes use of multi-subunit RNA polymerases to carry out a majority of RNA synthesis in the cell. In prokaryotes, the core enzyme has a molecular weight of ~450kD and is composed of five subunits: 2 α subunits, β, β', ω. The core enzyme is responsible for elongation and termination. In prokaryotes, the addition of subunit σ forms the holoenzyme and is required for initiation.

With the exception of mitochondrial and chloroplast RNA polymerases, transcription in all eukaryotes is carried out by three multisubunit RNA polymerases; RNAPI, RNAPII, and RNAPIII (Sweetser, Nonet *et al*. 1987). Each homolog is responsible for transcribing different types of genes; specifically, RNAPI synthesizes rRNA precursors, RNAPII synthesizes mRNA, and RNAPIII is responsible for synthesizing tRNA. Eukaryotic RNAPs share several conserved subunits with prokaryotic RNAP. Specifically, the two largest subunits of RNAPI, RNAPII, and RNAPIII have high sequence identity with the β and β' subunits of eubacteria

(Allison, Moyle *et al*. 1985; Biggs, Sesrles *et al*. 1985; Sweetser, Nonet *et al*. 1987).

The extensive conservation between all multisubunit RNAPs is probably due to the conservation of their basic function and internal mechanics. RNAPII shows the highest degree of sequence conservation with prokaryotic RNAP and shows the highest similarity in structure, biochemical function, and regulation (Allison, Moyle *et al*. 1985; Sweetser, Nonet *et al*. 1987). Since prokaryotic RNAP and RNAPII are homologous, studies on bacterial RNAP can lead to insights on transcription in higher organisms.

1.2 The Many Steps of Transcription

The process of transcription can be broken down into four main steps (Figure 1.1): 1) Open promoter complex formation, which is characterized by the binding of the RNAP holoenzyme to a promoter and the subsequent formation of a DNA bubble; 2) Initiation, characterized by the formation of a small 6-9 nucleotide (nt) transcript and subsequent release of the σ-subunit; 3) Elongation, which is the processive elongation of RNA by the addition of nucleotides into the growing RNA strand, and 4) Termination, which entails the release of both RNAP and the transcript from the DNA template.

1.2.1 Open Promoter Formation

During open promoter formation in prokaryotes, the RNAP holoenzyme (core+σ) recognizes a promoter sequence consisting of two conserved hexameric DNA sequences (-35 and -10 box). When promoter recognition and binding occurs, the double stranded DNA is melted at the promoter region leading to the formation of the transcription bubble. Upon transcription bubble formation, RNAP enters the initiation phase. The situation is

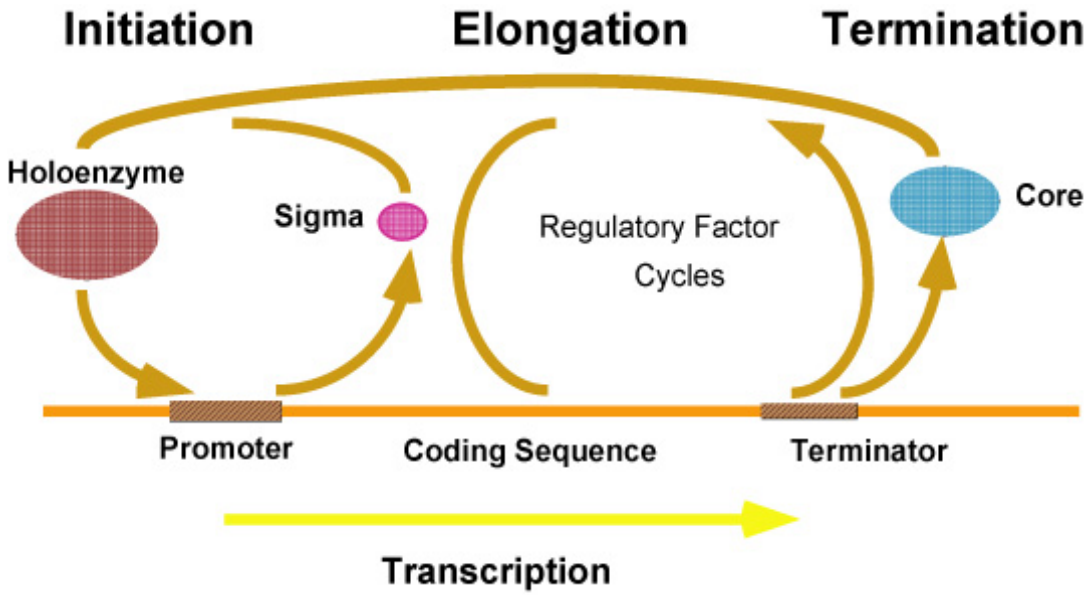


Figure 1.1 The Transcription Cycle. The RNAP holoenzyme is colored deep red, core RNAP is colored cyan, and the sigma subunit is colored pink. The holoenzyme is required to bind to the promoter (brown box). Sigma is released to form the core enzyme upon transition to the elongation phase.

significantly more complex in eukaryotes. Due to gene expression being significantly more regulated in eukaryotes, RNAPII is recruited and forms the open promoter complex by being phosphorylated and using a myriad of protein factors. Interestingly, eukaryotic RNAPs do not use a σ homolog for open promoter formation and initiation, but rather use several different proteins that act together as a σ ortholog (Thomas and Chiang 2006).

1.2.2 Transcription Initiation

Initiation is characterized by the synthesis of a short 6-9 nt transcript prior to the release of the σ-factor. During initiation, the synthesis of short transcript is believed to induce " DNA scrunching" by RNAP *(*Kapanidis, Margeat *et al. 2006;* Revyakin, Liu *et al*. 2006). In the scrunching model, RNAP draws the downstream DNA template into the enzyme, thereby inducing a DNA bulge in the single stranded DNA bubble. The stress induced by template scrunching is thought to cause release of the short transcript in a process known as abortive initiation (Straney and Crothers 1987). If release occurs, RNAP starts another round of synthesis. In productive initiation, meaning initiation leading to the transition to elongation, the physical stress of scrunching, combined with the interaction of the transcript with the σ-factor, leads to the release of σ from the core RNAP and the escape of RNAP from the promoter region (Kapanidis, Margeat *et al*. 2006).

1.2.3 Transcription Elongation

Elongation begins after a 10-12 nt transcript is formed and is characterized by the addition of NMPs into a nascent RNA chain by the RNAP ternary complex. The

elongation phase is processive, which means that RNAP must finish synthesizing the transcript before dissociating. Should the enzyme dissociate prior to completion of the transcript, it must release the incomplete transcript, rebind σ, and start again at open promoter formation. Because elongation is processive, it is incumbent for RNAP to be properly regulated. Elongation regulation is evident by the fact that RNAP can undergo several different conformational shifts which are governed by the interactions between the enzyme and extrinsic protein factors, DNA template sequences, and/or the transcript sequence. Each state is characterized by its own conformation and biochemical properties (Table 1) (Arndt and Chamberlin 1990; Erie, Hajiseyedjavadi *et al*. 1993; Mustaev, Kashlev *et al*. 1993; Yin, Artsimovitch *et al*. 1999; Davenport, Wuite *et al*. 2000; Foster, Holmes *et al*. 2001; Tolić-Nørrelykke, Engh *et al*. 2004).

1.2.3.1 The States of Transcription Elongation

The process of elongation can be broken down into three main pathways: synthesis pathways, regulatory pathways, and errant/rescue pathways (Figure 1.2). The main synthesis pathway consists of an activated state and an unactivated state. The activated state is characterized by rapid synthesis of the RNA chain and lowered fidelity. Conversely, the unactivated state, while competent to incorporate NTPs, does so at a much slower rate and with higher fidelity (Erie, Hajiseyedjavadi *et al*. 1993). The unactivated state is the primary state in which RNAP may enter regulatory or errant/rescue pathways. The regulatory and errant/rescue pathways are generally characterized by the cessation of NMP incorporation and stalling of elongation.

One of the primary regulatory states is pausing. Pausing is the temporary delay in

State	**Description**	**Reference**
Activated	Normal addition of NTPs to RNA; Rapid; Low Fidelity.	(Foster, Holmes *et al*. 2001; Holmes and Erie 2003)
Unactivated	Normal addition of NTPs to RNA; Slow; High Fidelity.	(Foster, Holmes *et al*. 2001; Holmes and Erie 2003)
Paused	Temporary cessation of synthesis; induced by DNA/RNA sequence.	
Stalled	Temporary cessation of synthesis due to lack of NTP substrate.	(Krummel and Chamberlin 1992)
Arrested	Unable to continue synthesize RNA; requires extrinsic protein factors to continue synthesis.	(Krummel and Chamberlin 1992)
Backtracked	3'-end of the RNA is translocated backwards relative to the catalytic site and extruded out of the enzyme.	(Surratt and Milan; Boukhov, Polyakov *et al*. 1992)
Hyper-translocated	forward translocation of RNAP relative to the 3'OH of the transcript.	*(Toulokhonov, Zhang et al*. 2007)
Cleavage	Hydrolysis of an internal phosphodiester bond to form a new 3'OH at the catalytic site.	(Surratt, Milan *et al*. 1991; Boukhov, Polyakov et al. 1992)

Table 1.1 Major Conformational States in Transcription Elongation

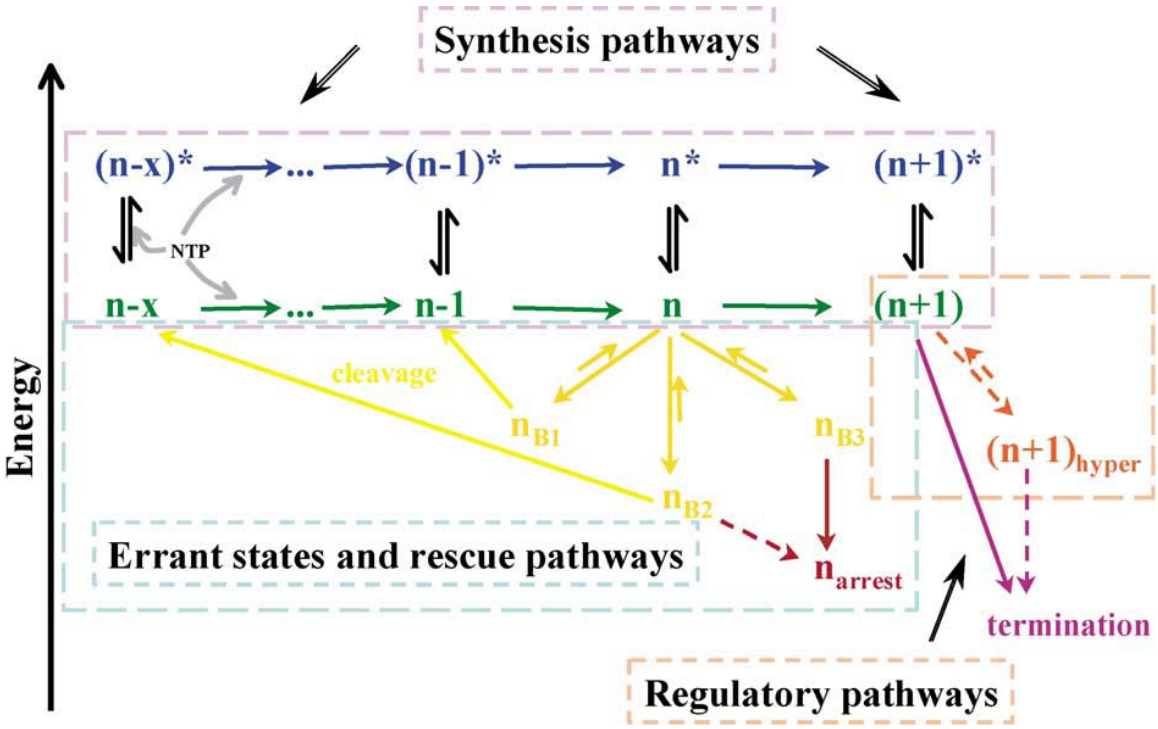


Figure 1.2 Catalytic and Regulatory States Used in Transcription Elongation.

Elongation normally occurs along either an activated path (n^*, blue), which is induced by NTP binding to an allosteric site, or an unactivated path (n, green). The unactivated path is sampled in the absence of bound NTPs. The enzyme can further decay into various backtracked states (n_{B1}, n_{B2}, n_{B3}, yellow). The backtracked states can be rescued back to the elongation state upon transcript cleavage; an alternative fate is that the back tracked states can further decay to a fully arrested state (n_{arrest}, red) which can only be recovered with the assistance of extrinsic protein factors. As an alternative to the decaying to the backtracked state, the unactivated state can undergo a transition to the termination phase (purple) of transcription or enter the hypertranslocated state ($(n+1)_{hyper}$, orange), which can then undergo a transition to termination.

RNA chain elongation due to the enzyme encountering certain DNA sequences. The purposes for pausing are manifold. Pausing has been shown to allow the synchronization of transcription and translation in prokaryotes, allow enough time for regulatory factors to interact with RNAP, and is a precursor to both transcriptional arrest (complete cessation of elongation) and transcriptional termination (described later) (Landick, 2006). Pausing also leads to other regulatory states, such as hyper-translocation, which is thought to be on path to termination. In certain DNA sequence contexts, RNAP pauses and the nascent RNA forms a stable hairpin structure which causes the 3'-end of the RNA to fray or hyper-translocate due to the forward translocation of RNAP relative to the 3'-OH of the transcript.

Both the synthesis and regulatory pathways can lead directly into the errant pathway in which RNAP enters conformations that are no longer competent for elongation without major changes to the elongation complex. A major errant state is the backtracked state. In the backtracked state, the 3'-end of the RNA is translocated backwards relative to the catalytic site and extruded out of the enzyme (Komissarova and Kashlev 1997; Komissarova and Kashlev 1997) (Figure 1.3). Under certain reaction conditions, such as NTP deprivation, RNAP can enter an arrest or 'dead end' state, where even high concentrations of NTPs are unable to rescue productive synthesis. Under arrest and backtracking conditions, RNAP must cleave the transcript in order to return to productive synthesis. The arrested/backtracked intermediate places an internal phosphodiester bond proximal to the catalytic Mg^{+2}. Bond cleavage can occur intrinsically (Surratt, Milan *et al*. 1991; Izban and Luse 1992) or be stimulated by GreA or GreB (Komissarova and Kashlev 1997). Upon transcript cleavage, the 3' fragment is

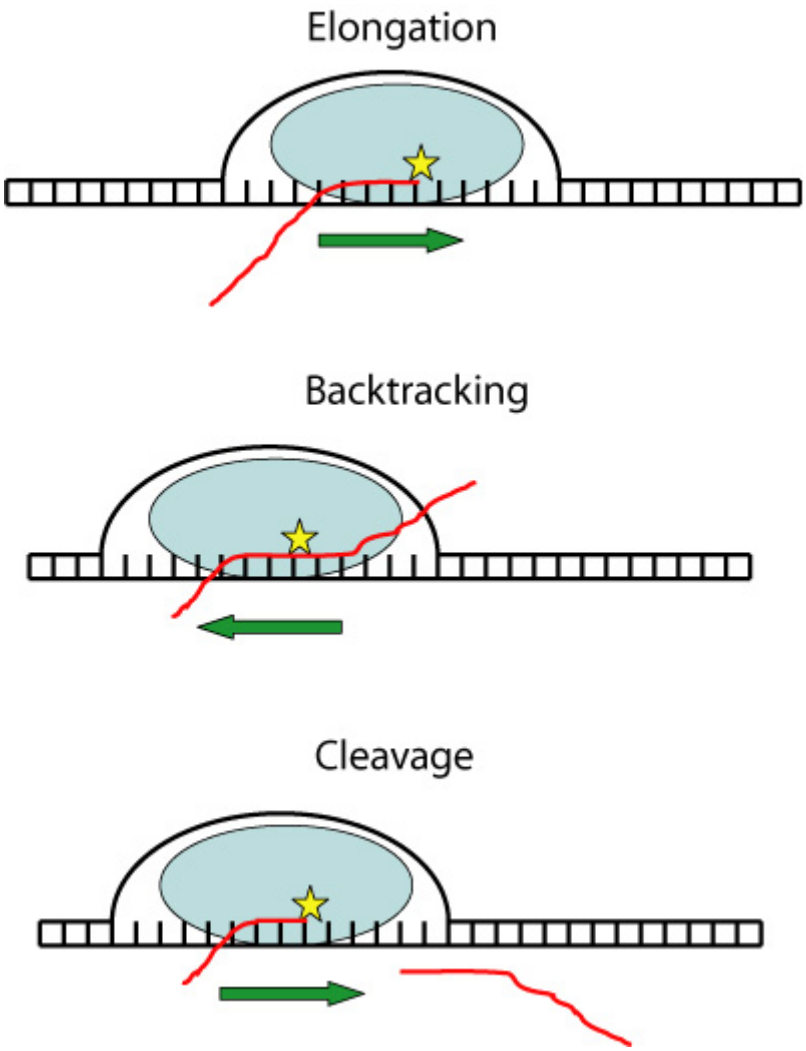


Figure 1.3 Cartoon Model of Elongation, Backtracking, and Cleavage States. RNAP is represented as a cyan oval with its catalytic site denoted by the star. Under normal conditions, the enzyme translocates along the DNA (black) in the direction of the arrow and forms a nascent RNA chain (red). Upon falling into an errant state, the enzyme will stop forward translocate and reverse direction thereby extruding the RNA from the enzyme. RNAP can be rescued from this state by cleavage of the RNA transcript and normal productive synthesis can resume.

hydrolyzed and released from the enzyme leaving a new 3'-OH on the 5' RNA fragment which remains bound to RNAP. The enzyme is able to resume elongation in the unactivated state.

1.2.4 Termination

At some point, RNAP will come to the end of the gene and must terminate. Termination falls into two main categories, intrinsic and extrinsic termination. In intrinsic termination RNAP encounters a template DNA sequence consisting of a palindromic GC-rich region followed by an AT-rich region. The resulting RNA transcript folds into a stable hairpin structure. The hairpin sequence is followed by a sequence of several uridines. The combination of the hairpin and uridine rich sequence causes RNAP to dissociate from the DNA and release its transcript. The alternate method of termination, extrinsic termination, requires the use of the external protein factor Rho. Rho interacts with specific sequences on the nascent RNA transcript. Upon binding, Rho actively translocates along the transcript and eventually breaks the DNA/RNA hybrid (Richardson 1993). The common feature between intrinsic and extrinsic termination is that it is the nascent RNA transcript that governs termination rather than the DNA template.

1.3 RNA Polymerase Structure

In the past 10 years, the level of structural detail has dramatically increased and has lead to unprecedented access to the inner workings of both prokaryotic and eukaryotic RNAPs and even lead to a Nobel Prize being awarded to Roger Kornberg in

2006 for his work on solving crystal structures of *S. cerevisiae* RNAPII elongation complexes. Several crystal structures of prokaryotic RNAPs have been determined for *T. aquaticus* core enzyme (Zhang, Campbell *et al*. 1999) and *T. thermophilus* holoenzyme (Vassylyev, Sekine *et al*. 2002). In addition, prokaryotic RNAPs have been co-crystalized with various small molecules, such as rifampicin, streptolydigin, and ppGpp (Campbell, Korzheva *et al*. 2001; Bushnell, Cramer *et al*. 2002; Artsimovitch, Patlan *et al*. 2004; Temiakov, Zenkin *et al*. 2005; Vassylyev, Vassylyeva *et al*. 2007).

Recent crystal structures of both RNAPII and *T. thermophilus* RNAP have been solved with the DNA, the DNA/RNA hybrid, and various NTPs bound in the catalytic site. These structures have yielded significant information regarding the structure/function relationship of elongation complexes (Cramer, Bushnell *et al*. 2001; Gnatt, Cramer *et al*. 2001; Kettenberger, Armache *et al*. 2004; Westover, Bushnell *et al*. 2004; Wang, Bushnell *et al*. 2006; Vassylyev, Vassylyeva *et al*. 2007; Vassylyev, Vassylyeva *et al*. 2007).

1.3.1 Structural Elements of RNA Polymerase

Crystal structures of both prokaryotic and eukaryotic RNAP reveals an enzyme that has a basic overall structure resembling a crab claw. The two largest subunits, β and β', form the "pincers" of the enzyme (Figure 1.4). There is a deep channel that runs the length of the claw and is referred to as the main channel. There are several other important structures found in the main channel that have been shown to be required for proper elongation and catalysis (Figure 1.5). The bridge helix, which is part of the β' subunit and spans the main channel, is thought to play an important role in translocation

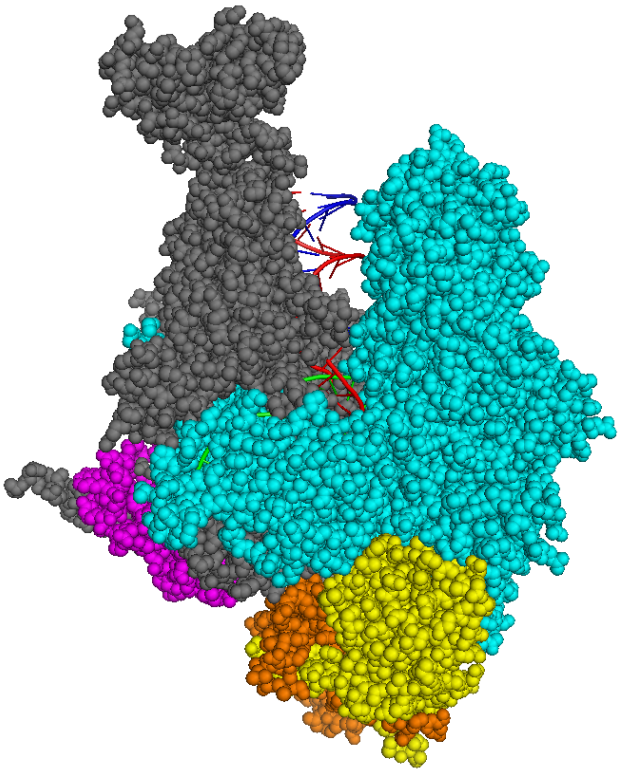


Figure 1.4 Overall Spacefilled Model of *T. thermophilus* RNAP Elongation Complex. The overall structure resembles a crabclaw. The β-subunit is colored in cyan, β'-subunit is colored in grey, the two α-subunits are colored yellow and orange, and the ω-subunit is colored purple. The DNA can be seen in the main channel. The template strand is colored red, the non-template strand is blue, and the RNA is colored green.

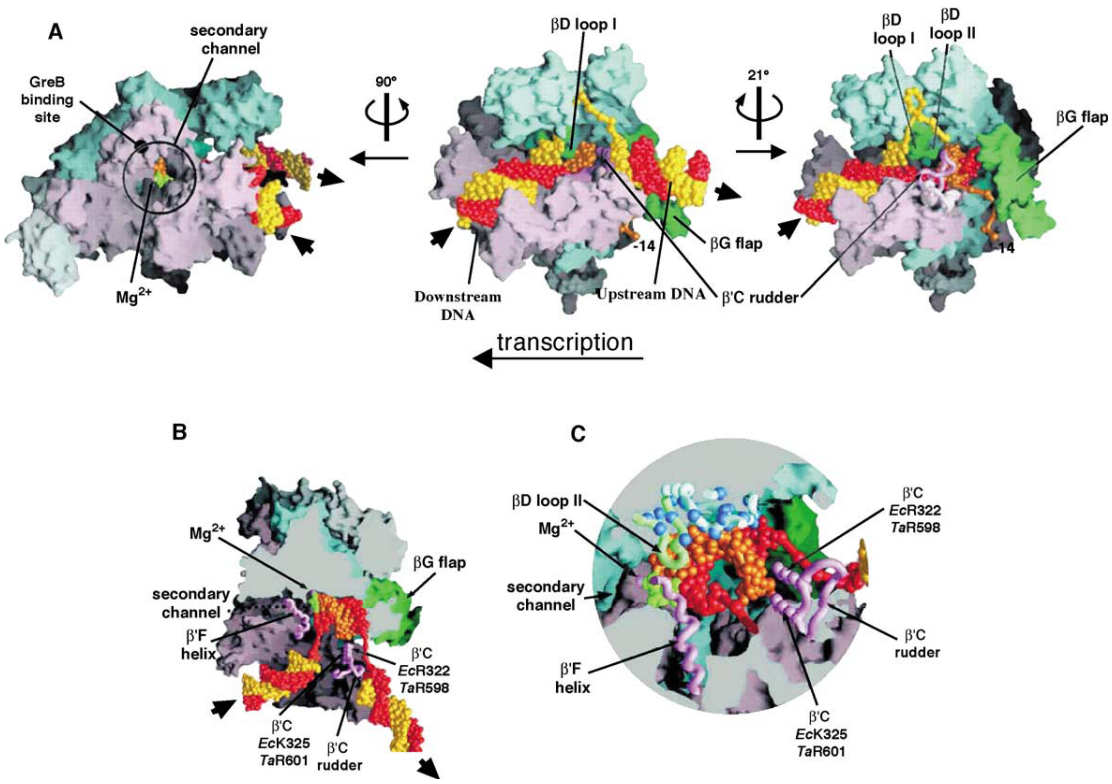


Figure 1.5 Various Views of the Bacterial Elongation Complex Based on the *T. aquaticus* Structure (Korzheva, Mustaev *et al*. 2000). RNAP is represented in as a molecular surface model (β', pink; β, cyan; α, green; ω, white). The DNA (template strand, red; nontemplate strand, yellow) enters the main channel and the DNA strands separate at βD-loopI (fork loop 2). The template strand enters the catalytic site (catalytic Mg^{+2} is purple sphere) and the RNA (orange) is synthesized. Several important structural elements mentioned in the main text are labeled. PDB ID: 2PPD

(Epshtein, Mustaev *et al*. 2002; Artsimovitch, Chu *et al*. 2003; Temiakov, Zenkin *et al*. 2005; Tuske, Sarafianos *et al*. 2005). The trigger loop, which is a mobile element located under the bridge helix, is required for proper catalysis (Temiakov, Zenkin *et al*. 2005). The catalytic site is located at the junction of β and β'. The catalytic site is composed of 3 invariant aspartic acid residues that are located in a motif with a NADFDGD amino acid sequence. The aspartic acids chelate two magnesium cations that are essential for catalysis.

Another major structural element seen in all RNAP crystal structures is the secondary channel (Figure 1.5). This channel, which is located in the β'-subunit, is a ~12 Å wide and ~45 Å deep tunnel that leads directly into the catalytic site (Zhang, Campbell *et al*. 1999; Gnatt, Cramer *et al*. 2001). During elongation, the main channel is filled with the DNA, therefore, it is generally believed that the secondary channel is the main route for NTP entry into the catalytic site. (Zhang, Campbell *et al*. 1999; Cramer, Bushnell *et al*. 2001; Gnatt, Cramer *et al*. 2001; Batada, Westover *et al*. 2004). However, the size of the channel precludes more than one nucleotide from being bound inside the channel which may lead to an NTP trafficking problem to the catalytic site (Korzheva, Mustaev *et al*. 2000). Potential trafficking problems raise the possibility that RNAP has evolved more than one pathway for NTPs to enter the catalytic site (Foster, Holmes *et al*. 2001; Burton, Feig *et al*. 2005; Gong, Zhang *et al*. 2005; Zhang, Zobeck *et al*. 2005). Apart from its role in NTP entry, the secondary channel appears to be able to accommodate the 3'-end of an extruded RNA formed during backtracking of the elongation complex. The movement causes a displacement of the 3'-end of the RNA into the secondary channel (Reeder and Hawley 1996; Komissarova and Kashlev 1997;

Nudler, Mustaev *et al*. 1997; Artsimovitch and Landick 2000). Additionally, several transcription factors, such as GreA, GreB, and DksA in prokaryotes (Opalka, Chlenov *et al*. 2003; Sosunova, Sosunov *et al*. 2003; Perederina, Svetlov *et al*. 2004) and TFIIS in eukaryotes (Kettenberger, Armache *et al*. 2004), have been shown interact with RNAP through the secondary channel. In both of these cases, the secondary channel becomes occluded. This observation raises the possibility that NTPs are able to enter the catalytic site via a different route when the enzyme is in a regulatory and or errant state.

The transcription bubble, which consists of the DNA/RNA hybrid and the melted DNA bubble, is located in the main channel of RNAP (Figure 1.6A and Figure 1.6B). The boundaries of the transcription bubble are delineated by the rudder, and fork loop 2, which has been suggested to maintain the downstream edge of the bubble (Kettenberger, Armache *et al*. 2004). The template DNA follows a path from the bubble opening, across the bridge helix, at which point it makes a 90º turn. The RNA/DNA hybrid leads from the catalytic site and is maintained for approximately nine base pairs. The rudder, which is thought to maintain the upstream edge of the transcription bubble, separates the RNA from the DNA. The RNA is then extruded through the RNA exit channel which is composed of the "lid" and "saddle" structural elements in yeast RNAP II and the βG flap in prokaryotes (Figure 1.6) (Zhang, Campbell et al. 1999; Cramer, Bushnell et al. 2001; Kettenberger, Armache et al. 2004).

The biochemical data indicate that RNAP is capable of entering into many different states. Structurally, these different states are the result of the enzyme adopting different conformations. Structural elements, such as the bridge helix and β-pincer, have been seen in several different conformations and are thought to play a very important role

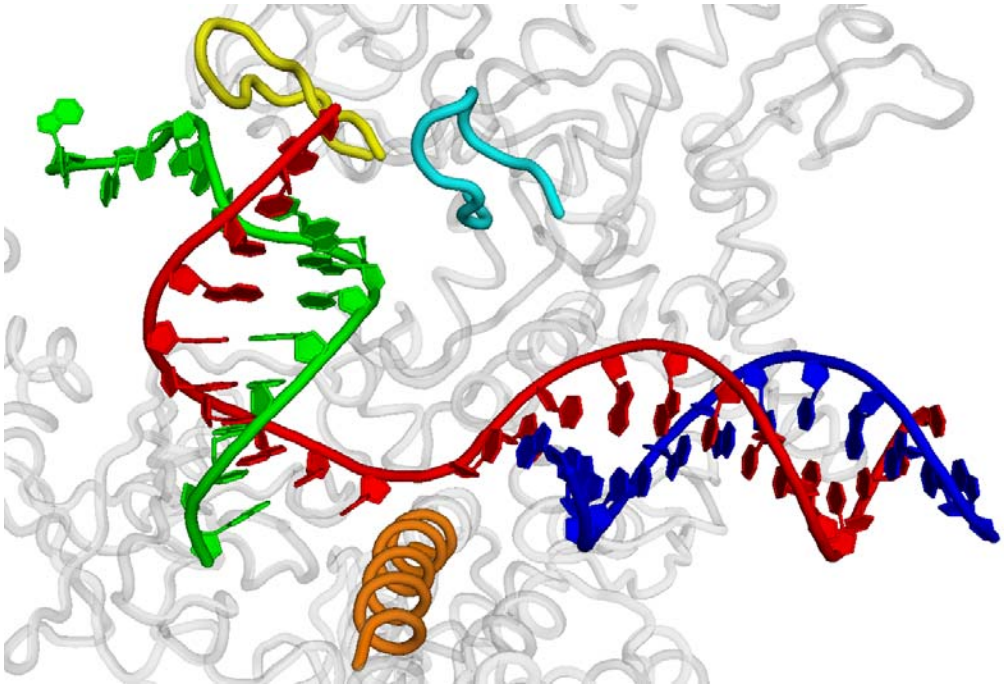


Figure 1.6 Path of the Nucleic Acids in the Transcription Elongation Complex.

A. View of the path of nucleic acids from the *T. thermophilus* elongation complex. All subunits have been removed with the exception of β' (transparent grey). Several key structures found in β' are represented in color (bridge helix, orange; rudder, cyan; and lid/saddle, yellow). The DNA (template strand, red; nontemplate strand, blue) come into the enzyme via the main channel. The template strand makes a 90° turn at the catalytic site and forms a DNA/RNA hybrid (RNA, green). The hybrid then exits the enzyme. PDB ID: 2PPD

in translocation (Zhang, Campbell *et al*. 1999; Vassylyev, Sekine *et al*. 2002; Vassylyev, Vassylyeva *et al*. 2007). Other critical structural elements found in the main channel have been seen in multiple conformations. For example, the trigger loop has been seen in multiple conformations in both *T. thermophilus* and *S. cerevisiae* elongation complexes (Figure 2.18). This loop is catalytically essential and, upon interaction with the catalytic site, occludes the secondary channel. In addition to the trigger loop, fork loop 2 is able to adopt different conformations (Figure 2.18). Interestingly, based on kinetic experiments, fork loop 2 has been suggested to constitute an NTP binding site (Foster, Holmes *et al*. 2001; Holmes and Erie 2003).

Due to the large conformational complexity found in RNAP, it is vital that we be able to tease out all the steps in elongation to be able to fully understand the inner workings of the enzyme. Specifically, we can use pre-steady state kinetics to determine the steps involved in the nucleotide addition cycle. These data, combined with the various structures that have been published in the past 10 years, will allow for a detailed picture of transcription elongation.

BIBLIOGRAPHY

Allison, L. A., M. Moyle, et al. (1985). "Extensive Homology among the Largest Subunits of Eukaryotic and Prokaryotic RNA Polymerases." Cell **42**(2): 599-610.

Arndt, K. M. and M. J. Chamberlin (1990). "RNA Chain Elongation by Escherichia coli RNA Polymerase: Factors Affecting the Stability of Elongating Ternary Complexes." J. Mol. Bio. **213**(1): 79-108.

Artsimovitch, I., C. Chu, et al. (2003). "A New Class of Bacterial RNA Polymerase Inhibitor Affects Nucleotide Addition." Science **302**(5645): 650-654.

Artsimovitch, I. and R. Landick (2000). "Pausing by Bacterial RNA Polymerase is Mediated by Mechanistically Distinct Classes of Signals." Proc. Nat. Acad. Sci. USA **97**(13): 7090-7095.

Artsimovitch, I., V. Patlan, et al. (2004). "Structural Basis for Transcription Regulation by Alarmone ppGpp." Cell **117**(3): 299-310.

Batada, N. N., K. D. Westover, et al. (2004). "Diffusion of Nucleoside Triphosphates and Role of the Entry Site to the RNA Polymerase II Active Center." Proc. Nat. Acad. Sci. USA **101**(50): 17361-17364.

Biggs, J., L. L. Sesrles, et al. (1985). "Structure of the Eukaryotic Transcription Apparatus: Features of the Gene for the Largest Subunit of Drosophila RNA Polymerase II." Cell **42**(2): 611-621.

Boukhov, S., A. Polyakov, et al. (1992). "GreA Protein: A Transcription Elongation Factor from *Escherichia coli*." Proc. Nat. Acad. Sci. USA **89**(19): p8899-8902.

Burton, Z. F., M. Feig, et al. (2005). "NTP-Driven Translocation and Regulation of Downstream Template Opening by Multi-Subunit RNA Polymerases." Biochem. Cell. Bio. **83**: 486-496.

Bushnell, D. A., P. Cramer, et al. (2002). "Structural Basis of Transcription: α-Amanitin–RNA Polymerase II Cocrystal at 2.8 Å Resolution." Proc. Nat. Acad. Sci. USA **99**(3): 1218-1222.

Campbell, E. A., N. Korzheva, et al. (2001). "Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase." Cell **104**(6): 901-912.

Chistopher K. Surratt, S. C. Milan, et al. (1991). "Spontaneous Cleavage of RNA in Ternary Complexes of *Escherichia coli* RNA Polymerase and its Significance for the Mechanism of Transcription." Proc. Nat. Acad. Sci. USA **88**(18): 7983-7987.

Cramer, P., D. A. Bushnell, et al. (2001). "Structural Basis of Transcription: RNA Polymerase II at 2.8 Ångstrom Resolution." Science **292**(5523): 1863-1876.

Davenport, R. J., G. J. L. Wuite, et al. (2000). "Single-Molecule Study of Transcriptional Pausing and Arrest by E. coli RNA Polymerase." Science **287**(5462): 2497-2500.

Epshtein, V., A. Mustaev, et al. (2002). "Swing-Gate Model of Nucleotide Entry into the RNA Polymerase Active Center." Mol. Cell **10**(3): 623-634.

Erie, D. A., O. Hajiseyedjavadi, et al. (1993). "Multiple RNA Polymerase Conformations and GreA: Control of the Fidelity of Transcription." Science **262**(5135): 867-873.

Foster, J. E., S. F. Holmes, et al. (2001). "Allosteric Binding of Nucleoside Triphosphates to RNA Polymerase Regulates Transcription Elongation." Cell **106**: 243-252.

Gnatt, A. L., P. Cramer, et al. (2001). "Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 A Resolution." Science **292**(5523): 1876-1882.

Gnatt, A. L., P. Cramer, et al. (2001). "Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 Å Resolution." Science **292**(5523): 1876-1882.

Gong, X. Q., C. Zhang, et al. (2005). "Dynamic Error Correction and Regulation of Downstream Bubble Opening by Human RNA Polymerase II." Mol. Cell **18**(4): 461-470.

Holmes, S. F. and D. A. Erie (2003). "Downstream DNA Sequence Effects on Transcription Elongation: Allosteric Binding of Nucleoside Triphosphates Facilitates Translocation Via a Ratchet Motion." J. Bio. Chem. **278**(37): 35597-35608.

Izban, M. G. and D. S. Luse (1992). "The RNA Polymerase II Ternary Complex Cleaves the Nascent Transcript in a 3'→5' Direction in the Presence of Elongation Factor SII." Gene. Dev. **6**(7): 1342-1356.

Kapanidis, A. N., E. Margeat, et al. (2006). "Initial Transcription by RNA Polymerase Proceeds Through a DNA-Scrunching Mechanism." Science **314**(5802): 1144-1147.

Kettenberger, H., K.-J. Armache, et al. (2004). "Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS." Mol. Cell **16**(6): 955-965.

Komissarova, N. and M. Kashlev (1997). "RNA Polymerase Switches between Inactivated and Activated States By Translocating Back and Forth along the DNA and the RNA." J. Bio. Chem. **272**(24): 15329-15338.

Komissarova, N. and M. Kashlev (1997). "Transcriptional Arrest: *Escherichia coli* RNA Polymerase Translocates Backward, Leaving the 3'-end of the RNA Intact and Extruded." Proc. Nat. Acad. Sci. USA **94**(5): 1755-1760.

Korzheva, N., A. Mustaev, et al. (2000). "A Structural Model of Transcription Elongation." Science **289**(5479): 619-625.

Krummel, B. and M. J. Chamberlin (1992). "Structural Analysis of Ternary Complexes of *Escherichia coli* RNA Polymerase Individual Complexes Halted along Different Transcription Units Have Distinct and Unexpected Biochemical Properties." J. Mol. Bio. **225**(2): 221-237.

Landick, R. (2006). "The Regulatory Roles and Mechanism of Transcriptional Pausing." Biochem. Soc. Trans. **34**(6): 1062-1066.

Mustaev, A., M. Kashlev, et al. (1993). "Active Center Rearrangement in RNA Polymerase Initiation Complex." J. Bio. Chem. **268**(26): 19185-19187.

Nudler, E., A. Mustaev, et al. (1997). "The RNA–DNA Hybrid Maintains the Register of Transcription by Preventing Backtracking of RNA Polymerase." Cell **89**(1): 33-41.

Opalka, N., M. Chlenov, et al. (2003). "Structure and Function of the Transcription Elongation Factor GreB Bound to Bacterial RNA Polymerase." Cell **114**(3): 335-345.

Perederina, A., V. Svetlov, et al. (2004). "Regulation Through the Secondary Channel: Structural Framework for ppGpp-DksA Synergism During Transcription." Cell **118**(3): 297-309.

Reeder, T. C. and D. K. Hawley (1996). "Promoter Proximal Sequences Modulate RNA Polymerase II Elongation by a Novel Mechanism." Cell **87**(4): 767-777.

Revyakin, A., C. Liu, et al. (2006). "Abortive Initiation and Productive Initiation by RNA Polymerase Involve DNA Scrunching." Science **314**(5802): 1139-1143.

Richardson, J. P. (1993). "Transcription Termination." Crit. Rev. Biochem. Mol. Bio. **28**(1).

Sosunova, E., V. Sosunov, et al. (2003). "Donation of Catalytic Residues to RNA Polymerase Active Center by Transcription Factor Gre." Proc. Nat. Acad. Sci. USA **100**(26): 15469-15474.

Straney, D. C. and D. M. Crothers (1987). "A Stressed Intermediate in the Formation of Stably Initiated RNA Chains at the *Escherichia coli lac* UV5 Promoter." J. Mol. Bio. **193**(2): 267-278.

Surratt, C. K., S. C. Milan, et al. (1991). "Spontaneous Cleavage of RNA in Ternary Complexes of *Escherichia coli* RNA Polymerase and its Significance for the Mechanism of Transcription." Proc. Nat. Acad. Sci. USA **88**(18): 7983-7987.

Sweetser, D., M. Nonet, et al. (1987). "Prokaryotic and Eukaryotic RNA Polymerases have Homologous Core Subunits." Proc. Nat. Acad. Sci. USA **84**(5): 1192-1196.

Temiakov, D., N. Zenkin, et al. (2005). "Structural Basis of Transcription Inhibition by Antibiotic Streptolydigin." Mol. Cell **19**(5): 655-666.

Thomas, M. C. and C.-M. Chiang (2006). "The General Transcription Machinery and General Cofactors." Crit. Rev. Biochem. Mol. Bio. **41**(3): 105-178.

Tolić-Nørrelykke, S. F., A. M. Engh, et al. (2004). "Diversity in the Rates of Transcript Elongation by Single RNA Polymerase Molecules." J. Bio. Chem. **279**(5): 3292-3299.

Toulokhonov, I., J. Zhang, et al. (2007). "A Central Role of the RNA Polymerase Trigger Loop in Active-Site Rearrangement during Transcriptional Pausing." Mol. Cell **27**(3): 406-420.

Tuske, S., S. G. Sarafianos, et al. (2005). "Inhibition of Bacterial RNA Polymerase by Streptolydigin: Stabilization of a Straight-Bridge-Helix Active-Center Conformation." Cell **122**(4): 541-552.

Vassylyev, D. G., S. Sekine, et al. (2002). "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution." Nature **417**(6890): 712-719.

Vassylyev, D. G., M. N. Vassylyeva, et al. (2007). "Structural Basis for Transcription Elongation by Bacterial RNA Polymerase." Nature **448**(7150): 157.

Vassylyev, D. G., M. N. Vassylyeva, et al. (2007). "Structural Basis for Substrate Loading in Bacterial RNA Polymerase." Nature **448**(7150): 163-168.

Wang, D., D. A. Bushnell, et al. (2006). "Structural Basis of Transcription: Role of the Trigger Loop in Substrate Specificity and Catalysis." Cell **127**(5): 941-954.

Westover, K. D., D. A. Bushnell, et al. (2004). "Structural Basis of Transcription: Nucleotide Selection by Rotation in the RNA Polymerase II Active Center." Cell **119**(4): 481-489.

Yin, H., I. Artsimovitch, et al. (1999). "Nonequilibrium Mechanism of Transcription Termination from Observations of Single RNA Polymerase Molecules." Proc. Nat. Acad. Sci. USA **96**(23): 13124.

Zhang, C., K. L. Zobeck, et al. (2005). "Human RNA Polymerase II Elongation in Slow Motion: Role of the TFIIF RAP74 α1 Helix in Nucleoside Triphosphate-Driven Translocation." Mol. Cell. Bio. **25**(9): 3583-3595.

Zhang, G., E. A. Campbell, et al. (1999). "Crystal Structure of *Thermus aquaticus* Core RNA Polymerase at 3.3 Å Resolution." Cell **98**(7): 811-824.

CHAPTER 2:

MAIN CHANNEL STRUCTURES ACT AS A GATE KEEPER FOR NTP ADDITION IN TRANSCRIPTION

2.1 Introduction

The central role of RNA polymerase (RNAP) in transcription is to catalyze the processive synthesis of the growing RNA transcript. It has been observed that RNAP can exist in multiple conformations during transcription elongation. Two states of interest, which are thought to play an important role in regulation, are an "activated" state, which is characterized by rapid synthesis, and an "unactivated" state, which has a significantly slower rate of synthesis. These states affect pausing, termination, and the operation of extrinsic protein factors (Erie, Hajiseyedjavadi *et al*. 1993; Matsuzaki, Kassavetis *et al*. 1994; Yin, Artsimovitch *et al*. 1999; Davenport, Wuite *et al*. 2000; Erie 2002; Tolić-Nørrelykke, Engh *et al*. 2004).

From studies of pausing and termination, it has been shown that the sequence of the downstream DNA plays an important role in the regulation of RNAP. For example, the formation of paused complexes is directly dependent on the downstream DNA and adjacent sites. Any alterations in the underlying sequence affects pausing (Telesnitsky and Chamberlin 1989; Reynolds, Bermbdez-Cruz *et al*. 1992; Palangat, Hittinger *et al*. 2004; Herbert, Porta *et al*. 2006).

Recently, crystal structures of RNAP have allowed a structural explanation of the downstream DNA sequence effects on transcription. Two main structural features of

RNAP are thought to play an important role in transcription regulation: the main channel, which is filled with the downstream DNA, and a negatively-charged, funnel shaped pore (called the secondary channel) that leads from the surface of the enzyme to the active site. These two channels are separated by the bridge helix and trigger loop, which make intimate contact with the downstream DNA and partially overlap with the catalytic site. In recent crystal structures of both prokaryotic and eukaryotic elongation complexes, both of these structures have been seen in multiple conformations; specifically, the trigger loop was found to be an integral component of the catalytic site (Wang, Bushnell *et al*. 2006; Vassylyev, Vassylyeva *et al*. 2007).

Using transient state kinetics to study nucleotide addition in *E. coli* RNAP, our lab has previously identified an allosteric site that is utilized during elongation (Foster, Holmes *et al*. 2001). The data for these experiments were fit to a non-essential activation mechanism in which there is rapid equilibrium binding to an allosteric site. Because it is known that the downstream DNA can help regulate transcription, Holmes and Erie (2003) examined the effects of changing the downstream DNA sequence on transcription elongation. From these experiments, they further refined the mechanism previously proposed by Foster *et al.* (2001) (Figure 2.1a, 2.1b). Furthermore, based on structural analysis of several crystal structures, they proposed that the putative allosteric site exists in the main channel of the enzyme (Holmes and Erie 2003).

The suggested site has several important features that are consistent with it being a NTP binding site (Figure 2.2). The site consists of a flexible loop structure (fork loop 2, aka βD-loop I) that contains a number of glycines in both prokaryote and eukaryotes (Figure 2.2B). Furthermore, fork loop 2 is surrounded by a β-sheet on one side and

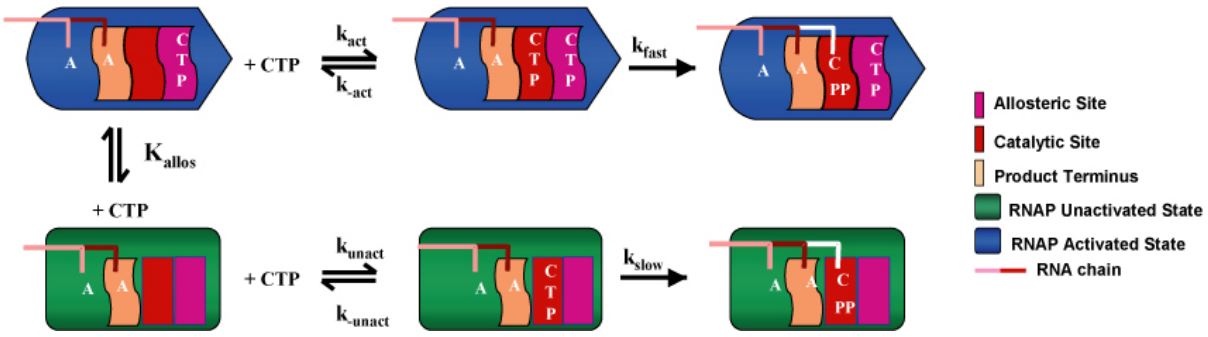


Figure 2.1A

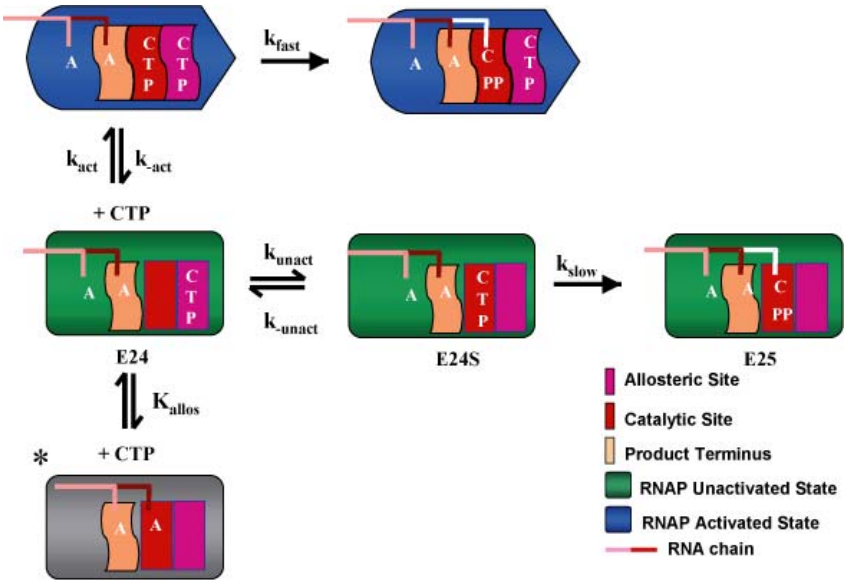


Figure 2.1B

Figure 2.1 Two Kinetically Identical Non-essential Activation Mechanisms for NTP Incorporation.

Figure 2.1 Two Kinetically Identical Non-essential Activation Mechanisms for NTP Incorporation (Holmes and Erie 2003). The two paths of synthesis are designated in blue and green. The blue pathway represents the activated (fast) state of synthesis while the green pathway represents the unactivated (slow) state of synthesis. The red and magenta boxes represent the catalytic site and allosteric site, respectively. The salmon color box represents the product terminus binding site. The lines connecting to the designated nucleotide represent the RNA chain. PP_i stands for pyrophosphate. S_A and S_C represent NTP binding to the allosteric and catalytic site, respectively.

A. **Kinetic Mechanism that Assumes an Equilibrium Between the Pre- and Post-translocated State.** Binding of a templated NTP to the allosteric site shifts the enzyme to the activated state. The shifting of the enzyme into the activated state is shown in one step but is actually several steps and is embedded in K_{allos}. Specifically,

$$E + S \underset{}{\overset{K_{bind}}{\rightleftharpoons}} E{:}S_a \overset{K_t}{\rightleftharpoons} E^*S_a$$

$K_{allos} = K_{bind} * K_t$, where K_{bind} is the binding constant for the allosteric site and K_t is the transition from the unactivated state to the activated state. The grey box represents the pre-translocated state and was not used in determining this mechanism. The equilibrium between the pre- and post-translocated states, K_{trans}, is contained in k_{unact} and K_{allos}.

In this mechanism, if an NTP binds directly into the catalytic site without using the allosteric site, then synthesis proceeds along the unactivated path. The

distribution of complexes between the activated and unactivated state is determined by a simple competition between NTPs binding to the catalytic site or the allosteric site first.

B. Kinetic Mechanism in which the Shift Between Pre- and Post-translocation is Induced by NTP Binding. The mechanism is kinetically identical to the mechanism presented in **A.** The grey box represents the pre-translocated state. In this model, an NTP binds to the allosteric site and drives the enzyme to the post-translocated state.

In this mechanism, an NTP binds to the allosteric site, which facilitates translocation. At this point in the mechanism, there is a competition between the allosteric NTP entering the catalytic site (the unactivated state) or a second NTP binding directly to the catalytic site (activated state).

several α-helices on the other side. These features are indicative of "P-loops" (Walker, Saraste *et al*. 1982; Via, Ferrè *et al*. 2000; Leipe, Wolf *et al*. 2002). Additionally, a totally conserved Walker B motif is located in close proximity to fork loop 2 (Figure 2.2C). The Walker B motif is a well conserved amino acid sequence and that interacts with the phosphate groups, found in many ATP binding proteins (Walker, Saraste *et al*. 1982).

Several models of transcription elongation suggest that RNAP uses the main channel in elongation to bind NTPs, coded for by the downstream DNA, to facilitate translocation of RNAP along the DNA. Additionally, some models suggest the NTPs may enter the catalytic site via the main channel; however, a major caveat of these models is that there is no direct evidence for NTPs binding to the main channel during elongation. In recent crystal structures of yeast RNAPII and *T. thermophilus* elongation complexes, NTPs have only been observed bound in the catalytic site and secondary channel, but not in the main channel. From these structures and other biochemical data, it has been proposed that NTP entry to the catalytic site occurs exclusively through the secondary channel (Zhang, Campbell *et al*. 1999; Batada, Westover *et al*. 2004; Mukhopadhyay, Sineva *et al*. 2004; Westover, Bushnell *et al*. 2004).

Because there is no direct evidence that the main channel is used for NTP binding during elongation, we were interested in testing this possibility. Using transient-state kinetics, we investigated the role of amino acids in the main channel in NTP binding and nucleotide incorporation. Specifically, we characterized the incorporation of multiple nucleotides into a growing RNA transcript for wild-type (wt-RNAP) and a mutant in which amino acids R542-F545 in the β-subunit have been deleted (Δ-loop RNAP). In

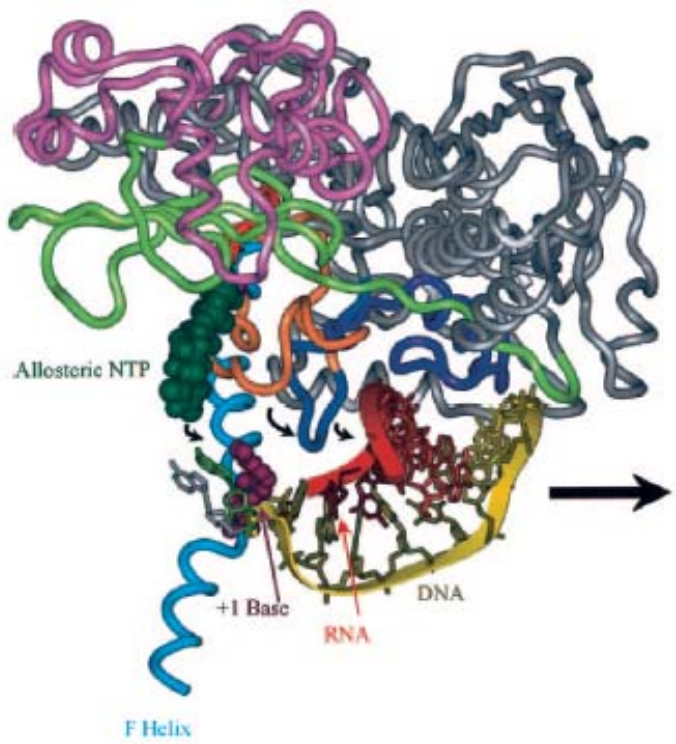


Figure 2.2A

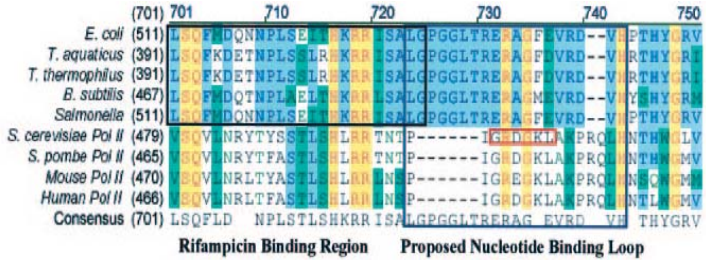


Figure 2.2B

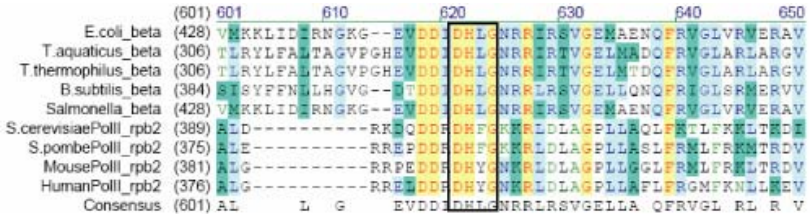


Figure 2.2C

Figure 2.2 Structure of Putative Allosteric Site (Holmes and Erie 2003).

Figure 2.2 Structure of Putative Allosteric Site (Holmes and Erie 2003).

A. The DNA template strand is shown in yellow and the RNA in red. The bridge helix is depicted in cyan, the β-sheet region and flanking helices are colored in green and pink, respectively. A Walker B motif is in red. The rifampicin binding region is colored in blue and fork loop 2, which is the proposed NTP binding site, is colored in orange. A modeled in UTP (dark-green space-fill) binds is able to base pair to the downstream DNA base (adenosine) (purple space-fill).

B. Sequence alignment of the β-subunit region that contains the fork loop 2 (Holmes and Erie 2003). The name of the organism from which the sequence is from is listed on the left with the number of the starting residue in parenthesis. The consensus sequence is listed on the bottom. The Walker B motif is boxed in black. Totally conserved residues are highlighted in yellow while residues that are conserved solely in prokaryotes are highlighted in blue. Well conserved residues are highlighted in green.

C. Sequence alignment of the β-subunit region that contains the Walker B motif (Holmes and Erie 2003). The descriptive details are the same as presented in **B.**

wt-RNAP, several of these residues are totally conserved in all RNAPs and comprise part of the fork loop 2, which lies across from the downstream template DNA. As discussed below, deletion of these residues has a moderate effect on single nucleotide incorporation (i+1 position); however, the rate of incorporation of the second nucleotide (i+2 position) is dramatically reduced. Our results indicate that RNAP can catalyze synthesis along two pathways: a fast pathway in which NTPs first bind to fork loop 2 and a slow path which is independent of this loop. Herein, we present the first evidence indicating that RNAP utilizes the main channel during elongation by binding NTPs, in a template dependent manner, to fork loop 2. Furthermore, we provide evidence that NTPs bound to this site can be shuttled into the catalytic site from the main channel and subsequently incorporated. Using the knowledge gleaned from our kinetic data in conjunction with the crystal structures of both eukaryotic and prokaryotic TECs, we propose a structural model for nucleotide binding in the main channel and its subsequent shuttling into the catalytic site.

2.2 Kinetics of Multiple Nucleotide Incorporation for wt-RNAP

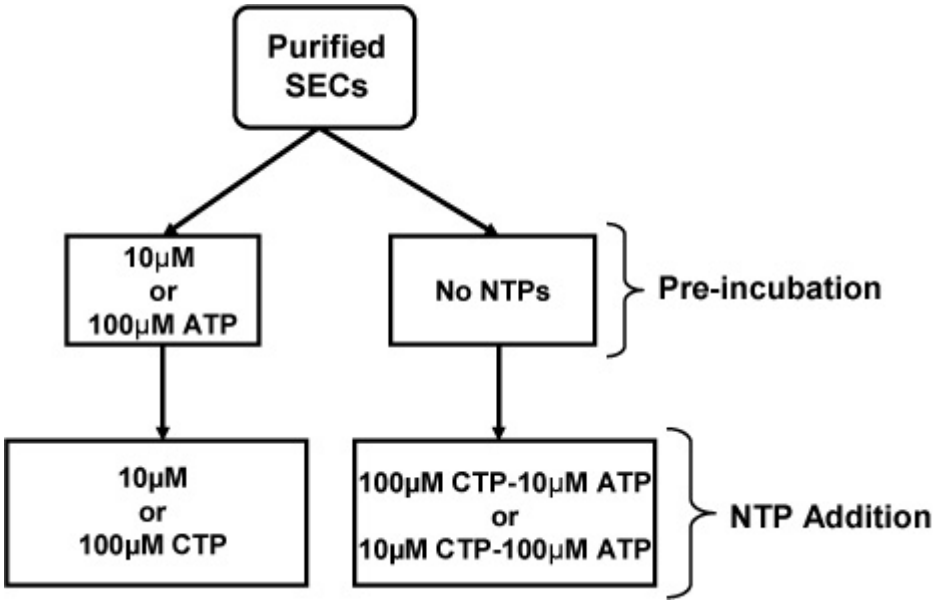
All experiments used a derivative of the transcript template DE13 (See Methods). The first CMP to be incorporated is located at position +25 in all templates. The sequence of the first 30 nucleotides of DE13 is as follows: pppAUGUAGUAAGGAGGUUGUAUGGAACAGCGC. The first cytidine to be incorporated is underlined. Due to this design, stalled elongation complexes (SEC) can be formed at position +24 by initiating transcription with addition of only ATP, GTP, and UTP to OPCs (See Methods). SECs can be purified from free NTPs and subsequent

NTPs can be added back in order to perform kinetic experiments using a rapid quench flow device. In this work, the kinetics of multiple nucleotide incorporations into a growing RNA transcript were characterized for wild-type wtRNAP and a mutant in which amino acids R542-F545 in the β-subunit have been deleted (Δ-loop RNAP). Specifically we measured the rate of CMP and AMP incorporation at template positions +25 (i+1; catalytic site) and +26 (i+2, putative allosteric site), respectively, under two conditions: one in which CTP and ATP were added simultaneously and one in which the SECs were incubated with ATP prior to initiating the reaction with the addition of CTP (Figure 2.3).

Using single nucleotide addition experiments with ATP, Holmes and Erie proposed that a putative NTP binding site is located in the main channel of RNAP and acts as an allosteric site (Holmes and Erie 2003). As discussed in the previous section, the proposed site has several important features that indicate that it is consistent with it being an NTP binding site. To address the potential role for this main channel site in nucleotide binding and incorporation, we conducted a series of kinetic experiments on RNAP, in which we measured the consequences of altering the order of addition of NTPs on the incorporation of multiple nucleotides to stalled elongation complexes (SECs).

2.2.1 Pre-incubation of ATP Increases the Rate of CMP Incorporation

To examine the effect of pre-incubating RNAP with the i+2 nucleotide (ATP) on the rates of incorporation of the i+1 nucleotide (CMP), we performed experiments in which SECs were pre-incubated with either 10μM or 100μM ATP prior to the addition of a low concentration of CTP (10μM). Simultaneous addition experiments were done as



DE13-A27g: pppAUGUAGUAAGGAGGUUGUAUGGAA**C**AGCGC

Figure 2.3 Experimental Design to Test for NTP Preloading. Purified SECs are split into two groups; one in which 10μM or 100μM ATP is added to the complexes prior to the addition of 10μM or 100μM CTP via rapid quench, and one in which 100μM CTP and 10μM ATP or 10μM CTP and 100μM AT are added simultaneously by rapid quench techniques. The sequence of the first 30nt of the DE13-A27g transcript is shown. The first CMP to be incorporated is in bold.

controls using the same NTP concentrations as in the pre-incubation experiments. As seen previously, simultaneous addition of 10 μM CTP and 10 μM or 100μM ATP does not significantly affect the rate of CMP incorporation (Foster, Holmes *et al*. 2001); however, comparison of the simultaneous addition and pre-incubation data reveal a clear increase in the rate of CMP incorporation when SECs are pre-incubated with 10 μM or 100 μM ATP prior to the addition of 10 μM CTP (Figure 2.4). Interestingly, there is an increase in the extent of complexes that incorporate CMP rapidly with increasing concentrations of pre-incubated ATP. These results are similar to those seen in experiments with RNAPII, in which the simultaneous addition of high concentrations of the i+2 nucleotide along with the i+1 nucleotide increased the rate of i+1 NMP incorporation, and indicate that the i+2 nucleotide (ATP) can facilitate the incorporation of the i+1 nucleotide (CMP) (Nedialkov, Gong *et al*. 2003; Gong, Zhang *et al*. 2005; Zhang, Zobeck *et al*. 2005). Taken together, these data indicate that ATP (i+2) can bind at a site other than the catalytic site and enhance the incorporation of both CMP (i+1) and AMP (i+2).

2.2.2 Pre-incubation of ATP Increases the Rate of AMP Incorporation

To examine the effect of pre-incubating the complexes with ATP on the rate of AMP incorporation, we used a high concentration of CTP (100μM) and a low concentration of ATP (10μM) so that CMP incorporation would not be rate limiting for AMP incorporation. The results show that pre-incubating SECs with 10μM ATP prior to initiating the reaction with CTP causes a dramatic increase in the rate of AMP incorporation relative to initiating the reaction with the simultaneous addition of CTP and

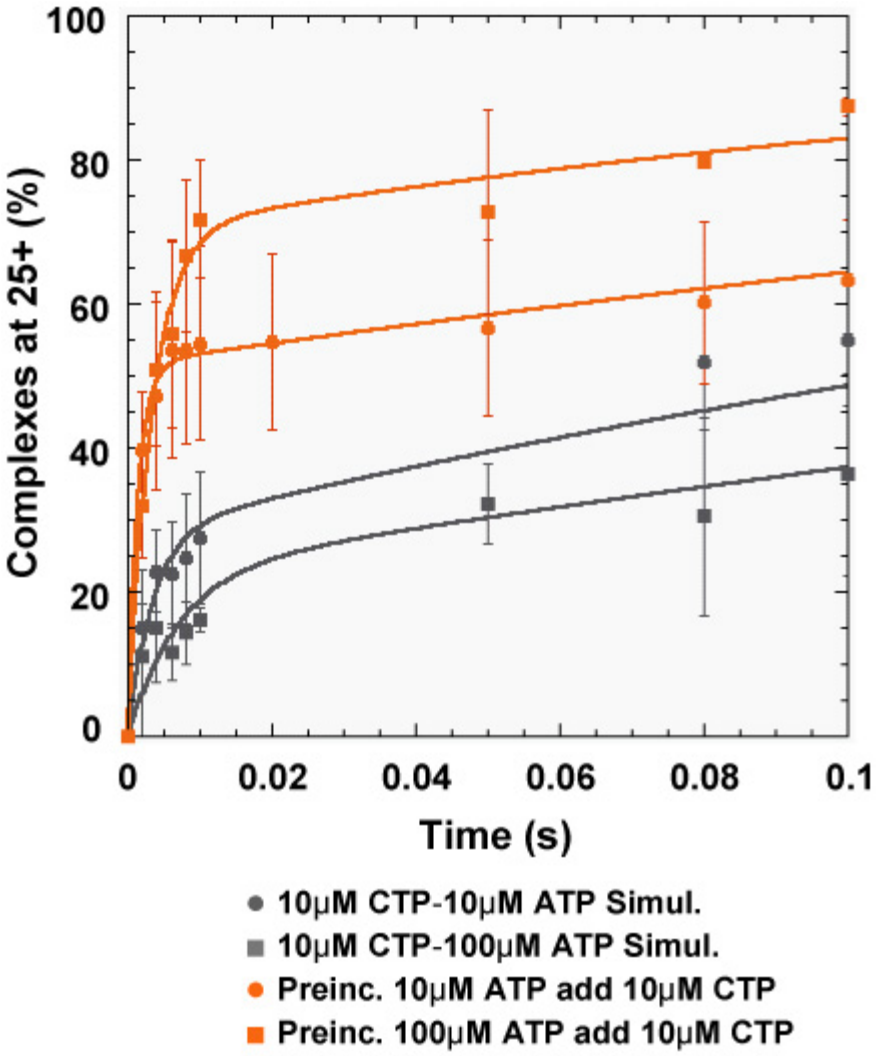


Figure 2.4 Pre-incubation with ATP Increases the Rate and Extent of CMP Incorporation. Percentage of ternary complexes at position +25 under the following conditions: Simultaneous addition of 10μM CTP-10μM ATP (grey circles); Simultaneous addition of 10μM CTP-100μM ATP (grey squares); Pre-incubation of 10μM ATP and reaction initiation with 10μM CTP (orange circles); Pre-incubation with 100μM ATP and reaction initiation with 10μM CTP (orange squares).

ATP (Figure 2.5). In fact, in the pre-incubation experiment, the rate of AMP incorporation at position +26 (i+2) is limited by the incorporation of CMP at position +25 (i+1) even though the CTP concentration is ten times higher than the ATP concentration (Figure 2.4). Because CMP must be incorporated before AMP, these results indicate that RNAP can bind ATP at a site other than the catalytic site. The ATP is templated at the i+2 position, which is situated at the proposed location for the allosteric site. This result suggests that the allosteric site functions in a template dependent manner; however these experiments do not conclusively show this to be the case.

2.2.3 Pre-incubation of ATP is not Dependent on CTP Concentration or Template Position

To assure that the pre-incubation effect is not dependent on ATP or CTP concentration, or template position, we conducted additional experiments that varied these parameters. We were interested in determining if the CTP concentration affected the rate of AMP incorporation when pre-incubated. Figure 2.6 clearly shows that even when the CTP concentration is reduced 10-fold to 10μM, AMP incorporation is limited by the rate of CMP incorporation. This result clearly indicates that initiating the reaction with 10μM CTP does not affect AMP incorporation when it is pre-incubated.

As additional controls, we performed pre-incubation experiments to determine if the identity of the i+2 NTP or the template position affected the ability of RNAP to preload the i+2 NTP. To perform these experiments, SECs stalled at position +24 were "walked" to position +26 with the addition of 10μM CTP and 10μM ATP. After stalling, the i+1 and i+2 (template positions +27 and +28) sites code for GMP and CMP,

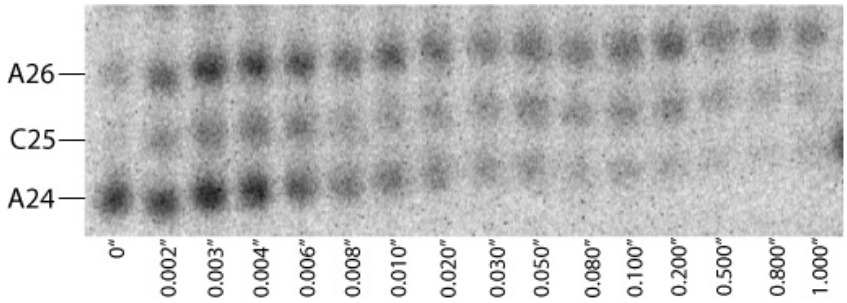


Figure 2.5A

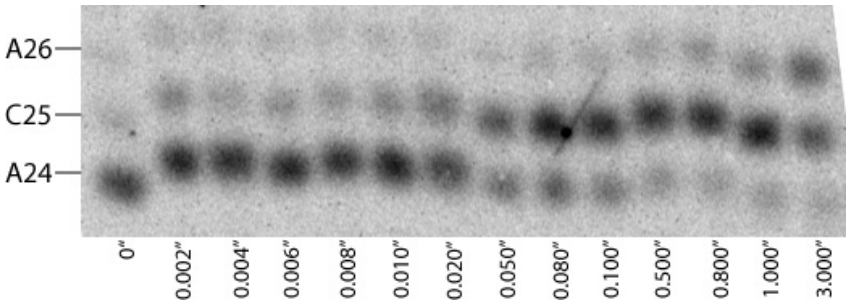


Figure 2.5B

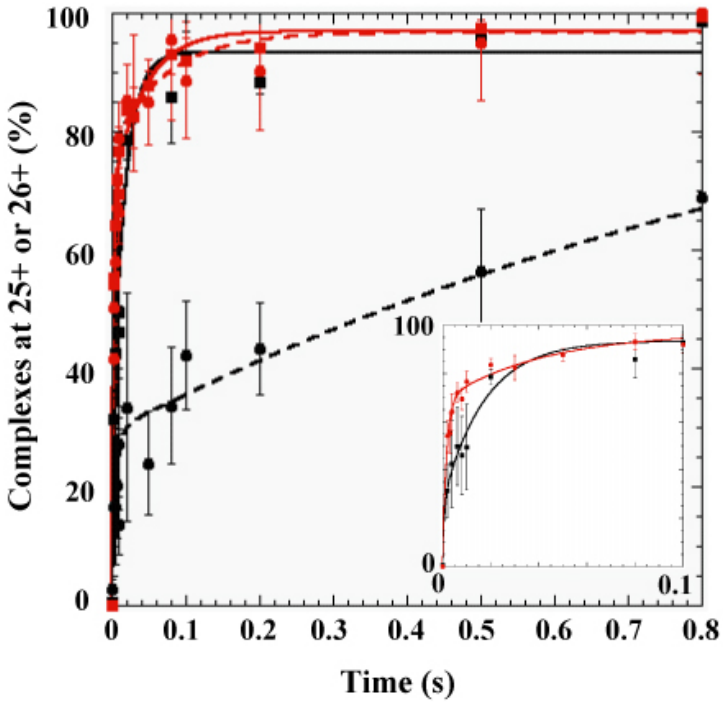


Figure 2.5C

Figure 2.5 RNA Polymerase Preloads NTPs in a Template Specific Manner.

Figure 2.5 RNA Polymerase Preloads NTPs in a Template Specific Manner.

A. Representative gel of simultaneous addition of 100μM CTP and 10μM ATP. The gels show the time course for CMP incorporation at position +25 and AMP incorporation at position +26. The positions are indicated on the left of the gel. The times are presented in seconds and denoted at the bottom of the gel.

B. Same as **A.** but it is a representative gel of pre-incubation of 10μM ATP.

C. Plots of the percentage of ternary complexes that have incorporated a CMP and AMP under the conditions of pre-incubation or simultaneous addition of NTPs. The rate of CMP incorporation is not significantly affected by the pre-incubation of ATP (Simultaneous 25+: solid black circles and dashed black line; Pre-incubation 25+: open red circles and dashed red line). A dramatic increase in the rate of AMP incorporation is observed when 10μM ATP is pre-incubated (Simultaneous 26+: solid black squares and solid black line; Pre-incubation 26+: open red squares and solid red line). Inset shows the first 0.1s of the same reaction. The inset shows the first 0.1s of the reaction for the incorporation of 100μM CMP. Interestingly, the rate of CMP incorporation is enhanced by the presence of 10μM ATP even when the [CTP] =100μM. The effect isn't nearly as large as seen at lower CTP concentrations. The coloring scheme is the same as the main plot.

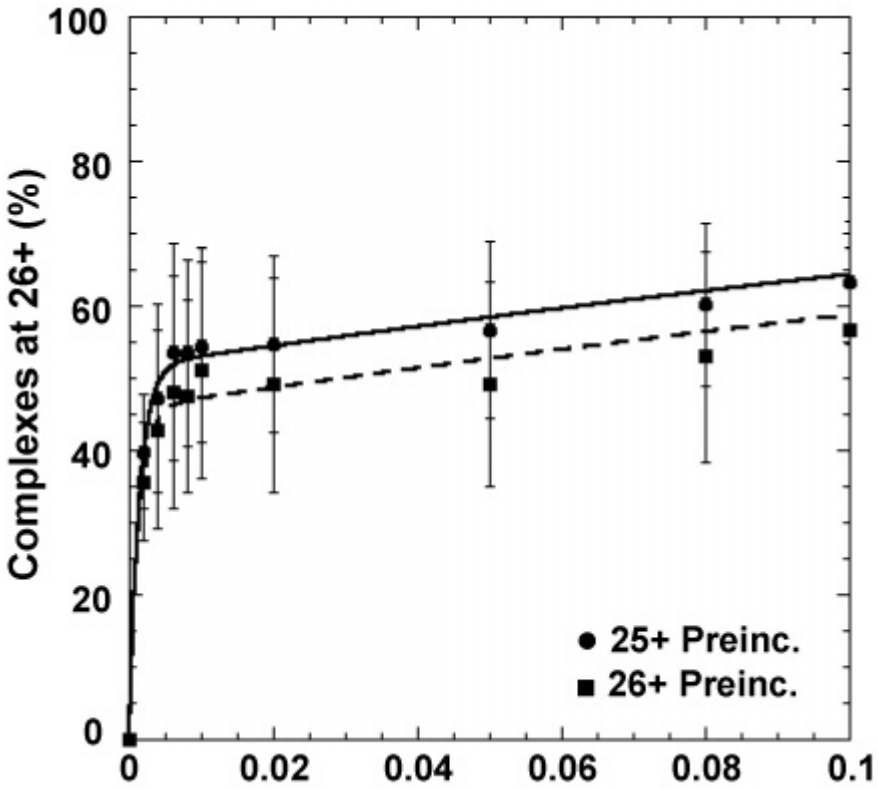


Figure 2.6 Pre-loading Effect is Independent of the Concentration of CTP. To determine if the concentration of the n+1 NTP affected the rate of incorporation for the pre-incubated n+2 NTP, SECs were pre-incubated with 10μM ATP and the reaction was initiated with the addition of 10μM CTP. The rates of CMP and AMP incorporation were measured. Closed circles are CMP incorporation and closed squares are AMP incorporation. The data were fit to double exponential curves. The solid line is the fit for the 25+ data and the dashed line is the fit for the 26+ data.

respectively. Because the complexes were not purified after walking, CTP is automatically the pre-incubated NTP. The reaction is then initiated with the addition of 10 μM GTP. Inspection of the data in Figure 2.7 clearly shows that CMP incorporation is limited by the incorporation of GMP.

From these data, we can conclude that the pre-incubation effect is independent of the i+2 NTP identity. Furthermore, the pre-incubation effect is seen at a different template position; therefore, it seems unlikely that the surrounding DNA sequence affects the ability of RNAP to preload the i+2 NTP in a template specific manner. This experiment also further strengthens the argument that the identity and concentration of the i+1 NTP does not affect pre-loading of the i+2 nucleotide.

2.2.4 Pre-incubation of ATP is not Affected by a Nonspecific Competitor

These experiments, however, do not demonstrate whether or not this site is coded for by the DNA template. To test this possibility, we performed pre-incubation experiments with 10μM ATP and 1mM GTP, which acts as a nonspecific competitor. If binding of ATP is at a nonspecific site that happens to have high affinity for NTPs, then bound NTPs should be able to be removed with a high enough concentration of a competitor. If binding is not specific, then a decrease in the rate of AMP incorporation would be expected in the presence of 1mM GTP; however; no significant decrease should be detected if binding of the i+2 NTP is specific. The results are presented in Figure 2.8. Notably, the rapid rate of incorporation of AMP is not significantly reduced in the presence of 1mM GTP. These results indicate that the NTP binding site is specific for the i+2 nucleotide and strongly suggests that the NTP interacts with the template DNA in

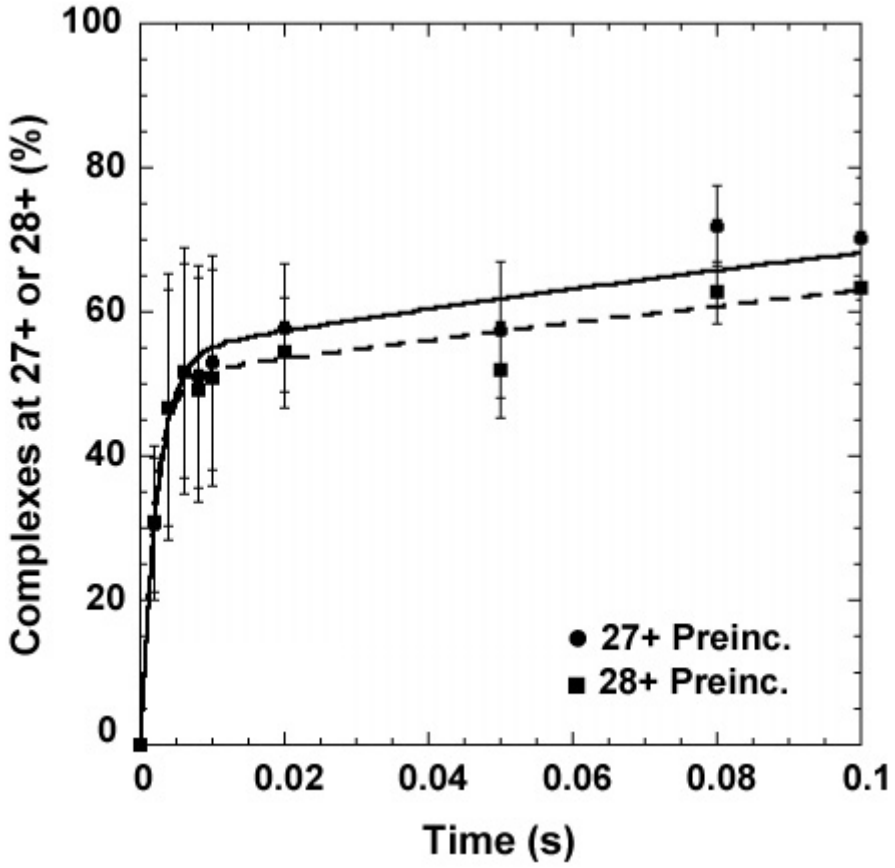


Figure 2.7 Pre-loading is not Dependent on Template Position or Template Identity. To test if template position and pre-incubated NTP identity matter, complexes were "walked" to position +26 by the addition of 10μM CTP and 10μM ATP. After walking, the identities of the n+1 (position +27) and n+2 (position +28) nucleotides after stalling are G and C, respectively. To initiate the reaction 10μM GTP is added to the complexes and the rates of GMP and CMP incorporation are measured. Closed circles are GMP incorporation and closed squares are CMP incorporation. The data were fit to double exponential curves. The solid line is the fit for the 27+ data and the dashed line is the fit for the 28+ data.

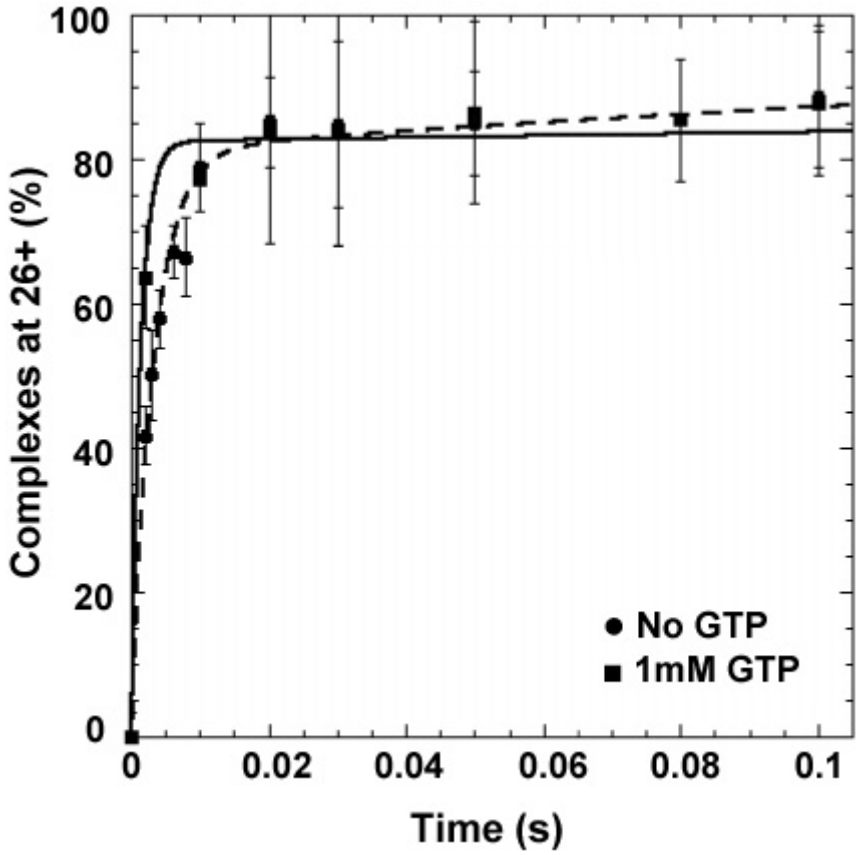


Figure 2.8 ATP Preloading is not Affected by a Nonspecific Inhibitor. The rate of AMP (i+2) incorporation is not significantly affected when SECs are simultaneously pre-incubated with 10μM ATP and 1mM GTP prior to the addition of CTP (Pre-incubation 10μM ATP: open circles; Pre-incubation 10μM ATP & 1mM GTP: closed circles). The data were fit to double exponential curves. The solid line is the fit to the data in which no GTP was added and the dashed line is the fit to the data in which 1mM GTP was added.

the main channel. Similar results were obtained if 1mM GTP was replaced with 1mM sodium triphosphate or 1mM dTTP. These two non-competitive inhibitors also exhibit additional effects on transcription that will be discussed in Chapter 4.

2.3 Kinetics of Multiple Nucleotide Incorporation for Δ-loop RNAP

The results of all the experiments up to this point are consistent with an NTP binding site being present at the i+2 register, which is located in the main channel of the enzyme. As mentioned previously, fork loop 2 has been suggested to be a NTP binding site and located in the β-subunit in the main channel across from the downstream DNA (Holmes and Erie 2003). To determine if fork loop 2 is a component of the NTP binding site, a mutant RNAP was created in which amino acids R542-F545 in the β-subunit have been deleted (Δ-loop RNAP). We then performed the same simultaneous addition and pre-incubation addition experiments with 100μM CTP and 10μM ATP. In both the simultaneous (Figure 2.9, *left*) and pre-incubation (Figure 2.9, *right*) addition experiments, the deletion of several residues comprising fork loop 2 has a modest effect on the rate of CMP incorporation at +25 (i+1). However, unlike wtRNAP, pre-incubation of Δ-loop RNAP with ATP does not result in an increase in the rate of AMP incorporation; instead, deletion of these residues causes a significant reduction in the rate of AMP incorporation at +26 relative wtRNAP (Figure 2.10). Interestingly, the simultaneous addition experiment also exhibits a similar dramatic reduction in AMP incorporation. The observation that this site is specific for the i+2 NTP, suggests that NTPs at the i+2 position can bind to fork loop 2, pair with the template DNA, and facilitate i+2 nucleotide incorporation.

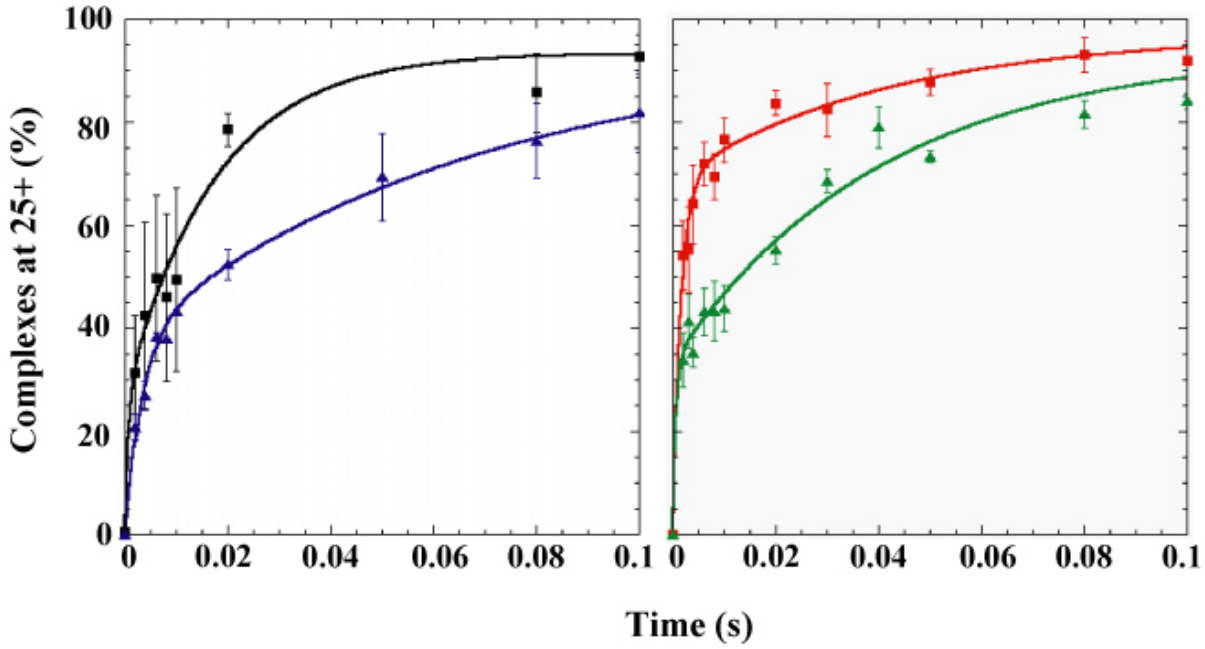


Figure 2.9 Δ-loop RNAP Exhibits Reduced Rates of Incorporation at the i+1 Position. The data presented in the wtRNAP plots are the same as in Figure 2.4A. Percentage of ternary complexes at position +25 under the following conditions: wtRNAP Simultaneous addition 100μM CTP-10μM ATP (black squares, *left*); Δ-loop RNAP Simultaneous addition 100μM CTP-10μM ATP (blue circles, *left*); wt RNAP Pre-incubation of 10μM ATP and initiation with 100μM CTP (red squares, *right*); Δ-loop RNAP Pre-incubation of 10μM ATP and initiation with 100μM CTP (green circles, *right*).

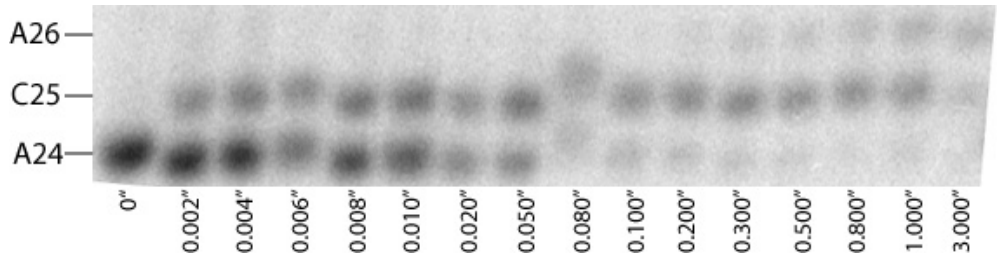


Figure 2.10A

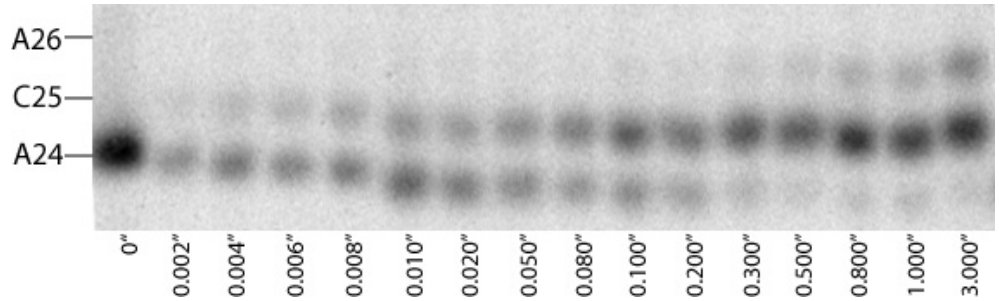


Figure 2.10B

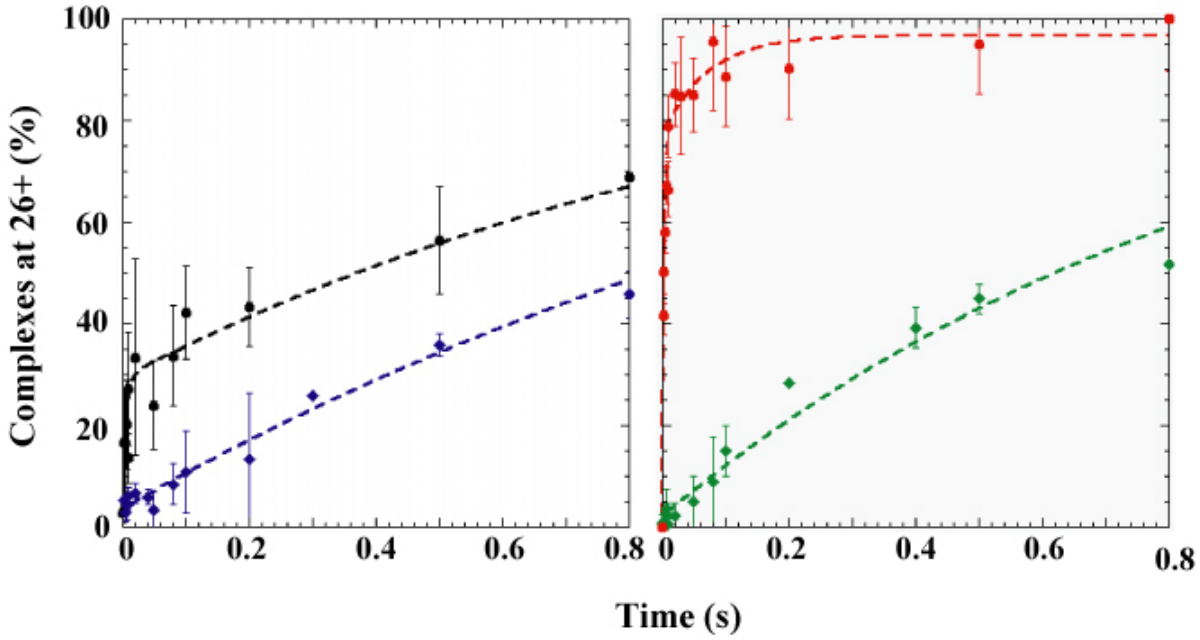


Figure 2.10C

Figure 2.10 Δ-loop RNAP Does Not Preload NTPs.

Figure 2.10 Δ-loop RNAP Does Not Preload NTPs. Δ-loop RNAP has a reduced rate of incorporation for AMP (position +26). The slow phase of both wild-type and Δ-loop RNAP are parallel indicating that they are the same.

A. Representative gel for Δ-loop RNAP simultaneous addition experiments with 100μM CTP and 10μM ATP. The gels show the time course for CMP incorporation at position +25 and AMP incorporation at position +27. The positions are indicated on the left of the gel. The times are presented in seconds and denoted at the bottom of the gel.

B. Same as in **B.** except that the representative gel is for Δ-loop RNAP pre-incubation experiments.

C. The Δ-loop data only fits to a single-exponential. wtRNAP Simultaneous: solid circles and solid line; wtRNAP Pre-incubation: solid squares and solid line; Δ-loop RNAP Simultaneous: open circles and dashed line; Δ-loop RNAP Pre-incubation: open squares and dashed line.

Inspection of the data for the simultaneous addition experiments reveals that wt-RNAP exhibits biphasic kinetics of AMP incorporation with a fast phase followed by a slow phase; whereas the Δ-loop RNAP shows only a single slow phase (Figure 2.10). Interestingly, the rate of the slow phase is not significantly affected but the fast phase is eliminated in the Δ-loop RNAP mutant. These results indicate that NTP incorporation consists of a fast pathway which utilizes the fork loop 2 and the slow pathway which is independent of this loop.

The disappearance of the fast phase in AMP incorporation is consistent with fork loop 2 being an allosteric site. However, this effect may be purely coincidental. The mutated fork loop 2 may lead to the reduction in elongation independent transcription that is not the result of lower a binding affinity for NTPs. To test for this possibility, Δ-loop RNAP was pre-incubated with a higher concentration of ATP (25μM). If the fork loop 2 does not or cannot preload NTPs, then the kinetics of Δ-loop RNAP would remain unchanged regardless of ATP concentration. If NTPs are able to bind to the mutant fork loop 2, then the kinetics for AMP incorporation should become increasingly biphasic with an increasing ATP concentration.

The curve for the pre-incubation of 25μM ATP shows a clear increase in the biphasic nature for AMP incorporation when compared to pre-incubation with 10μM ATP (Figure 2.11). Similar to wt-RNAP, the rate of slow phase is unaffected by the deletion of fork loop 2; however, the complexes in the fast phase incorporate AMP significantly slower than wt-RNAP. These results clearly indicate fork loop 2 is indeed part of an NTP binding site that responsible for the rapid incorporation of NTPs into a nascent RNA chain. Fork loop 2 being responsible for rapid NTP incorporation, and not

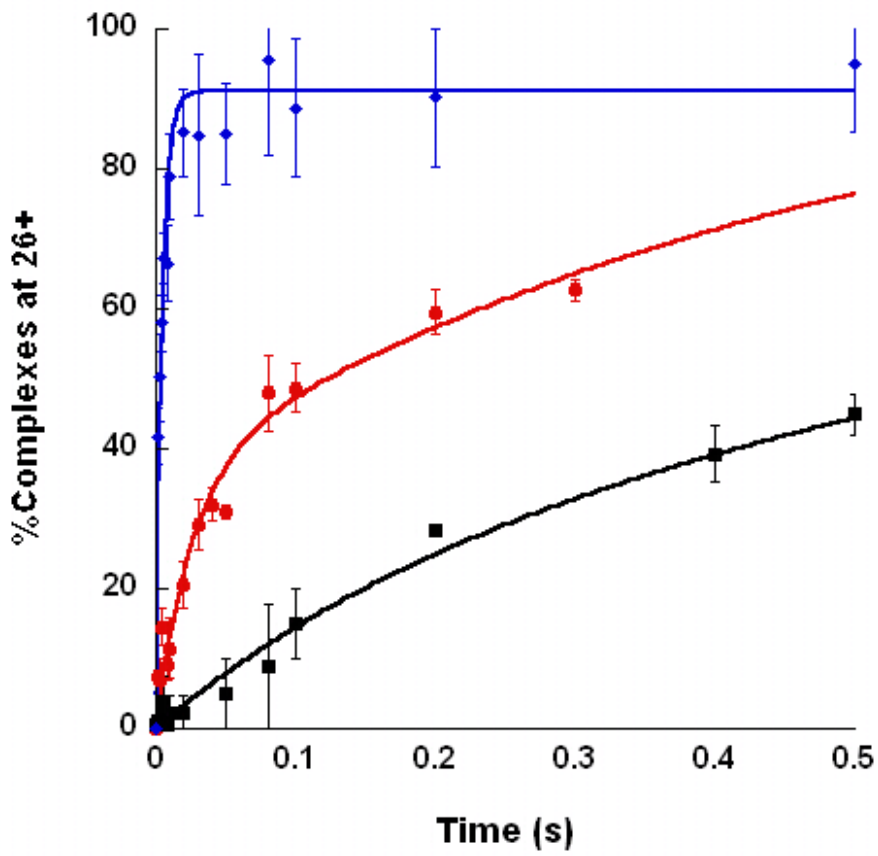


Figure 2.11 Fork loop 2 Exhibits Weakened Binding Affinity for NTPs. Pre-incubating Δ-loop RNAP with 25μM ATP prior to the addition 100μM CTP causes an increase in the rate of AMP when compared to complexes pre-incubated 10μM ATP. The 25μM AMP incorporation curve (red) is clearly biphasic and means that RNAP can exist in a fast and slow phase. Δ-loop RNAP-10μM ATP is in black; Δ-loop RNAP-25μM ATP is in red; wt-RNAP-10μM ATP is in blue.

slow incorporation, is consistent with there being two pathways in which NTPs are able to enter the catalytic site; one in which NTPs utilize fork loop 2 and one in which NTPs do not. The possibility of multiple NTP entry pathways into the catalytic site will be discussed later in this chapter.

2.4 The NTP Bound to Fork Loop 2 is Able to Load Directly Into the Catalytic Site

Taken together, these results indicate that the i+2 NTP can bind specifically and with high affinity to fork loop 2, pair with the template DNA, and facilitate the incorporation of both the i+1 and i+2 nucleotides. It is clear that the effect of the i+2 NTP on the incorporation of the i+1 nucleotide must be allosteric in nature; however, binding of the i+2 NTP in the main channel could increase the rate of i+2 nucleotide incorporation by an allosteric mechanism and/or by shuttling the NTP bound in the main channel into the catalytic site (Holmes and Erie 2003; Gong, Zhang *et al*. 2005).

To examine these possibilities, we performed experiments in which non-radiolabeled SECs were pre-incubated with 10μM [^{32}P]-ATP and the reaction was initiated with a mixture of 100μM CTP and 0μM, 100μM, 1mM, or 5mM unlabeled ATP. The same experiment was performed with the simultaneous addition of 100μM CTP, 10μM [^{32}P]-ATP, and the corresponding concentration of unlabeled ATP (See Experimental Procedures and Figure 2.12). The relative extents of [^{32}P]-AMP incorporation at position +26 were determined in both the pre-incubation and simultaneous addition experiments. If the NTP bound to fork loop 2 is shuttled into the catalytic site and incorporated, we would not expect to observe as much reduction in radioactively labeled SECs when the unlabeled SECs are pre-incubated with 10μM [^{32}P]-

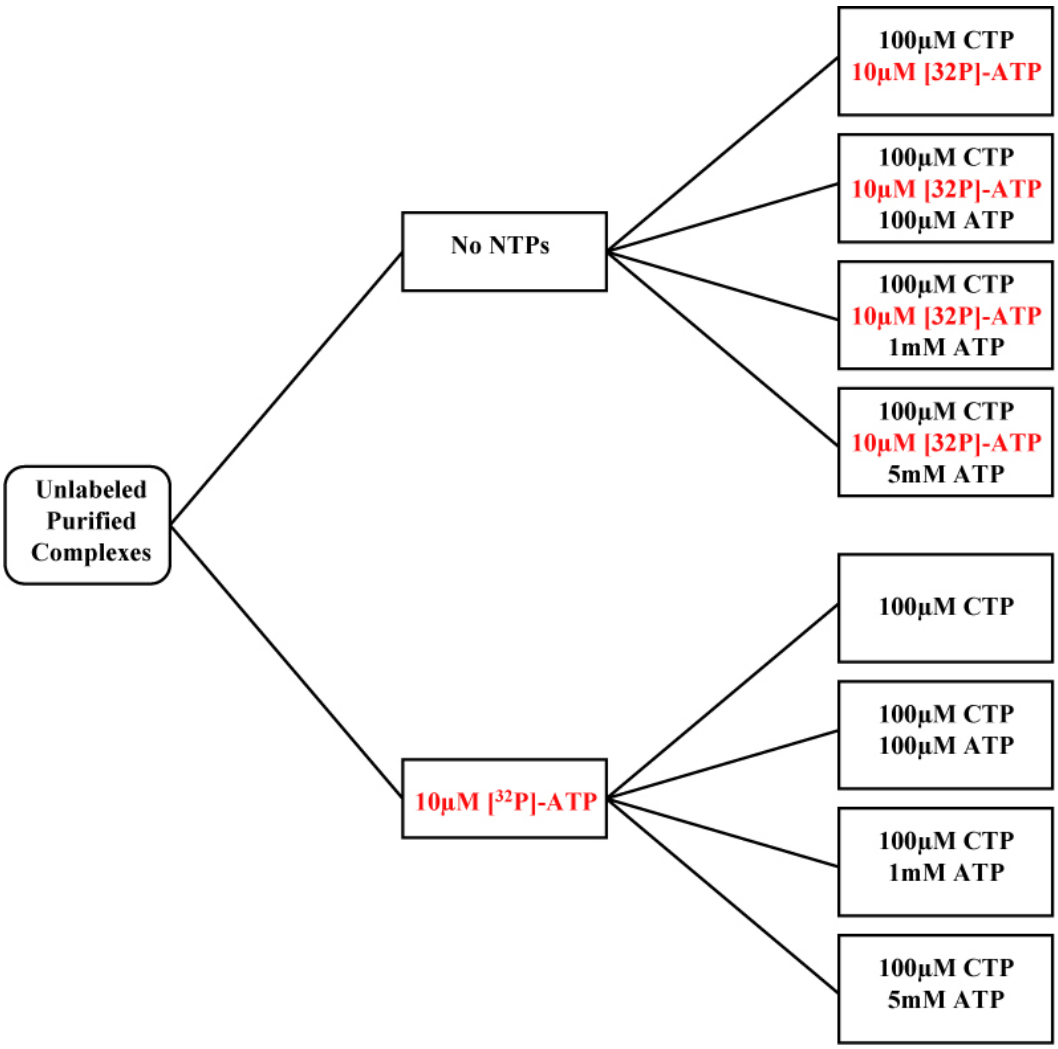


Figure 2.12 Preloaded NTP Incorporation Experimental Setup. Unlabeled purified SECs are split into two main groups. One reaction group is pre-incubated with 10µM α-[^{32}P]-ATP the other is not pre-incubated with any NTPs. The reactions are further aliquoted and the indicated NTP mixture is added to the SECs prior to quenching with 0.5M. The unlabeled ATP concentrations are a 0-fold, 10-fold, 100-fold, and 500-fold dilution of α-[^{32}P]-ATP, respectively. All reactions were performed in the presence of 10mM Mg^{+2}.

AMP and then unlabeled ATP added versus when the labeled and unlabeled ATP are added simultaneously. Conversely, if NTP binding is solely facilitating incorporation via an allosteric mechanism, the extent of labeling of the SECs should not depend on the order of addition of labeled and unlabeled ATP.

Inspection of the data in Figure 2.13a and Figure 2.13b reveal that the simultaneous addition of NTPs (grey bars) follows the expected reduction in [^{32}P]-AMP incorporation as a function of the dilution factor (10X=12.2%±1.1%, 100X=1.33%±0.48%, 500X=Not Detectable (ND)). In the pre-incubation experiments (cross-hatched bars), there is a significantly higher amount of [^{32}P]-AMP incorporation relative to the simultaneous addition of NTPs (10X=16.1%±4.3%, 100X=6.66%±1.13%, 500X=4.37%±0.47%). Figure 2.14 shows the multiplicative factor by which the experimental percent incorporated deviates from the expected percentage incorporation as determined by the dilution (Protection Factor). It is clear from Figure 2.14 that pre-incubated complexes incorporate a significantly higher amount of [^{32}P]-AMP than would be predicted by dilution alone. If the amount of [^{32}P]-AMP were the same as what would be expected based on the dilution, then this ratio would be unity. For the simultaneous addition experiment, the data show a Protection Factor of ~1 in all cases in which it is applicable (10X=1.22, 100X=1.34, 500X=Not Applicable (NA). This result means that the amount of [^{32}P]-AMP is based solely on the dilution factor. For the pre-incubation experiments, the data are strikingly different. The amount of [^{32}P]-AMP incorporation is higher than what would be expected by the dilution factor alone (10X=1.61, 100X=6.66, 500X=21.9). This effect is especially apparent at the higher concentrations of unlabeled ATP.

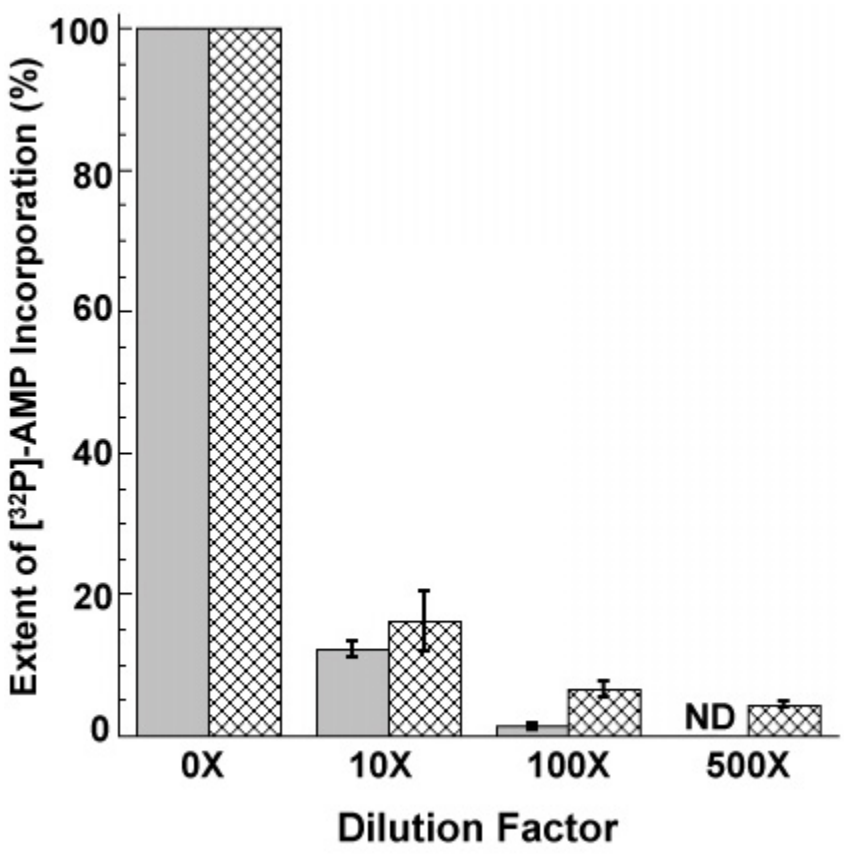


Figure 2.13A

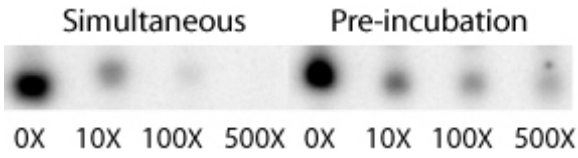


Figure 2.13B

Figure 2.13 NTPs Can Enter the Catalytic Site of RNAP Via the Main Channel.

Figure 2.13 NTPs Can Enter the Catalytic Site of RNAP via the Main Channel.

A. Bar graph of the relative extent of [^{32}P]-AMP incorporation as a function of the dilution factor. The amount of AMP incorporation in each lane is normalized to the lane with the highest amount of incorporation. Solid bars, simultaneous addition of all NTPs; cross-hatched bars, pre-incubation of 10μM [^{32}P]-AMP. ND means "Not Detectable".

B. Representative gel the [^{32}P]-AMP pre-incubation experiment. The dilution factor is indicated below each lane.

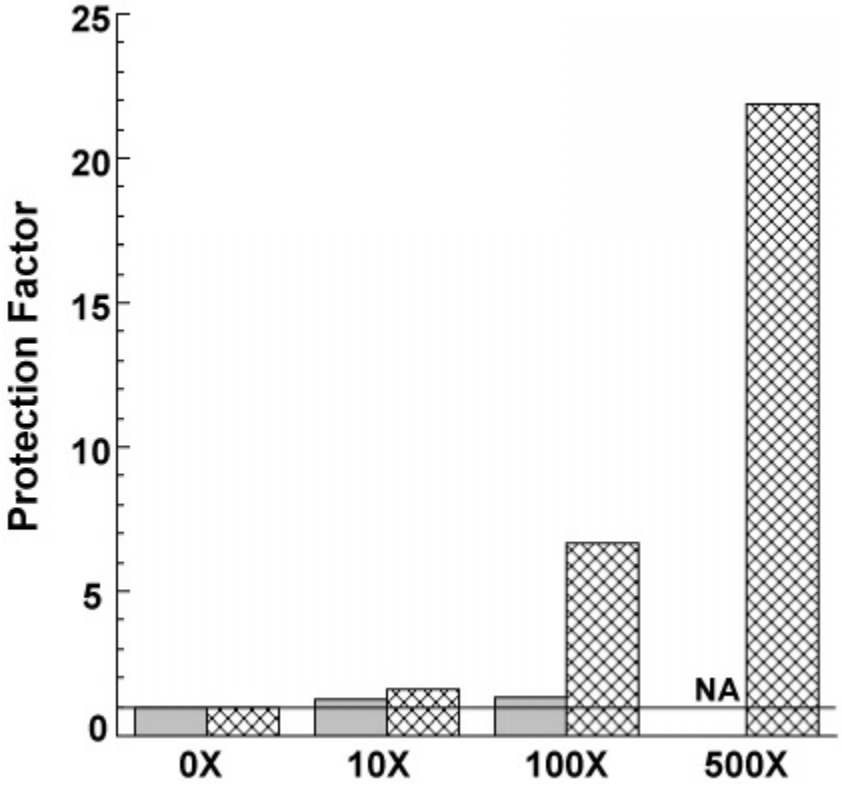


Figure 2.14 Deviation from the Expected Amount of [^{32}P]-AMP Incorporation.

Protection Factor is defined as (% actual [^{32}P]-AMP)/(% expected [^{32}P]-AMP). Solid bars, simultaneous addition of all NTPs; cross-hatched bars, pre-incubation of 10μM [^{32}P]-AMP. The line at Protection Factor=1 is if the data did not deviate from the expected amount of incorporation. NA means "Not Applicable"

2.4.1 Kinetic Simulations of Pre-loading Experiments

Our results are consistent with the idea that NTPs are able to enter the catalytic site via the main channel; however, if the main channel were the exclusive route for NTP entry, we would expect a significantly higher percentage of complexes to incorporate [^{32}P]-AMP than we observe. Instead, we observe an initial decrease in [^{32}P]-AMP incorporation at low unlabeled ATP concentrations followed by an approach to a near zero-order dependence on unlabeled ATP at higher concentrations. To explain these observations, we performed kinetic simulations using KinTekSim (Barshop, Wrenn *et al*. 1983) in which we tested several different mechanisms (Figure 2.15a, 2.15b, 2.16a, and 2.16b). In the course of performing the kinetic simulations we found that a binding constant of 20μM for the allosteric site and a k_{cat}, which is the rate of product formation after entry into the catalytic site, was not rate limiting, provide for the best fits to the data The value of 20μM is in excellent agreement with the experimentally determined binding constant of 22μM for the allosteric site (Holmes and Erie 2003). The other kinetic parameters used in the simulations are presented in Table 2.1. Interestingly, from our simulations it was found that k_{fast} having a value of ~20-25X greater than k_{slow} gave the best fits. The experimentally determined values of k_{fast} and k_{slow} show that k_{fast} is ~18X faster than k_{slow} (Holmes and Erie 2003). This ratio is remarkably close to the simulated ratio.

The first mechanism we tested assumed a rapid exchange in which the [^{32}P]-ATP is replaced by an unlabeled ATP, and the allosteric site was the exclusive route into the catalytic site (Figure 2.15a). This mechanism will decrease in the amount of [^{32}P]-AMP with increasing ATP concentration and does not give a zero-order dependence regardless

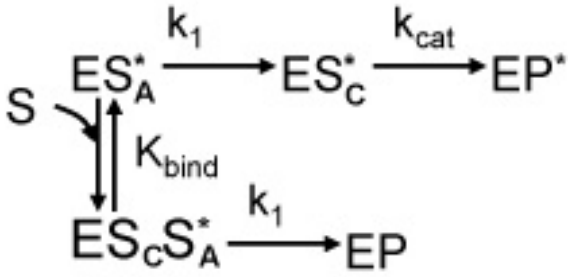


Figure 2.15A Simple Competition Mechanism

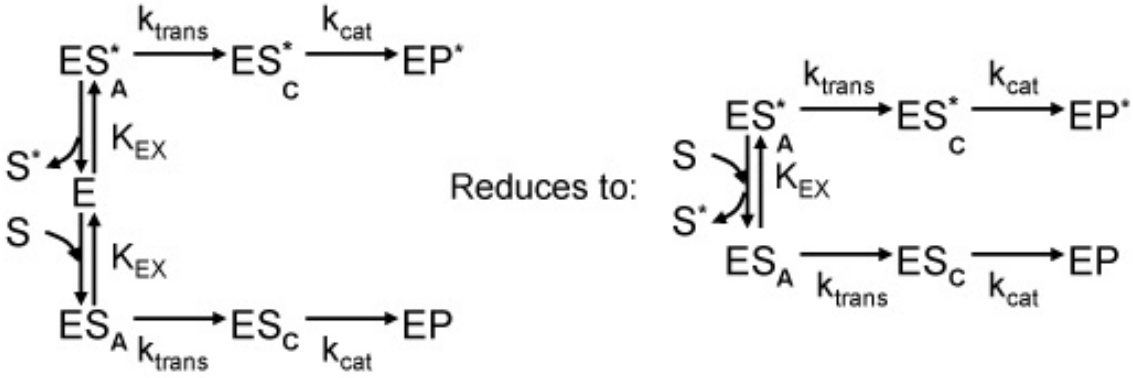


Figure 2.15B Rapid Exchange Mechanism

Figure 2.15 Mechanisms in which Simulations were not in Agreement with Experimental Data. E stands for enzyme; S represents the NTP substrate and P is the RNA product. The astericks denotes the substrate is radiolabeled. The subscript A or C denotes the location of the bound substrate; allosteric site or catalytic site, respectively. K_{bind} is the dissociation constant where $K_{bind}=k_{off}/k_{on}$ where k_{off} is the off rate and k_{on} is the on rate. k_{cat} is the catalytic rate, $K_{EX}=K_{bind}$ is the exchange constant and is the same thing as the dissociation constantk_{trans} is the rate of transfer of the allosteric NTP into the catalytic site. k_1 is the rate of incorporation

of the ATP concentration. The next mechanism that we tested involved a simple competitive pathway mechanism in which there is a rapid equilibrium for unlabeled substrate binding to the catalytic site and both the labeled and unlabeled ATP compete for entry into the catalytic site via two different pathways (Figure 2.15b). This simple mechanism gives the same extents of [^{32}P]-AMP incorporation as the rapid exchange mechanism. It is not possible for these two mechanisms to show zero-order dependence at any ATP concentration because, as the concentration of unlabeled ATP increases, then the path that incorporates ATP will become the dominant path. In order to show zero-order dependence for ATP concentration, there must be a saturation of a binding site accompanied with a competing path that exists after the binding of unlabeled ATP.

The simplest mechanism that results in [^{32}P]-AMP incorporation with a zero-order dependence only at high ATP concentrations is a competitive pathway mechanism similar to the one previously described in which one pathway is utilized by preloaded [^{32}P]-ATP and the second pathway that utilizes ATP/[^{32}P]-ATP from bulk solution and becomes saturated at high unlabeled ATP concentrations (Figure 2.16A and Figure 2.16B). When the second path is not saturated at low unlabeled ATP concentrations, the path preloaded with [^{32}P]-ATP will be dominant. Furthermore, the amount of [^{32}P]-AMP incorporation will be strongly affected by low unlabeled ATP concentrations because the second path is the preferred path; however, once the second pathway is saturated, high concentrations of ATP will no longer have a significant effect on [^{32}P]-AMP incorporation and incorporation reduces to a simple competitive mechanism that is independent of NTP concentrations. In the case of *E. coli* RNAP, ~4.37%±0.47% of complexes incorporate [^{32}P]-ATP from the main channel regardless of the ATP

Binding Const.	Kinetic Const.	Value
K_{bind}	k_{off}	$1x10^{8}$ s^{-1}
	k_{on}	$5x10^{4}$ s^{-1}
N/A	k_{fast}	100 s^{-1}
	k_{slow}	4 s^{-1}
N/A	k_{cat}	1000 s^{-1}
	k_{app}	10 s^{-1}

Table 2.1 Kinetic Parameters Used in Simulations

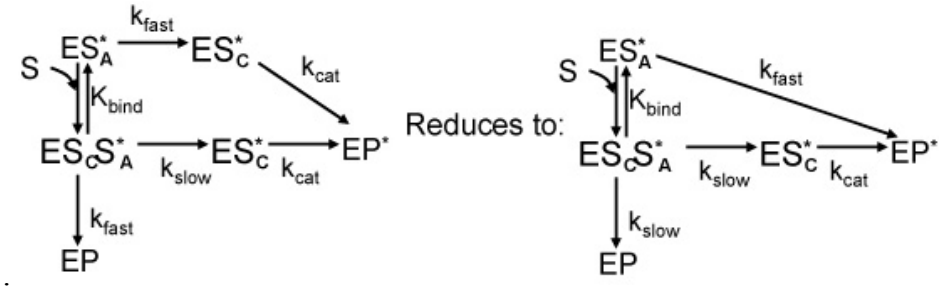


Figure 2.16A

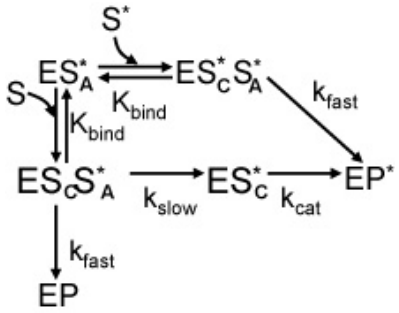


Figure 2.16B

Figure 2.16 Mechanisms in Which Simulations Agreed with Experimental Data. E stands for enzyme; S represents the NTP substrate and P is the RNA product. The astericks denotes the substrate is radiolabeled. The subscript A or C denotes the location of the bound substrate; allosteric site or catalytic site, respectively. K_{bind} is the dissociation constant where $K_{bind}=k_{off}/k_{on}$ where k_{off} is the off rate and k_{on} is the on rate. k_{cat} is the catalytic rate, K_{EX} is the exchange constant and is the same thing as the dissociation constant. k_{trans} is the rate of transfer of the allosteric NTP into the catalytic site. k_{fast} is the rate of NTP entry into catalytic site via the primary path (secondary channel) and k_{slow} is the rate of entry NTP entry into the catalytic site through the secondary path (main channel).

concentration. The extents for the various mechanisms that were simulated are presented in Figure 2.17 and are compared to the experimentally determined extents presented in Figure 2.13. It is clear from the results that the saturated competitive path mechanism is in remarkably good agreement with our observations; in fact, the simulated data is within error of the experimental results. The agreement is especially noteworthy for the data at high unlabeled ATP concentrations. It should be noted that, from our simulations, we determined that there are two kinetically identical mechanisms that follow the saturating competitive pathway mechanism.

Figure 2.16A presents a mechanism in which the NTP bound in the allosteric site undergoes a fast incorporation in the absence of a NTP bound in the catalytic site. However, in the presence of unlabeled NTP, there is a competition between the allosteric NTP being incorporating and an unlabeled NTP binding to the catalytic site. If an unlabeled NTP binds to the catalytic site, then the rate of the allosteric NTP incorporation drops significantly. While certainly possible, this mechanism seems unlikely because of the reduction in the rate of incorporation of the allosteric NTP is dependent on the presence of a NTP in the catalytic site.

Another, more attractive, mechanism that explains our data is one in which there is competitive binding of labeled and unlabeled NTP to the catalytic site. Whichever NTP binds enters the catalytic site first, is rapidly incorporated (k_{fast}). This path must be independent of the allosteric site because that site is already bound with an NTP. Upon binding to the catalytic site, there is a certain population that transfers the NTP from the allosteric site into the catalytic site via a slow pathway (k_{slow}). The slow path occurs regardless if a labeled or unlabeled NTP binds to the catalytic site, but, due to our

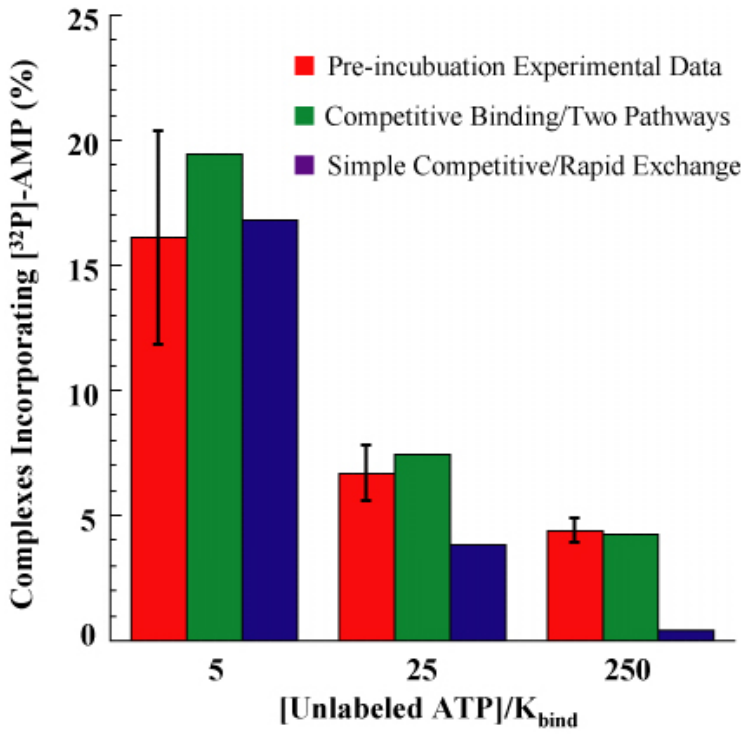


Figure 2.17 Comparison of Simulations with Experimental Pre-incubation Data. Actual data (red) show a very similar profile for the competitive binding/two pathway mechanisms (green) (Figure 2.16A and 2.16B). In contrast, there is a poor correlation between the actual data and the mechanisms utilizing either a rapid exchange at the allosteric site or a simple competition between pathways (blue) (Figure 2.15A and 2.15B).

experimental design we cannot see this exchange when a labeled NTP bind to the catalytic site because it is exchanging a labeled NTP for another labeled NTP.

It is clear from these simulations that the NTP bound to fork loop 2 and NTPs in solution are able to compete for the catalytic site, which is consistent with there being two pathways for NTP entry into the catalytic site. One pathway utilizes fork loop 2 and the other pathway is independent and is most likely the secondary channel.

2.5 Structural Model of NTP Entry into the Catalytic Site

Comparing several key structural elements of RNAP known to be involved in translocation and/or catalysis from yeast RNAPII crystal structures (Figure 2.18) reveals that they can exist in several conformations. One of the most notable structural shifts is that of the fork loop 2. This loop has been seen in three different conformations. Interestingly, the shift in the fork loop 2 conformation is also accompanied by a change in the DNA structure (Fig. 2.19). When fork loop 2 is in the "open" position, a concomitant opening of the DNA by one base pair is observed. The bubble opening would allow for an NTP to bind to fork loop 2 and base pair with the DNA or, conversely, NTP binding could induce a shift in fork loop 2, leading to DNA bubble opening. Other major conformational shifts are observed for the bridge helix and trigger loop (Fig. 2.20). Similar structural shifts in these same elements have also been observed in the recently published *T. thermophilus* elongation complex. In both RNAPII (Wang, Bushnell *et al*. 2006) and bacterial RNAP(Vassylyev, Vassylyeva *et al*. 2007), the trigger loop, which is located below the bridge helix (Figure 2.18, 2.19, 2.21), has been observed "closed" on the active site and makes direct contact with the bound catalytic NTP. When the trigger

Figure 2.18 Main Channel Structural Elements Involved in NTP Preloading. Multiple conformations of fork loop 2 and the trigger loop in *S. cerevisiae* RNAPII. The DNA template strand (grey), non-template strand (dark blue), and RNA (red) are from 1Y77. The catalytic site (beige) and bound GTP (purple) are from 2E2H. The bridge helix is colored orange (1Y1V). The trigger loop is able to exist in both an "open" conformation (green, 1Y1V) and a "closed" conformation (magenta, 2E2H). Fork loop 2 has been seen in three distinct conformations: "open" (yellow, 2E2I), partially "closed" (green, 1Y1V), and "closed" (light blue, 1Y77). The spheres represent the residues deleted in the Δ-loop RNAP mutant.

Figure 2.19A

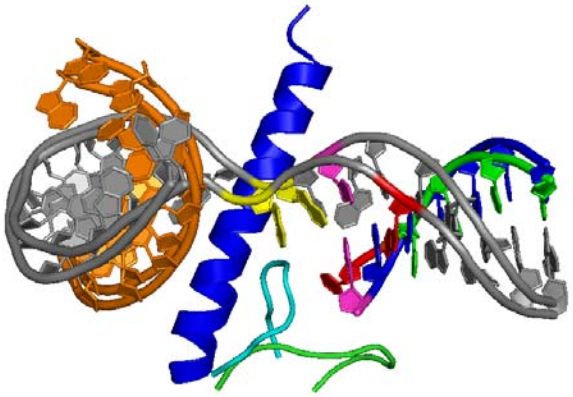


Figure 2.19B

Figure 2.19 Movement of Fork Loop 2 Coincides with the Opening of the Transcription Bubble in *S. cerevisiae* RNAPII.

Figure 2.19 Movement of Fork Loop 2 Coincides with the Opening of the Transcription Bubble in *S. cerevisiae* RNAPII.

A. Protein structures and the corresponding non-template DNA strand belonging to the same crystal structure are colored the same (green, 2E2I; dark blue, 1Y77). The i+2 position of both template strands is designated in yellow and does not have a base pairing partner. The base pair that is broken with the corresponding fork loop 2 movement is designated in magenta (base paired) and red (non-base paired). The template strands are colored grey and the RNAs are colored orange.

B. The same structure presented in **A.** rotated 90º out of the plane of the page.

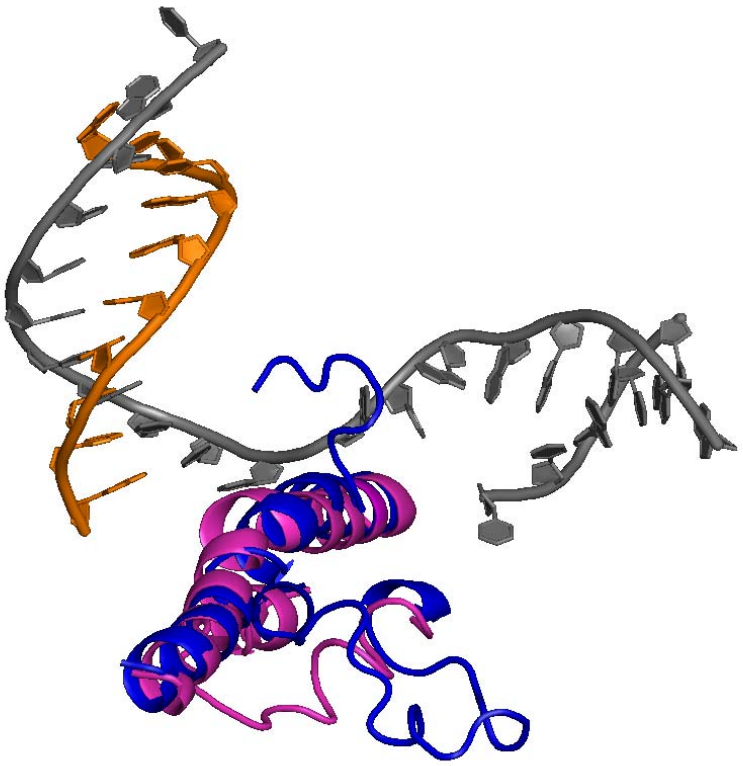


Figure 2.20 Concerted Movement of Bridge-helix and Trigger Loop in *S. cerevisiae* RNAPII. When the trigger loop makes a conformational shift from the open position (1Y1V, purple) to a closed position (2E2H, blue) there is a concomitant unwinding of the bridge helix (structure colors correspond to the trigger loop colors). Both DNA strands are colored grey and the RNA is in blue.

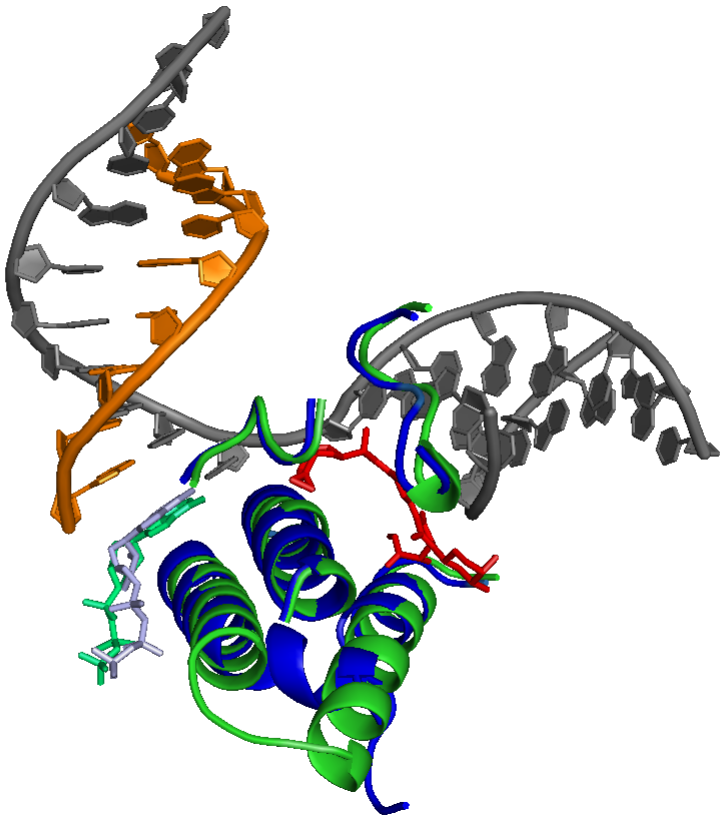


Figure 2.21 Trigger Loop Movements in *T. thermophilus* RNA Polymerase. The DNA strands are colored in grey and the RNA is in blue and are from PDB 2PPD (Vassylyev, Vassylyeva *et al*. 2007). The green trigger loop is from PDB 2PPD and the magenta trigger helix is from PDB 2O5J (Vassylyev, Vassylyeva *et al*. 2007). The nucleotides seen in the catalytic site are from the same structures that have the respective color only slightly paled. The streptolydigin is pale yellow and is from 2PPD. When an NTP is bound in the catalytic site, the trigger loop assumes a "closed" conformation and is involved in catalytis; however, streptolydigin locks the trigger loop in an "open" conformation. Streptolydigin allows for the isolation of the "open" conformation which is most likely transient.

loop is in the "closed" conformation, fork loop 2 becomes disordered in RNAPII and is seen in different conformations in bacterial crystal structures (Wang, Bushnell *et al*. 2006; Vassylyev, Vassylyeva *et al*. 2007).

2.5.1 Model for Allosteric Site Usage in Transcription Elongation

Previous structural and kinetic studies have proposed several different types of Brownian ratchet models (Holmes and Erie 2003; Temiakov, Patlan *et al*. 2004; Yin and Steitz 2004; Bar-Nahum, Epshtein *et al*. 2005). Our results are consistent with a Brownian ratchet model. To that end, we propose a model in which fork loop 2 acts primarily as an allosteric site, but the bound NTP can be transferred into the catalytic site from the main channel (Figure 2.22). In this model, the allosteric site is only utilized when RNAP is in the pre-translocated state. Binding of the allosteric NTP stabilizes the opening of the downstream DNA bubble by one base (Step 1). Opening of the DNA bubble destabilizes trigger loop interactions with the catalytic site allowing the trigger loop to adopt an "open" conformation which opens the secondary channel allowing for future NTP entry (Figure 2.23A and 2.23B). The trigger loop interacts with bridge helix, fork loop 2, and the allosteric NTP which, together, act as a latch that temporarily locks the trigger loop open (Step 3). Step 3 is consistent with previous work which shows that the antibiotic strepolydigin, which irreversibly locks the trigger loop in the "open" conformation, does not affect translocation or substrate binding to the catalytic site (Temiakov, Zenkin *et al*. 2005; Vassylyev, Vassylyeva *et al*. 2007). The β-pincer movement, in conjunction with the opened downstream DNA, would allow for rapid translocation of RNAP along the DNA.

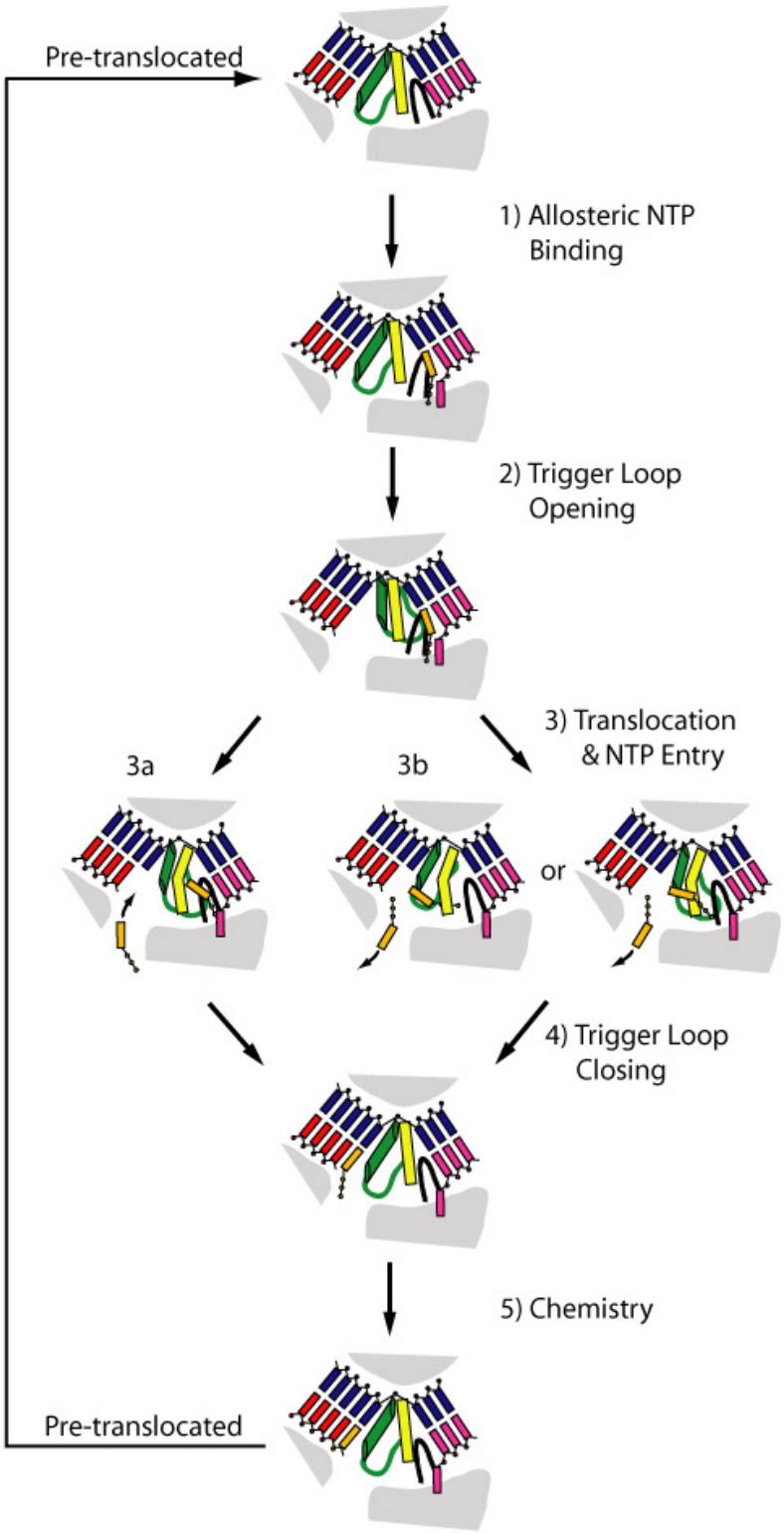


Figure 2.22 Cartoon Model of Allosteric Site Usage During Elongation.

Figure 2.22 Cartoon Model of Allosteric Site Usage During Elongation. The template DNA strand is colored blue and the non-template strand is colored in pink. The RNA chain is colored in red. The next incoming NTP is colored in orange. Fork loop 2 is in black, the bridge helix is yellow and the trigger loop is colored in green. The helical region of the trigger loop that makes contact with the catalytic site is represented by the cylinder attached to the green loop region. The enzyme starts off the NTP addition cycle in the pre-translocated state. Step 1) Opening of the downstream DNA bubble and destabilization of trigger loop by binding of the allosteric NTP to fork loop 2. Step 2) The trigger loop is "locked open" by interactions with allosteric site. Step 3a) Translocation and NTP entry via the secondary channel. Step 3b) Translocation and entry of the allosteric NTP into the catalytic site via the main channel; active displacement of catalytic NTP. Step 4) The trigger loop closes down on the NTP in the catalytic site. Step 5) Chemistry occurs, thereby leaving RNAP in a pre-translocated state.

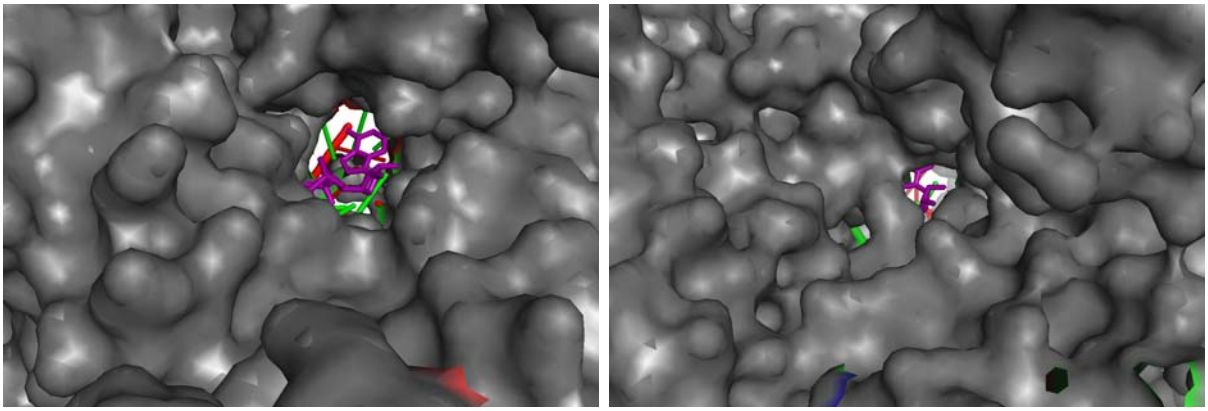


Figure 2.23A **Figure 2.23B**

Figure 2.23 Opening of the Trigger Loop to the Leads to the Opening of the Secondary Channel. The opened secondary channel allows an NTP (purple) access to the active site whereas a closed secondary channel (right) would restrict access.

A. The opened channel is from PDB 2PPB.

B. The closed channel is from PDB 2O5J.

Due to translocation, fork loop 2 looses its affinity for the allosteric NTP. The loss of affinity for the bound NTP acts like a latch release for the trigger loop. Concurrently, translocation frees the catalytic site for NTP entry. NTP entry into the catalytic site can occur via two pathways. The first path, which is most likely the primary path, is through the secondary channel (Step 3a); however, in a small fraction of complexes, the allosteric NTP is able to enter the catalytic site using the secondary pathway. The secondary pathway is via the main channel. We propose a model in which NTPs are either handed-off to the trigger loop and shuttled under the bridge helix or, alternatively, transferred into the catalytic site by going over the top of the bridge helix and actively displacing any NTP previously bound in the catalytic site (Step 3b). Trigger loop closing provides several catalytically essential residues to the active site and induces catalysis and the cycle repeats (Step 5).

Importantly, NTP binding to fork loop 2 is not essential for productive synthesis, but synthesis is significantly slower. If fork loop 2 is not utilized, DNA bubble opening would not be stabilized which would keep the equilibrium shifted towards the pre-translocated state. Furthermore, the trigger loop would not be latched open thereby spending more time in the "closed" conformation which blocks NTP entry via the secondary channel.

2.5.2 Model for NTP Entry Through the Main Channel

The question remains of how the NTP that is bound to fork loop 2 is able to transit directly into the catalytic site. There are two possibilities, both of which merit examination. The first model we propose is a shuttling mechanism in which the NTP

bound to fork loop 2 is transferred via a hand-off mechanism to the trigger loop. As mentioned previously, the allosteric site would lose its affinity for the bound NTP upon translocation because the site is no longer specific to that NTP. Instead of diffusing back into bulk solution, the NTP could make and maintain interactions with the trigger loop and subsequently become exclusively associated with the trigger loop. The trigger loop then transfers the NTP under the bridge helix and into the catalytic site. The hand-off may not be the primary pathway for NTP entry into the catalytic site but occurs only when the i+2 NTP concentration in solution is low because at higher i+2 NTP concentrations this pathway will be have to compete with NTPs entering via the secondary channel.

There are several lines of evidence that argue for this possibility. Recent crystal structures of both eukaryotic and prokaryotic RNAP have shown that the trigger loop makes extensive contacts with the catalytic nucleotide; however, the binding affinity to the active site appears to be primarily determined by catalytic site residues and not the trigger loop; therefore, it is reasonable that the trigger loop could bind an NTP whether it is interacting with the active site or in the main channel. Inspection of the streptolydigin bound *T. thermophilus* elongation complex shows the trigger loop stabilized in an "open" conformation (Vassylyev, Vassylyeva *et al*. 2007). Interestingly, streptolydigin makes numerous contacts simultaneously with both fork loop 2 and the trigger loop (Figure 2.24A, 2.24B, and 2.24C). This observation is consistent with the potent elongation inhibitor streptolydigin acting as a transition state analog for nucleotide transfer that locks fork loop 2 and the trigger loop in the conformation that is responsible for this transition.

Another, and potentially more attractive, model involves the NTP being loaded

Figure 2.24A

Figure 2.24B

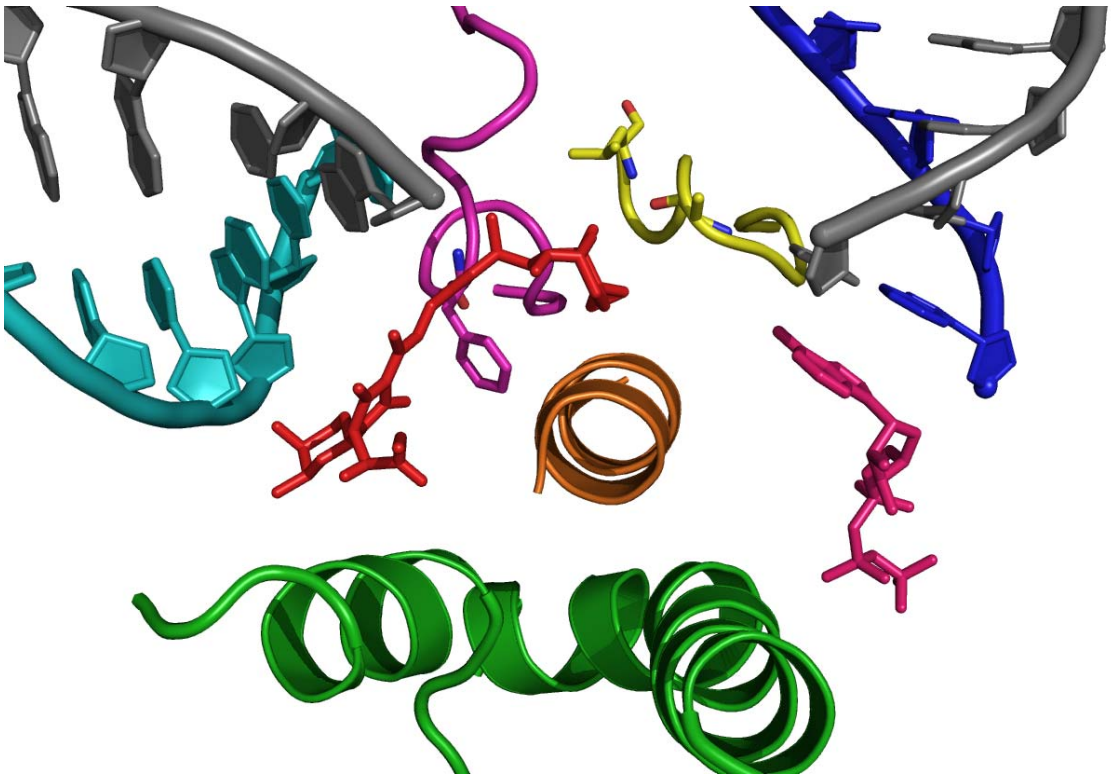


Figure 2.24C

Figure 2.24 *T. thermophilus* Elongation Complex with Streptolydigin and AMP-CPP Bound.

Figure 2.24 *T. thermophilus* Elongation Complex with Streptolydigin and AMP-CPP Bound. The color scheme for both **A.** and **B.** is as follows: The DNA template strand is colored grey; the non-template strand is in blue; RNA is red. The bridge helix is colored orange while the trigger loop is in green. Fork loop 2 is cyan and the spheres represent the residues removed in Δ-loop RNAP. AMP-CPP (non-incorporatable ATP analog) is purple and streptolydigin is magenta.

A. Streptolydigin makes simultaneous contacts with fork loop 2, bridge helix, and the trigger loop. Streptolydigin looks like it is a transition state analog for NTP hand-off from fork loop 2/DNA to the trigger loop. Both the streptolol ring (seen interacting with the bridge helix) and acetamide moiety (seen interacting with the trigger loop) are multi-cyclic and contain several oxygen and nitrogen atoms that can easily hydrogen bond with various residues. These oxygen and nitrogen atoms could easily mimic the nucleoside base and the phosphate groups. Importantly, the trigger loop can be seen extended past the bridge helix. There is an 11-a.a. unstructured region in the trigger helix. It is not unreasonable for this unstructured region to help in the transfer of the NTP from the allosteric site to the trigger loop.

B. The same structure seen in **A.** but rotated 180º about the Z-axis.

C. The same view seen in **B.**, but zoomed in. Residues that make interactions with the streptolol moiety are rendered in stick form. The color scheme is the same as as in **A.** and **B.** with the following exceptions: Fork loop 2 is magenta, βD-loopII is yellow, and AMP-CPP is pink. The i+1 and i+2 template positions were removed for clarity.

into the catalytic site from over the top of the bridge helix. A similar model for NTP entry has been suggested previously (Gong, Zhang *et al*. 2005). In this model, upon the loss of binding affinity for the allosteric site due to translocation of the DNA, the NTP can enter the catalytic site by either passively diffusing or being actively transported by the template DNA due to base pair and stacking interactions.

There are several lines of evidence that strongly argue for this model. Structurally, streptolydigin is reminiscent of a nucleotide triphosphate. The overall structure consists of a streptolol moiety and a tetramic acid moiety connected together by a hydrophobic linker (Figure 2.25); the streptolol moiety resembles a nucleobase; it is multicyclic and hydrophobic, although it is not planar. The tetramic acid moiety is significantly more polar in nature and contains several oxygen atoms that are able to form hydrogen bonds, which is reminiscent of the phosphate groups of a NTP. Additionally, streptolydigin has approximately the same dimensions as an NTP (~17Å for Stl vs ~15Å for NTPs). Taken together, streptolydigin may be acting as a nucleotide mimic.

Additionally, in the crystal structure of the *T. thermophilus* elongation complex the streptolol moiety is placed on top of the bridge helix and makes several contacts with the bridge helix, fork loop 2, and βD-loopII, which forms a "lid" over the catalytic site (Figure 2.24A, 2.24B, and 2.24C, yellow). The streptolol group (Figure 2.25) appears to be partially overlapping with the top of the catalytic site which could be reminiscent of an NTP entering the catalytic site. There is biochemical data that support this model. In a recent paper by Toulokhonov, *et al.*(2007), the authors show βD-loopII, which is directly connected to fork loop 2, is able to adopt multiple conformations which are dependent on whether the RNAP is in an active synthesis state or a paused state. Based on their results,

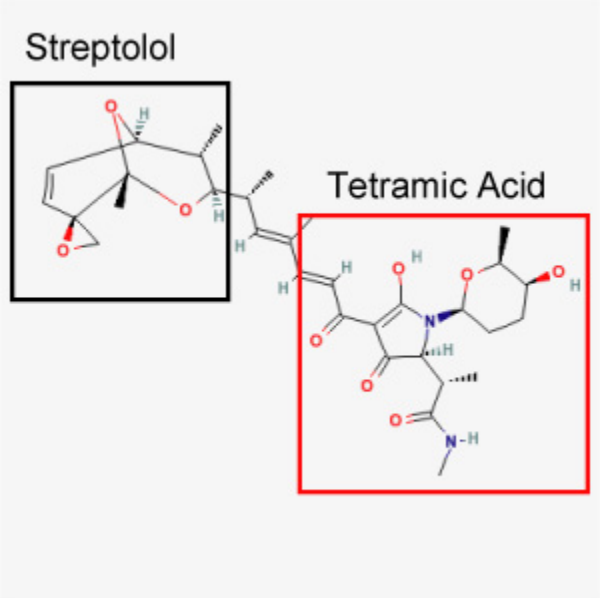


Figure 2.25 Structure of Streptolydigin. The streptolol moiety is boxed in black and the tetramic acid moiety is boxed in red.

they propose a structural model in which movement of βD-loopII between the active and paused conformation lead to movements in fork loop 2. In effect, βD-loopII allosterically controls fork loop 2 (Toulokhonov, Zhang *et al*. 2007). If βD-loopII is able to coordinate the movement of fork loop 2, then it is not unreasonable for the opposite to be true; during normal elongation, an NTP binding to fork loop 2 could allosterically move βD-loopII into a more open conformation which would no longer sterically restricts NTP traversal across the bridge helix.

There are several advantages to this model over the trigger loop shuttling model. The primary advantage is that the allosteric NTP never loses contact with the template DNA, thereby NTP specificity is maintained. We have noted in previous experiments that the slower, unactivated, state has a higher fidelity, which is consistent with the NTP never losing contact with the template DNA (Erie, Hajiseyedjavadi *et al*. 1993; Holmes, Santangelo *et al*. 2006). In the trigger loop model, the NTP loses contacts with the template DNA, therefore it is possible that the NTP could unbind and be replaced by a different NTP or even a dNTP; this could be problematic if an incorrect NTP were loaded directly into the catalytic site. Another possible advantage for this model is that loading over the bridge helix could more easily allow for active displacement of the catalytic NTP. In the bridge helix model, the NTP would enter via the main channel from the "top" of the catalytic site (Figure 2.24A, 2.24B, and 2.24C) which could more easily "push" any NTP/dNTP in the catalytic site out the secondary channel, which is located in opposition to the top of the catalytic site. While it is possible that trigger loop shuttling could result in the same active displacement, it does not make as much sense, structurally. If the trigger loop were to deliver an NTP into the catalytic site that was already occupied

by a second NTP, that NTP would not have an easy escape route due to the fact that the closed trigger loop conformation occludes the secondary channel. Furthermore, the shuttled NTP would have to be able to break NTP base pair and stacking interactions with the DNA and multiple protein-nucleic acid interactions in order to displace the NTP already present in the catalytic site.

2.6 Experimental Procedures

2.6.1 Sources of Protein and DNA

Wild-type His_6-tagged RNAP was purified from log phase *E. coli* strain RL916 (gift of R. Landick) using the protocol described previously (Burgess and Jendrisak 1975; Uptain and Chamberlin 1997). All DNA templates used were derived from pDE13 using a QuickChange Site-Directed Mutagenesis Kit (Stratagene) (Erie, Hajiseyedjavadi *et al*. 1993). All mutant templates were purified by a Qiagen Miniprep Kit (Qiagen) or by phenol-chloroform extraction and the sequences confirmed by the DNA sequencing facility at UNC. All DNA transcription templates were amplified by PCR. The 540 bp PCR product is biotinylated on one end and contains the λP_R promoter. The sequence of the first 30nt of DE13-A27g transcript is as follows:

DE13-A27g: pppAUGUAGUAAGGAGGUUGUAUGGAAC<u>C</u>AGCGC

The first cytidine to be incorporated is located at +25 (underlined).

His_6-tagged β-Δ(R542-F545) RNAP was made by site-directed mutagenesis using the QuickChange Site-Directed Mutagenesis Kit on the pRL-706 plasmid. pRL-706 was a gift from Richard Ebright and was made as described previously *(*Severinov, Mooney *et al*. 1997). Expression was carried out in the *E. coli* strain TOM100 (gift of T. Santangelo)

by induction with 1mM IPTG with cells in late log-phase (OD= ~0.8-1.0) and purified as previously described (Santangelo, Mooney *et al*. 2003) with the following modifications: 1) The sephacryl column purification was omitted and 2) purified RNAP was dialyzed into TGED+1M NaCl buffer [10mM Tris-HCl (pH 8.0), 5% glycerol, 1mM EDTA, 1mM DTT, 1M NaCl] prior to dialysis into Storage Buffer [20mM Tris-HCl (pH 8.0), 50% glycerol, 1mM EDTA, 1mM DTT, 100mM NaCl].

2.6.2 In Vitro Transcription Reactions

To form OPCs, 80nM RNAP was combined with 80nM template DNA bound to streptavidin magnetic beads in 1x transcription buffer [30mM HEPES-NaOH (pH 8.0), 10mM Mg^{+2} glutamate, 200mM K^{+} glutamate, 25μg/mL BSA, and 1mM DTT] and incubated at 37°C for 10 min. After incubation, SECs stalled at position +24 were formed by the addition of a mixture of 20μM ATP, 15μM UTP, 16μM GTP, and 4μM [α-^{32}P]-GTP (800Ci/mmol) to the OPCs at room temperature. The reaction was allowed to proceed for 1 min in the case of wtRNAP and 2 min in the case of β-Δ(R542-F545) RNAP. SECs were purified from free NTPs by holding the reaction tube next to a magnet and washing seven to ten times with cold 1x transcription buffer. The complexes were then resuspended in 1x transcription buffer to a final SEC concentration of ~14nM. The purified complexes were kept on ice until used in rapid quench kinetic experiments. In some experiments, the complexes were "walked" further downstream template by the addition of 10μM CTP, or a 10μM CTP and 10μM ATP mixture. Complexes were allowed to react for 1 minute and then purified by washing as previously described unless otherwise noted. All rapid quench experiments were conducted on a Kintek Rapid

Quench 3 device (Kintek Corporation, Austin, TX) at room temperature (~23°C). For each reaction time point, ~20μL of purified SECs were loaded into a sample injection loop and ~20μL of the required NTP solution was loaded on the other sample injection loop. After the reaction was allowed to proceed for the desired time the reactions were quenched with 0.5M EDTA-NaOH (pH 8.0). To make sure that the time between time points while conducting the experiments did not affect the results, the order in which the time points were collected was randomized. In addition, to ensure that complexes were still active throughout the time course of the experiment, a small aliquot of purified SECs were reacted with 1mM of all four NTPs and allowed to extend to full length transcript. The EDTA was removed from the samples by centrifuging the samples and holding them next to a strong magnet as described above. Complexes were resuspended in 100% formamide, boiled for two minutes, and loaded onto an 8M urea, 20% polyacrylamide gel.

The allosteric binding experiments were performed as follows: Formation and purification of OPCs and SECs was performed as previously described except 20μM ATP, 15μM UTP, 20μM GTP was used to initiate transcription. SECs were split into two sample groups. To one group, designated the pre-incubation sample group, [α-^{32}P]-ATP (800Ci/mmol) was added to a final concentration of 10μM. SECs were further aliquoted into four samples of equal volume. The second sample group, designated the Simultaneous sample group, was divided into three equal volume aliquots. For the Pre-incubation sample group, reactions were initiated with the following, respective, NTP mixtures: 100μM CTP; 100μM CTP, 100μM ATP; 100μM CTP, 1mM ATP, and 100 μM CTP, 5mM ATP. The Simultaneous sample group reactions were initiated with the following NTP mixtures: 100μM CTP, [α-^{32}P]-ATP (800Ci/mmol); 100μM CTP,

[α-^{32}P]-ATP (800Ci/mmol), 100μM ATP; 100μM CTP, [α-^{32}P]-ATP (800Ci/mmol), 1mM ATP, and 100μM CTP, [α-^{32}P]-ATP (800Ci/mmol), 5mM ATP. All reactions were quenched with 0.5M EDTA-NaOH (pH 8.0). Samples were processed as described above. All NTP concentrations are reported as final working concentrations. Each experiment was conducted between 2 and 5 times.

2.6.3 Data Quantification and Normalization of Transcription Rate Data

The amount of radiation in each sample lane was measured by a Molecular Dynamics Phosphoimager and analyzed using ImageQuant v5.2 software. The percentage of complexes at each position was calculated by dividing the amount of radiation at each template position by the total amount of radiation in bands ≥+24 for each lane. The data were normalized to 100% by dividing the amount of radiation at each time point by the highest percentage of complexes that incorporated past position +25 or +26. All experiments were conducted a minimum of three times per NTP concentration. For the allosteric binding experiments, the data were normalized to 100% by dividing the amount of radiation at position +26 in each lane by the sample with the highest amount of radiation.

2.6.4 Kinetic Fits to the Data

Fits of the data in each plot were performed using KaleidaGraph v4.01. The data were fit to both single and double-exponentials. Since the data is clearly biphasic in most cases, it was not surprising to that fitting to a double exponential function gave significantly better results. This is consistent with previous work on the kinetics of

nucleotide incorporation (Foster, Holmes *et al*. 2001; Holmes and Erie 2003). Error bars in the rapid quench experiments are reported as standard error, while error bars in the allosteric binding experiments are reported as standard deviation.

In an effort to determine a mechanism to our explain the allosteric loading experiments results, we using KinTekSim (Barshop, Wrenn *et al*. 1983) to simulate several different kinetic mechanisms. Specifically, we were interested in simulating the extents of [^{32}P]-AMP incorporation as a function of unlabeled ATP concentration. We were able to reduce the field of possible mechanisms down to three kinetically identical mechanisms. There are likely several steps between NTP binding and the actual catalysis event, but we cannot resolve them. The steps presented are the rate limiting steps in each path.

BIBLIOGRAPHY

Bar-Nahum, G., V. Epshtein, et al. (2005). "A Ratchet Mechanism of Transcription Elongation and Its Control." Cell **120**(2): 183-193.

Barshop, B. A., R. F. Wrenn, et al. (1983). "Analysis of Numerical Methods for Computer SImulation of Kinetic Processes: Development of KINSIM--A Flexible, Portable System." Anal. Biochem. (130): 134-145.

Batada, N. N., K. D. Westover, et al. (2004). "Diffusion of Nucleoside Triphosphates and Role of the Entry Site to the RNA Polymerase II Active Center." Proc. Nat. Acad. Sci. **101**(50): 17361-17364.

Burgess, R. R. and J. J. Jendrisak (1975). "Procedure for the Rapid, Large-Scale Purification of *Escherichia coli* DNA-Dependent RNA Polymerase Involving Polymin P Precipitation and DNA-Cellulose Chromatography." Biochem. **14**(21): 4634 - 4638.

Davenport, R. J., G. J. L. Wuite, et al. (2000). "Single-Molecule Study of Transcriptional Pausing and Arrest by E. coli RNA Polymerase." Science **287**(5462): 2497-2500.

Erie, D. A. (2002). "The Many Conformational States Of RNA Polymerase Elongation Complexes And Their Roles In The Regulation Of Transcription." Biochim. Biophys. Acta **1577**(2): 224-239.

Erie, D. A., O. Hajiseyedjavadi, et al. (1993). "Multiple RNA Polymerase Conformations and GreA: Control of the Fidelity of Transcription." Science **262**(5135): 867-873.

Foster, J. E., S. F. Holmes, et al. (2001). "Allosteric Binding of Nucleoside Triphosphates to RNA Polymerase Regulates Transcription Elongation." Cell **106**: 243-252.

Gong, X. Q., C. Zhang, et al. (2005). "Dynamic Error Correction and Regulation of Downstream Bubble Opening by Human RNA Polymerase II." Mol. Cell **18**(4): 461-470.

Herbert, K. M., A. L. Porta, et al. (2006). "Sequence-Resolved Detection of Pausing by Single RNA Polymerase Molecules." Cell **125**(6): 1083-1094.

Holmes, S. F. and D. A. Erie (2003). "Downstream DNA Sequence Effects on Transcription Elongation: Allosteric Binding of Nucleoside Triphosphates Facilitates Translocation Via a Ratchet Motion." J. Bio. Chem. **278**(37): 35597-35608.

Holmes, S. F., T. J. Santangelo, et al. (2006). "Kinetic Investigation of Escherichia coli RNA Polymerase Mutants That Influence Nucleotide Discrimination and Transcription Fidelity." J. Bio. Chem. **281**(27): 18677-18683.

Leipe, D. D., Y. I. Wolf, et al. (2002). "Classification and Evolution of P-loop GTPases and Related ATPases." J. Mol. Bio. **317**(1): 41-72.

Matsuzaki, H., G. A. Kassavetis, et al. (1994). "Analysis of RNA Chain Elongation and Termination by *Saccharomyces cerevisiae* RNA Polymerase III." J. Mol. Bio. **235**(4): 1173-1192.

Mukhopadhyay, J., E. Sineva, et al. (2004). "Antibacterial Peptide Microcin J25 Inhibits Transcription by Binding within and Obstructing the RNA Polymerase Secondary Channel." Mol. Cell **14**: 739-751.

Nedialkov, Y. A., X. Q. Gong, et al. (2003). "NTP-driven Translocation by Human RNA Polymerase II." J. Bio. Chem. **278**(20): 18303-18312.

Palangat, M., C. T. Hittinger, et al. (2004). "Downstream DNA Selectively Affects a Paused Conformation of Human RNA Polymerase II." J. Mol. Bio. **341**(2): 429-442.

Reynolds, R., R. M. Bermbdez-Cruz, et al. (1992). "Parameters Affecting Transcription Termination by *Escherichia coli* RNA Polymerase: I. Analysis of 13 Rho-independent Terminators." J. Mol. Bio. **224**(1): 31-51.

Santangelo, T. J., R. A. Mooney, et al. (2003). "RNA Polymerase Mutations that Impair Conversion to a Termination-Resistant Complex by Q Antiterminator Proteins." Gene. Dev. **17**(10): 1281-1292.

Severinov, K., R. Mooney, et al. (1997). "Tethering of the Large Subunits of *Escherichia coli* RNA Polymerase." J. Bio. Chem. **272**(39): 24137-24140.

Telesnitsky, A. and M. J. Chamberlin (1989). "Terminator-Distal Sequences Determine the in Vitro Efficiency of the Early Terminators of Bacteriophages T3 and T7." Biochem. **28**(12): 5210-5218.

Temiakov, D., V. Patlan, et al. (2004). "Structural Basis for Substrate Selection by T7 RNA Polymerase." Cell **116**(3): 181-391.

Temiakov, D., N. Zenkin, et al. (2005). "Structural Basis of Transcription Inhibition by Antibiotic Streptolydigin." Mol. Cell **19**(5): 655-666.

Tolić-Nørrelykke, S. F., A. M. Engh, et al. (2004). "Diversity in the Rates of Transcript Elongation by Single RNA Polymerase Molecules." J. Bio. Chem. **279**(5): 3292-3299.

Toulokhonov, I., J. Zhang, et al. (2007). "A Central Role of the RNA Polymerase Trigger Loop in Active-Site Rearrangement during Transcriptional Pausing." Cell **227**(3): 406-419.

Uptain, S. M. and M. J. Chamberlin (1997). "*Escherichia coli* RNA Polymerase Terminates Transcription Efficiently at Rho-Independent Terminators on Single-Stranded DNA Templates." Proc. Nat. Acad. Sci. **94**(25): 13548-13553.

Vassylyev, D. G., M. N. Vassylyeva, et al. (2007). "Structural Basis for Substrate Loading in Bacterial RNA Polymerase." Nature **448**(7150): 163-168.

Via, A., F. Ferrè, et al. (2000). "Three-dimensional View of the Surface Motif Associated with the P-loop Structure: *Cis* and *Trans* Cases of Convergent Evolution." J. Mol. Bio. **303**(4): 455-465.

Walker, J. E., M. Saraste, et al. (1982). "Distantly Related Sequences in the α- and β-Subunits of ATP Synthase, Myosin, Kinases and Other ATP-Requiring Enzymes and a Common Nucleotide Binding Fold." EMBOJ **1**(8): 945-951.

Wang, D., D. A. Bushnell, et al. (2006). "Structural Basis of Transcription: Role of the Trigger Loop in Substrate Specificity and Catalysis." Cell **127**(5): 941-954.

Westover, K. D., D. A. Bushnell, et al. (2004). "Structural Basis of Transcription: Nucleotide Selection by Rotation in the RNA Polymerase II Active Center." Cell **119**(4): 481-489.

Yin, H., I. Artsimovitch, et al. (1999). "Nonequilibrium Mechanism of Transcription Termination from Observations of Single RNA Polymerase Molecules." Proc. Nat. Acad. Sci. **96**(23): 13124.

Yin, Y. W. and T. A. Steitz (2004). "The Structural Mechanism of Translocation and Helicase Activity in T7 RNA Polymerase." Cell **116**(3): 393-404.

Zhang, C., K. L. Zobeck, et al. (2005). "Human RNA Polymerase II Elongation in Slow Motion: Role of the TFIIF RAP74 α1 Helix in Nucleoside Triphosphate-Driven Translocation." Mol. Cell. Bio. **25**(9): 3583-3595.

Zhang, G., E. A. Campbell, et al. (1999). "Crystal Structure of *Thermus aquaticus* Core RNA Polymerase at 3.3 Å Resolution." Cell **98**(7): 811-824.

CHAPTER 3:

A TOTALLY CONSERVED WALKER B MOTIF IS A COMPONENT OF THE MAIN CHANNEL ALLOSTERIC SITE

3.1 Introduction

NTP binding to enzymes is usually mediated through a divalent magnesium cation in the form of a NTP·Mg^{+2} complex. The positively charged magnesium coordinates with β and γ phosphates of the NTP. In order for the NTP·Mg^{+2} complex to efficiently bind to an enzyme, nature has evolved two core amino acid sequences, termed Walker motif A and Walker motif B, that are generally well conserved and were first discovered in ATP-binding enzymes (Walker, Saraste *et al*. 1982). The Walker A motif has a consensus sequence of -G X X X X G K-, where X is a non-conserved residue. This motif is well conserved in both ATP and GTP binding proteins (Leipe, Wolf *et al*. 2002). The Walker B motif has a consensus sequence of -X_h(2-4) D X_h(2-8) G-, where X_h is a hydrophobic residue and has been implicated in binding of the NTP·Mg^{+2} complex (Fry, Kuby *et al*. 1986; Black and Hruby 1992; Leipe, Wolf *et al*. 2002). This motif functions by having the invariant aspartic acid residue binding the Mg^{+2} through a water molecule, while the backbone of the invariant glycine interacts with the γ-phosphate of a NTP by hydrogen bonding via its amide hydrogen (Pai, A.Petsko *et al*. 1990; Bourne, Sanders *et al*. 1991).

Previous work based on the kinetics of single-nucleotide incorporation, along with structural analysis, has lead to the proposal that several structures located in the

main channel of RNAP comprise an NTP binding site that acts as an allosteric site for NTP incorporation (Foster, Holmes *et al*. 2001; Holmes and Erie 2003). The suggested site has several very important features that are consistent with it being an NTP binding site. The site consists of a flexible loop structure (fork loop 2, aka βD-loop I) that contains a number of glycines in both prokaryotic and eukaryotic RNAPs. In fact, the sequence is almost identical to a known Walker A motif found in rabbit muscle adenylate kinase (Kuby, Palmieri *et al*. 1984; Fry, Kuby *et al*. 1986). Furthermore, fork loop 2 is surrounded by a β-sheet on one side and several α-helices on the other. These features are indicative of "P-loops" (Walker, Saraste *et al*. 1982; Via, Ferrè *et al*. 2000; Leipe, Wolf *et al*. 2002). Additionally, a totally conserved Walker B motif is located in close proximity to fork loop 2 (Walker, Saraste *et al*. 1982) (Figure 3.1A and 3.1B). Inspection of the crystal structure of the *T. thermophilus* holoenzyme reveals that a magnesium ion is chelated to the invariant aspartic acid residue, which is a hallmark of a Walker B motif and is required for proper NTP binding (Vassylyev, Sekine *et al*. 2002).

Several models suggest that RNAP uses the main channel in elongation to bind NTPs, coded for by the downstream DNA, and facilitates translocation of RNAP along the DNA. Because there is no direct evidence that the main channel is used during elongation, we were interested in testing this possibility. We previously identified the fork loop 2 as a putative NTP binding site located in the main channel across from the downstream DNA in the β-subunit (Holmes and Erie 2003). Furthermore, in a previous chapter, we presented strong evidence that suggests that fork loop 2, which consists of a Walker A motif, has a strong effect on the kinetics of nucleotide incorporation. To further extend these results, we were interested in determining if the Walker B motif

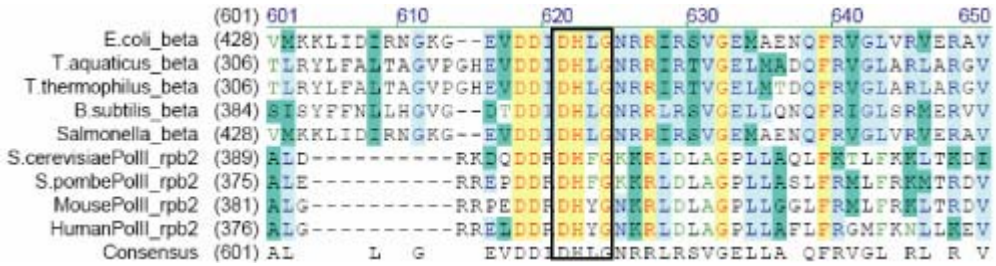


Figure 3.1A

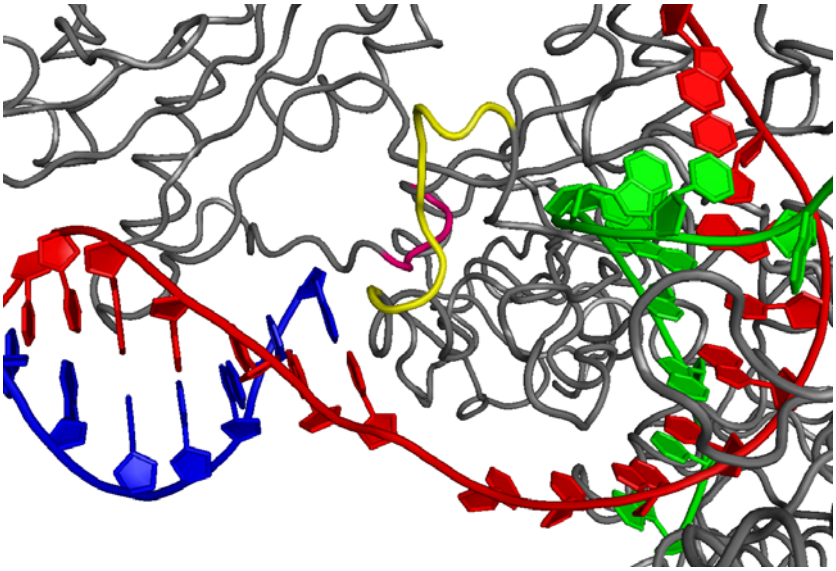


Figure 3.1B

Figure 3.1 The Primary Sequence and Location of the Walker B Motif.

Figure 3.1 The Primary Sequence and Location of the Walker B Motif.

A. Sequence alignment of the β-subunit region that contains the Walker B motif (Holmes and Erie 2003). The name of the organism from which the sequence is from is listed on the left with the number of the starting residue in parenthesis. The consensus sequence is listed on the bottom. The Walker B motif is boxed in black. Totally conserved residues are highlighted in yellow while residues that are conserved solely in prokaryotes are highlighted in blue. Well conserved residues are highlighted in green.

B. Structural view of the Walker B motif. The structure is of the *T. thermophilus* RNAP elongation complex (PDB 1PPD). All subunits except the β-subunit (grey) have been removed for clarity. Fork loop 2 is colored yellow and immediately behind it is the Walker B motif (mageneta). The DNA template strand (red) and non-template strand (blue) are in close proximity to both fork loop 2 and the Walker B motif. The RNA is shown in green.

located near fork loop 2 was important for proper function of the allosteric site. Using transient-state kinetics, we investigated the role of the amino acids that comprised the Walker B motif in NTP binding and incorporation. Specifically, we characterized the incorporation of multiple nucleotides into a growing RNA chain using mutant RNAP in which a double mutant of the Walker motif was introduced (β-(D446A/G449A)) (Walker RNAP).

In order to test if the Walker B motif is important for NTP addition, we measured the rate of CMP and AMP incorporation at template positions +25 (i+1) and +26 (i+2), respectively, under two conditions: one in which CTP and ATP were simultaneously added and one in which the SECs were incubated with ATP prior to initiating the reaction with the addition of CTP (Figure 3.2). These experiments are nearly identical to the ones carried out on Δ-loop RNAP (Figure 2.3).

3.2 Walker RNAP Shows Altered Kinetics

SECs are formed by using a template in which the first cytidine to be incorporated is at position +25 and initiating transcription with only ATP, UTP and GTP to form complexes stalled at template position +24. These complexes are purified from NTPs, then subsequent NTPs are added back and the rates of incorporation measured (See Experimental Procedures). Because we were particularly interested in examining the rate of AMP incorporation, we used a high concentration of CTP (100μM) and a low concentration of ATP (10μM) so that CMP incorporation would not be rate limiting for AMP incorporation.

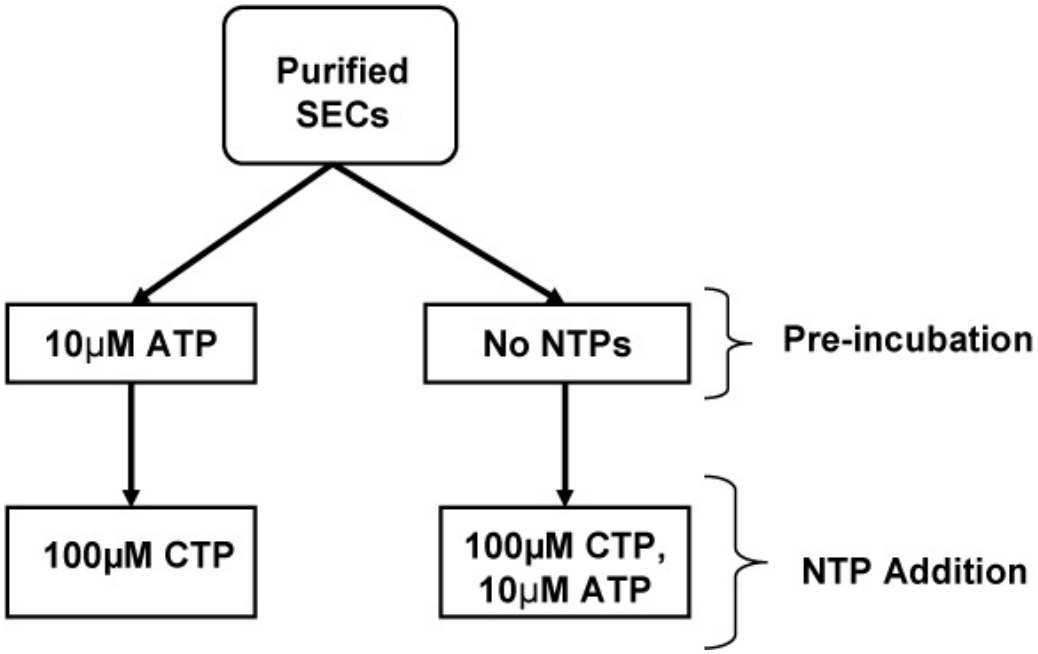


Figure 3.2 Experimental Designs to Test for NTP Preloading. Purified SECs are split into two groups; one in which 10μM ATP is added to the complexes prior to the addition of 100μM CTP via rapid quench, and one in which 100μM CTP and 10μM ATP are added simultaneously by rapid quench techniques.

3.2.1 Simultaneous Addition of NTPs Shows an Increase in the Rate of the Unactivated State

The data for the simultaneous addition of CTP and ATP are shown in Figure 3.3A and Figure 3.3B. The data reveal that replacement of the conserved aspartic acid and glycine residues has no significant effect on the rate of CMP incorporation at position +25 (i+1) compared to wtRNAP. The data are different for AMP incorporation. Inspection of the data for the simultaneous addition experiments for wt-RNAP reveals that it exhibits biphasic kinetics of AMP incorporation, which is characterized by a fast phase followed by a slow phase. Comparing the rate of AMP incorporation for wt-RNAP to that of Walker-RNAP, a marked enhancement is observed; specifically the rate of the slow phase is increased by a factor of 4.2±1.0. Interestingly, the extent of complexes in the fast or slow state is unchanged between Walker-RNAP and wt-RNAP. Additionally, as can be seen in Figure 3.3B, the rate of the fast phase is also unaffected (Figure 3.3B). Therefore, this mutation only affects complexes in the unactivated state by increasing their rates of incorporation.

3.2.2 Pre-incubation with ATP Shows a Decrease in the Extent of Complexes in the Activated State

To perform the experiments, Walker-RNAP SECs stalled out at position +24 on the DE13-A27g template were pre-incubated with 10μM ATP prior to initiating the reaction with 100μM CTP (Figure 3.2). Inspection of the data in Figure 3.4 shows that there is no appreciable effect on the rate of CMP incorporation at the concentration used in these experiments. An observed decrease in either the extent or rate of CMP incorporation may be observed if a lower concentration of CTP were used (i.e. 10μM).

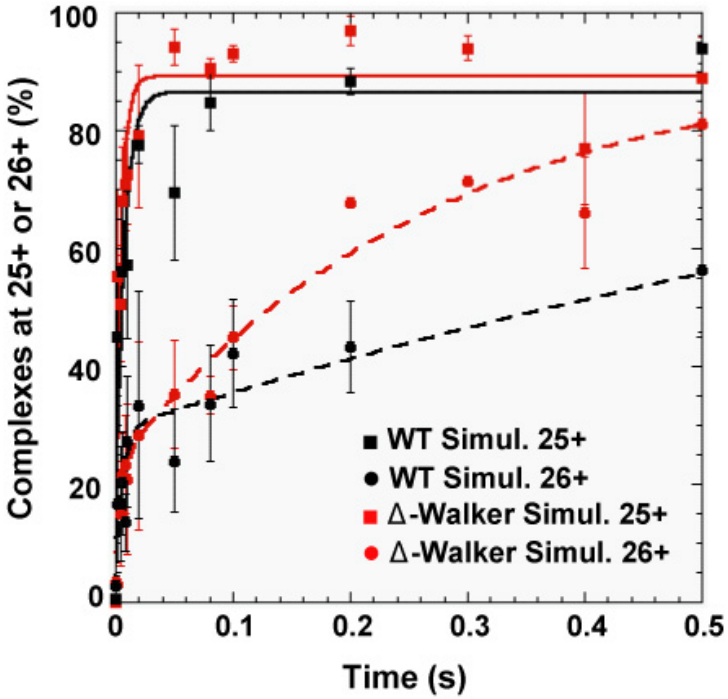


Figure 3.3A

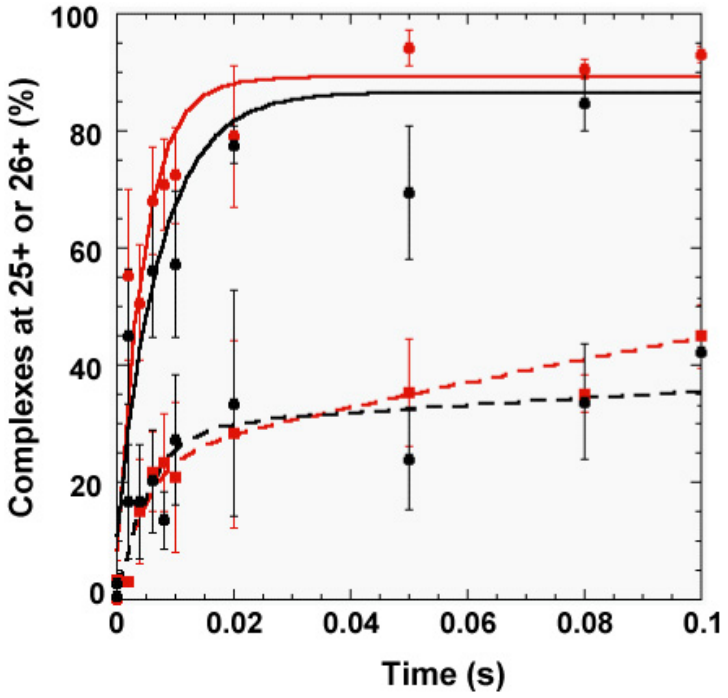


Figure 3.3B

Figure 3.3 Walker-RNAP Exhibits Altered Kinetics for AMP Incorporation in Simultaneous Addition Experiments.

Figure 3.3 Walker-RNAP Exhibits Altered Kinetics for AMP Incorporation in Simultaneous Addition Experiments. Plots of the wt-RNAP data are colored black, while plots of the Walker-RNAP are colored in red. Fits to 25+ data are solid and fits to 26+ data are dashed.

A. Simultaneous addition of 100μM CTP and 10μM ATP. There is no significant change in the rate of CMP incorporation between wt-RNAP; however, the slow phase of the Walker-RNAP shows a clear rate increase.

B. The same plot shown in **A.** only the time scale is 0.1s.

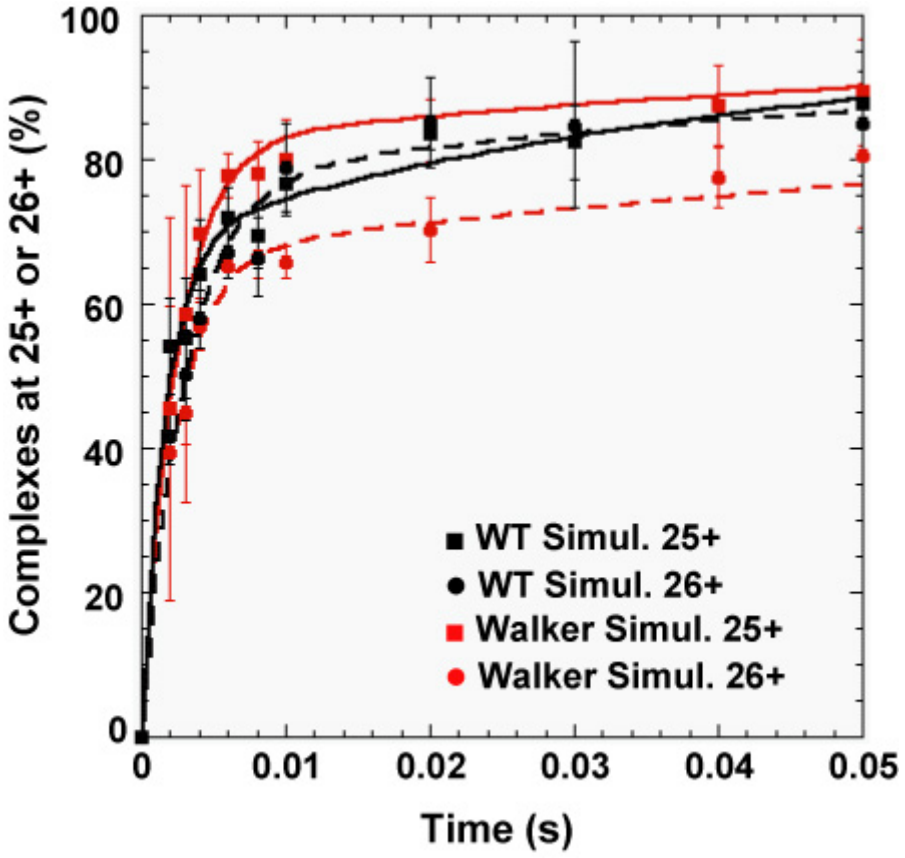


Figure 3.4 Walker-RNAP Exhibits Reduced Extent for AMP Incorporation When ATP is Pre-incubated. Plots of the wt-RNAP data are colored black, while plots of the Walker-RNAP are colored in red. Fits to 25+ data are solid and fits to 26+ data are dashed. The rate of CMP incorporation is not significantly affected in Walker-RNAP. AMP exhibits a small decrease in its extent but not rate of incorporation.

Upon inspection of the data for AMP incorporation at position 26+ (i+2), a small, but significant decrease in the extent is observed; however the rates of the fast and slow phases are not significantly affected.

On the surface, Walker-RNAP appears to give contradictory results depending on if the NTPs were added simultaneously or pre-incubated. In the simultaneous addition experiments, there is no significant alteration in either the rate or the extent of the complexes in the activated state; however, there is a significant increase in the rate, but not the extent, of complexes in the unactivated state. In direct opposition to those results, the pre-incubation experiments show no significant reduction in the rate of AMP incorporation; instead, a decrease in the extent of complexes that rapidly incorporate AMP is observed. The data are consistent with the mutation lowering binding affinity of the NTP for the allosteric site.

3.2.3 Structural Implications of Walker-RNAP

In Chapter 2, we proposed a model for NTP binding in the main channel through the use of fork loop 2. In the model, fork loop 2 acts primarily as an allosteric site by stabilizing the opening of the downstream DNA bubble and facilitating translocation. Translocation allows for NTP entry into the catalytic site via the secondary channel; however, the fork loop 2 bound NTP is also capable of being transferred into the catalytic site from the main channel and is responsible for the slow phase on elongation.

The increase in the rate for the slow phase in the simultaneous addition experiments can be explained by the allosteric site having a reduced binding affinity for the allosteric NTP. An NTP that was not as tightly bound to the allosteric site would

spend less time associated with allosteric site and, therefore, could be transferred into the catalytic site via the main channel pathway. Additionally, the reduction in the extent of complexes in the fast phase for pre-incubated elongation complexes can be explained by a reduced binding affinity for NTPs. With a reduced binding affinity, a smaller percentage of complexes would facilitate translocation because they would not have an NTP bound to the allosteric site. Therefore, the extent of complexes in the fast phase would be reduced, but the reaction would be able to go to completion, given enough time.

A key prediction of this model is that a higher percentage of complexes would load NTPs through the main channel. It would be interesting to perform the same competitive incorporation experiment as performed on wt-RNAP in Chapter 2 in which complexes are pre-incubated with 10μM [^{32}P]-ATP prior to the addition if 100μM CTP and increasing concentrations of unlabeled ATP. If the slow phase represents the population of complexes that incorporate the allosteric NTP, then a Walker-RNAP should incorporate a higher percentage of [^{32}P]-ATP. There should also be a concomitant increase in the rate of [^{32}P]-AMP incorporation. Performing quench-flow experiments to determine if this is true would be interesting, although the amount of [^{32}P]-ATP used could prove to be prohibitive.

3.3 Experimental Procedures

3.3.1 Sources of Protein and DNA

All DNA templates used were derived from pDE13 using a QuickChange Site-Directed Mutagenesis Kit (Stratagene) (Erie, Hajiseyedjavadi *et al*. 1993). All mutant templates were purified by a Qiagen Miniprep Kit (Qiagen) or by phenol-chloroform

extraction and the sequences confirmed by the DNA sequencing facility at UNC. All DNA transcription templates were amplified by PCR. The 540bp PCR product is biotinylated on one end and contains the λP_R promoter. The sequence of the first 30nt of DE13-A27g transcript is as follows:

DE13-A27g: pppAUGUAGUAAGGAGGUUGUAUGGAA<u>C</u>AGCGC

The first cytosine to be incorporated is located at +25 (underlined).

Wild-type His_6-tagged RNAP was purified from log phase *E. coli* strain RL916 (gift of R. Landick) using the protocol described previously (Burgess and Jendrisak 1975; Uptain and Chamberlin 1997). His_6-tagged β-D446A/G449A RNAP was made by site-directed mutagenesis using the QuickChange Site-Directed Mutagenesis Kit on the pIA-509 plasmid. pIA-509 was a gift from Irina Artsimovitch and was made as described previously (Artsimovitch, Vassylyeva *et al*. 2005). Expression, induction, and purification were carried out as described in chapter 4

3.3.2 In Vitro Transcription Reactions

To form OPCs, 80nM RNAP was combined with 80nM template DNA bound to streptavidin magnetic beads in 1x transcription buffer [30mM HEPES (pH 8.0), 10mM Mg^{+2} glutamate, 200mM K^+ glutamate, 25μg/mL BSA, and 1mM DTT] and incubated at 37°C for 10 min. After incubation, SECs stalled at position +24 were formed by the addition of a mixture of 20μM ATP, 15μM UTP, 16μM GTP, and 4μM [α-^{32}P]-GTP (800Ci/mmol) to the OPCs at room temperature. The reaction was allowed to proceed for 1 min in all experiments. SECs were purified from free NTPs by holding the reaction tube next to a magnet and washing with cold 1x transcription buffer a total of seven to ten

times. The complexes were then resuspended in 1x transcription buffer to a final SEC concentration of ~14nM. The purified complexes were kept on ice until used in rapid quench kinetic experiments. All rapid quench experiments were conducted on a Kintek Rapid Quench 3 device at room temperature (~23°C). For each reaction time point, ~20µL of purified SECs were loaded into a sample injection loop and ~20µL of the required NTP solution was loaded on the other sample injection loop. Reactions were quenched with 0.5M EDTA-NaOH (pH 8.0) or 1M HCl, as indicated in each experiment, after the reaction was allowed to proceed for the desired time. To make sure that the interval between time points while conducting the experiments did not affect the results, the order in which the time points were collected was altered. In addition, to ensure that complexes were still active throughout the time course of the experiment, a small aliquot of purified SECs were reacted with 1mM of all four NTPs and allowed to extend to full length transcript. The EDTA was removed from the samples by centrifuging the samples and holding them next to a strong magnet as described previously. To remove the HCl from the samples, 10µg of yeast tRNA, along with a 2X reaction volume of 100% ethanol, were added to each time point. The reactions were incubated at -20°C for a minimum of 15 minutes. The reactions were then centrifuged and the HCl solution removed. Complexes for both the EDTA and HCl quench reactions were resuspended in 100% formamide, boiled for two minutes, and loaded onto an 8M urea, 20% polyacrylamide gel.

3.3.3 Data Quantification and Normalization of Transcription Rate Data

The amount of radiation in each sample lane was measured by a Molecular Dynamics Phosphoimager and analyzed using ImageQuant v5.2 software. The percentage of

complexes at each position was calculated by dividing the amount of radiation at each template position by the total amount of radiation in bands ≥+24 for each lane. The data were normalized to 100% by dividing the amount of radiation at each time point by the highest percentage of complexes that incorporated past position +25 or +26. All experiments were conducted a minimum three times per NTP concentration. For the allosteric binding experiments, the data were normalized to 100% by dividing the amount of radiation at position +26 in each lane by the sample with the highest amount of radiation.

3.3.4 Kinetic Fits to the Data

Fits of the data in each plot were performed using KaleidaGraph v4.01. The data were fit to both single and double-exponentials. Since the data is clearly biphasic in most cases, it was not surprising to that fitting to a double exponential function gave significantly better results. This is consistent with previous work on the kinetics of nucleotide incorporation (Foster, Holmes et al. 2001; Holmes and Erie 2003). Error bars in the rapid quench experiments are reported as standard error, while error bars in the allosteric binding experiments are reported as standard deviations.

BIBLIOGRAPHY

Artsimovitch, I., M. N. Vassylyeva, et al. (2005). "Allosteric Modulation of the RNA Polymerase Catalytic Reaction Is an Essential Component of Transcription Control by Rifamycins." Cell **122**(3): 351-363.

Black, M. E. and D. E. Hruby (1992). "Site-directed Mutagenesis of a Conserved Domain in Vaccinia Virus Thymidine Kinase: Evidence for a Potential Role in Magnesium Binding." J. Bio. Chem. **267**(10): 6801-6806.

Bourne, H. R., D. A. Sanders, et al. (1991). "The GTPase Superfamily: Conserved Structure and Molecular Mechanism." Nature **349**(6305): 117-127.

Burgess, R. R. and J. J. Jendrisak (1975). "Procedure for the Rapid, Large-Scale Purification of *Escherichia coli* DNA-Dependent RNA Polymerase Involving Polymin P Precipitation and DNA-Cellulose Chromatography." Biochem. **14**(21): 4634 - 4638.

Erie, D. A., O. Hajiseyedjavadi, et al. (1993). "Multiple RNA Polymerase Conformations and GreA: Control of the Fidelity of Transcription." Science **262**(5135): 867-873.

Foster, J. E., S. F. Holmes, et al. (2001). "Allosteric Binding of Nucleoside Triphosphates to RNA Polymerase Regulates Transcription Elongation." Cell **106**: 243-252.

Fry, D. C., S. A. Kuby, et al. (1986). "ATP-binding Site of Adenylate Kinase: Mechanistic Implications of its Homology with Ras-encoded p21, F1-ATPase, and Other Nucleotide-binding Proteins." Proc. Nat. Acad. Sci. **83**(4): 907-911.

Holmes, S. F. and D. A. Erie (2003). "Downstream DNA Sequence Effects on Transcription Elongation: Allosteric Binding of Nucleoside Triphosphates Facilitates Translocation Via a Ratchet Motion." J. Bio. Chem. **278**(37): 35597-35608.

Kuby, S. A., R. H. Palmieri, et al. (1984). "Studies on Adenosine Triphosphate Transphosphorylases. Amino Acid Sequence of Rabbit Muscle ATP-AMP Transphosphorylase." Biochem. **23**(11): 2393-2399.

Leipe, D. D., Y. I. Wolf, et al. (2002). "Classification and Evolution of P-loop GTPases and Related ATPases." J. Mol. Bio. **317**(1): 41-72.

Pai, E. F., U. K. G. A.Petsko, et al. (1990). "Refined Crystal Structure of the Triphosphate Conformation of H-ras p21 at 1.35Å Resolution: Implications for the Mechanism of GTP Hydrolysis." EMBOJ **9**(8): 2351-2359.

Uptain, S. M. and M. J. Chamberlin (1997). "*Escherichia coli* RNA Polymerase Terminates Transcription Efficiently at Rho-Independent Terminators on Single-

Stranded DNA Templates." Proc. Na. Acad. Sci. **94**(25): 13548-13553.

Vassylyev, D. G., S. Sekine, et al. (2002). "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution." Nature **417**(6890): 712-719.

Via, A., F. Ferrè, et al. (2000). "Three-dimensional View of the Surface Motif Associated with the P-loop Structure: *Cis* and *Trans* Cases of Convergent Evolution." J. Mol. Bio. **303**(4): 455-465.

Walker, J. E., M. Saraste, et al. (1982). "Distantly Related Sequences in the α- and β-Subunits of ATP Synthase, Myosin, Kinases and Other ATP-Requiring Enzymes and a Common Nucleotide Binding Fold." EMBOJ **1**(8): 945-951.

CHAPTER 4:

PHOSPHODIESTER BOND FORMATION AND PYROPHOSPHATE RELEASE

4.1 Effects of Quenching Elongation Reactions with HCl

All known DNA and RNA polymerases employ two divalent metal ions in order to perform the phosphoryl transfer reaction (Brautigam and Steitz 1998). One metal ion is chelated by active site residues and is responsible for the reduction of the 3'OH group of the substrate nucleic acid chain (DNA or RNA). The other metal ion is transferred into the catalytic site bound to the phosphate moiety of the incoming NTP and is used in the nucleophilic attack of the α-phosphorous atom.

Previous work on RNA dependent RNA polymerases has shown that quenching with EDTA or HCl leads to the elongation reaction being quenched at different stages of incorporation (Arnold and Cameron 2004; Arnold, Gohara *et al*. 2004). The difference between EDTA and HCl quench is due to how each molecule quenches the reaction. EDTA quenches the reaction by chelating the Mg^{+2} that is associated with the incoming NTP; NTPs are unable to be incorporated because one of the two required Mg^{+2} is no longer available for use in catalysis. However, elongation complexes that have an NTP already bound in the active site prior to the addition of EDTA are immune to quenching because EDTA cannot chelate the Mg^{+2} ion; therefore, the reaction is able to continue towards completion. HCl quenches the incorporation reaction instantaneously upon addition by denaturing the enzyme, which demonstrates the time of bond formation.

Because RNAP uses the divalent metal ion system to perform its catalytic functions, quenching the reaction with EDTA or HCl will lead to stopping of the elongation reaction at different stages.

4.1.2 Quenching with HCl or EDTA Alters AMP Incorporation

Previous work by Burton and coworkers has shown that quenching elongation reactions with EDTA and HCl alters the kinetics of nucleotide addition in RNAPII (Gong, Zhang *et al*. 2005; Zhang, Zobeck *et al*. 2005). Due to these previously published results, we were interested in examining the effect of pre-incubating the i+2 nucleotide (ATP) on the rates of incorporation of the i+1 nucleotide (CMP), and the i+2 nucleotide (AMP) when complexes were quenched with HCl instead of EDTA. To that end, we performed experiments in which SECs were pre-incubated with either 10μM prior to the addition of CTP (100μM) and then quenched the reaction by the addition of 1M HCl. Simultaneous addition experiments were done as controls using the same NTP concentrations as the pre-incubation experiments (Figure 4.1).

The data for AMP incorporation exhibits biphasic kinetics with a fast phase followed by a slow phase. The biphasic kinetics are readily apparent in the simultaneous addition experiments. The results for both pre-incubation simultaneous addition experiments show a marked decrease in the extent of complexes in the fast phase for AMP incorporation compared to reactions quenched with EDTA (Figure 4.2A and 4.2B). Interesting, there is a ~40% reduction in the percentage of complexes in the fast chase for both the pre-incubation and simultaneous addition experiments when quenched with HCl instead of EDTA. This result means that, of the complexes in which ATP is isomerized

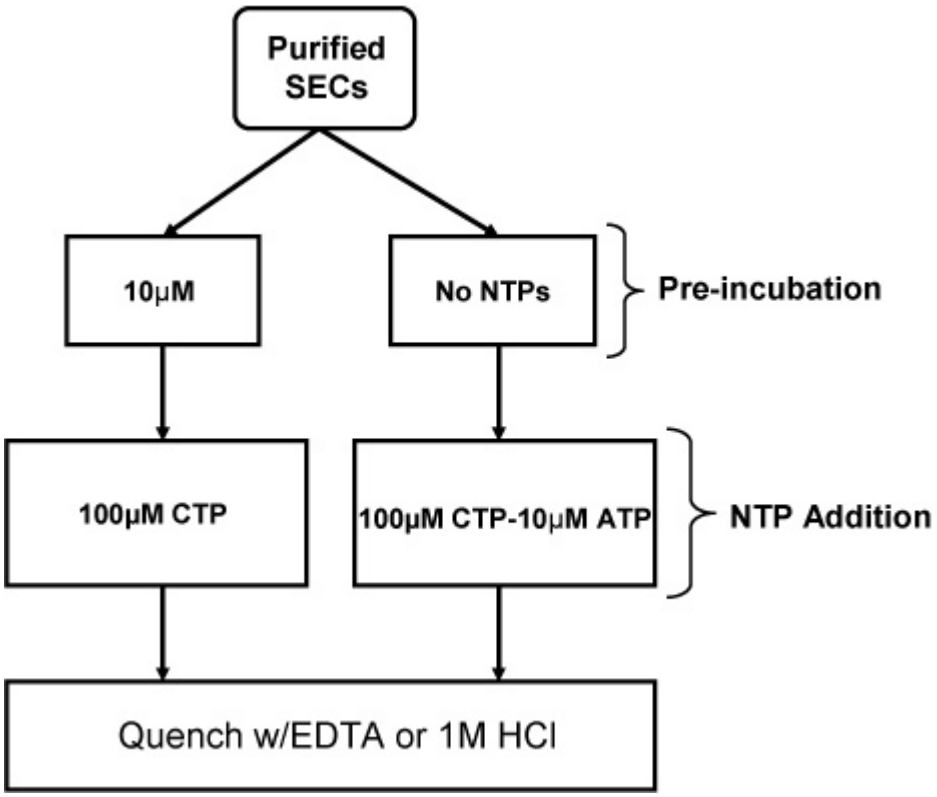


Figure 4.1 Experimental Design to Test for Differences in Quench Conditions. Purified SECs are split into two groups; one in which 10μM ATP is added to the complexes prior to the addition of 100μM CTP via rapid quench, and one in which 100μM CTP and 10μM ATP are added simultaneously by rapid quench techniques. The reactions are either quenched with 0.5M EDTA or with 1M HCl.

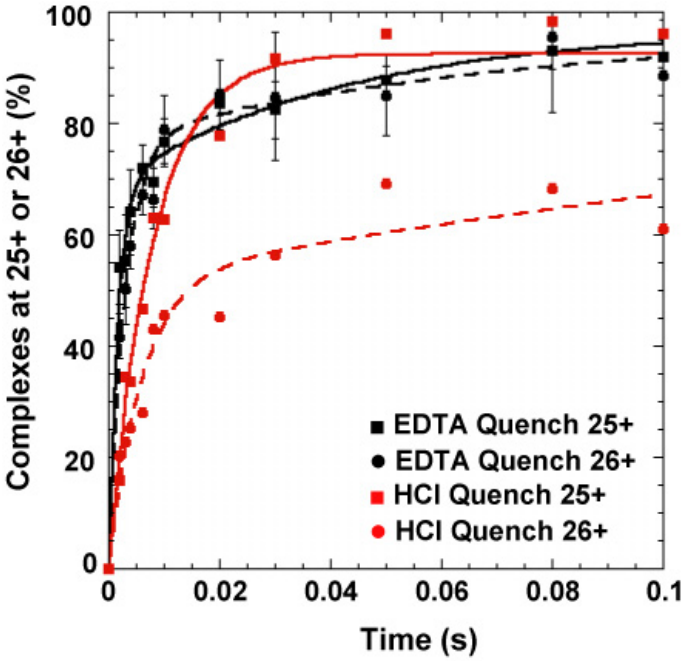


Figure 4.2A

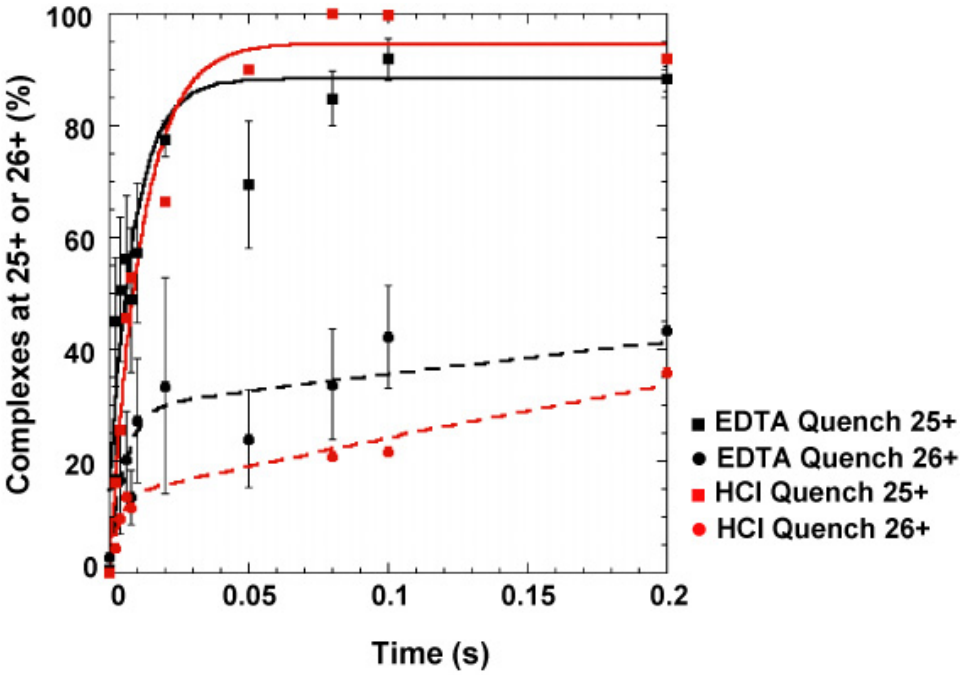


Figure 4.2B

Figure 4.2 Quenching Reactions with EDTA or HCl Alters the Extents of AMP Incorporation.

Figure 4.2 Quenching Reactions with EDTA or HCl Alters the kinetics AMP Incorporation. All EDTA quenched reactions are designated in black while all HCl quenched reactions are designated in red. Fits for CMP incorporation data (25+) are solid while curve fits for AMP incorporation data (26+) are dashed.

A. Plots are for pre-incubation experiments. The extent and rate of AMP incorporation is reduced by quenching with HCl when compared to experiments using EDTA as the quench. The EDTA quench data are presented in Chapter 2.

B. Plots are for the simultaneous addition experiments. The extent and rate of AMP incorporation is reduced by quenching with HCl when compared to experiments using EDTA as the quench. The EDTA quench data are presented in Chapter 2.

into the catalytic site, only half of them incorporate AMP prior to being quenched with HCl. This result is in stark contrast to what is observed for RNAPII(Gong, Zhang *et al*.2005).

Analogous experiments to the simultaneous addition experiments were carried out with human RNAPII. When quenched with HCl, RNAPII elongation complexes no longer exhibited biphasic kinetics. Instead, the fast phase for NMP incorporation was completely abolished, leaving complexes only in the slow phase (Gong, Zhang *et al*. 2005; Zhang, Zobeck *et al*. 2005). Prokaryotic RNAP does not exhibit a complete disappearance of the fast phase. This suggests that binding of an NTP into the catalytic site is faster in prokaryotic RNAP compared to eukaryotic RNAPII.

The rate of AMP incorporation for the fast phase in the pre-incubation experiments is significantly reduced in experiments in which reactions were quenched with HCl instead of EDTA. Specifically, there is a 49%±16% reduction in the rate of AMP when quenched with HCl. Given that quenching with HCl gives the rate of bond formation, then the ~50% reduction in rate is due to the rate of bond synthesis being ~50% slower than the rate of NTP binding into the active site (EDTA quench). The rate reduction is consistent with the reduction in the extent of AMP incorporation; if NTP incorporation is half as fast at the rate of NTP binding to the catalytic site, then only half of the complexes will incorporate an NTP. Additionally, there is a reduction in the rate of AMP incorporation for the simultaneous addition experiments; however, there is too much error in the measurements to be able to obtain an actual rate.

4.2 The Nonspecific Competitors dTTP and $Na_5(PO_4)_3$ Act as Pyrophosphate Release Inhibitors

The NTP addition cycle in transcription can be reduced to four basic steps: NTP binding to the catalytic site, phosphodiester bond formation, pyrophosphate release, and translocation of the RNAP along the DNA. Of particular interest is the translocation step. Translocation is the movement of RNAP along the template DNA by one base, thereby removing the 3'end of the RNA/DNA hybrid from the catalytic site and allowing the next templated NTP to enter.

There are three models that have been proposed posits the mechanism of translocation: a Brownian ratchet mechanism, a powerstroke mechanism, and a NTP-driven translocation mechanism. The Brownian ratchet mechanism describes a rapid equilibrium between the pre- and post-translocated state, with NTP binding only to the post-translocated state. In the powerstroke mechanism, the enzyme utilizes the energy from bond formation to drive translocation forward (Bar-Nahum, Epshtein *et al*. 2005; Sousa 2005). In the NTP-driven translocation model, binding of NTPs to the downstream DNA facilitates or drives translocation. Previous work done on *E. coli* RNAP and human RNAPII strongly supports the NTP-driven translocation model (Foster, Holmes *et al*. 2001; Nedialkov, Gong *et al*. 2003; Zhang and Burton 2004; Gong, Zhang *et al*. 2005; Zhang, Zobeck *et al*. 2005).

In the course of investigating NTP-driven translocation, Burton and coworkers discovered that RNAPII can under go "isomerization reversal" of the catalytic NTP. Isomerization reversal is defined as the release of the substrate NTP that had been previous bound, or "isomerized", into the catalytic site, but not yet incorporated (Gong, Zhang *et al*. 2005). To observe isomerization reversal, the potent translocation inhibitor α-amanitin and high concentrations of the i+2 NTP were used. In their experiments, the

authors observed that, in the presence of α-amanitin and high concentrations of the i+2 nucleotide, which binds in the main channel, induces the release of the i+1 NTP previously isomerized into the catalytic site. Here, we report that, in the presence of high concentrations of the nonspecific competitors dTTP and $Na_5(PO_4)_3$, *E. coli* RNAP pre-incubated with 10μM ATP (i+2 nucleotide) appears to undergo active site isomerization reversal.

4.2.1 dTTP Induces Isomerization Reversal and Inhibits Pyrophosphate Release

To examine the effect of pre-incubating the i+2 nucleotide (ATP) and a nonspecific competitor on the rates of incorporation of the i+1 (CMP) and i+2 (AMP) nucleotides, we performed experiments in which SECs were pre-incubated with 10μM ATP and 1mM dTTP prior to the initiating the elongation reaction with CTP and quenching with 0.5M EDTA. A high concentration of CTP (100μM) was used so that CMP incorporation would not be rate limiting for AMP incorporation. These experiments are very similar to experiments described in chapter 2 in which complexes were pre-incubated with 10μM ATP and 1mM GTP (Figure 2.8). The results are presented in Figure 4.3. Surprisingly, the results from this experiment are very different from the results using GTP as the nonspecific competitor. While the incorporation of CMP (solid black line) is unaffected by the presence of 1mM dTTP, the rate on AMP (dashed black line) incorporation is strikingly different from experiments in which dTTP was not included. At early time points, AMP rapidly incorporates; however, after rapid incorporation, a clear reversal in AMP incorporation is observed. This reversal continues until ~0.08s, at which point AMP is slowly reincorporated.

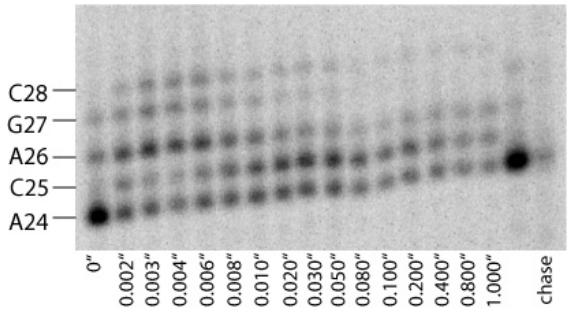


Figure 4.3A

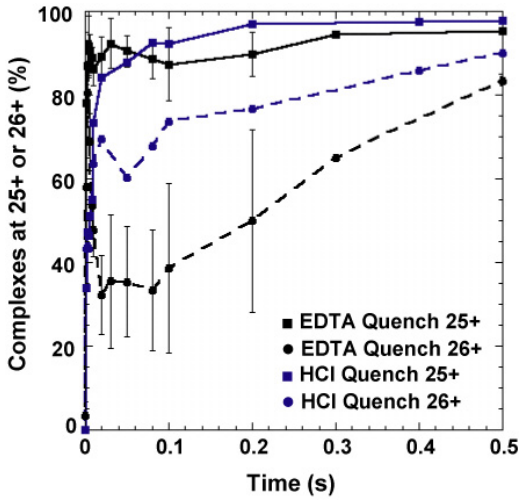


Figure 4.3B

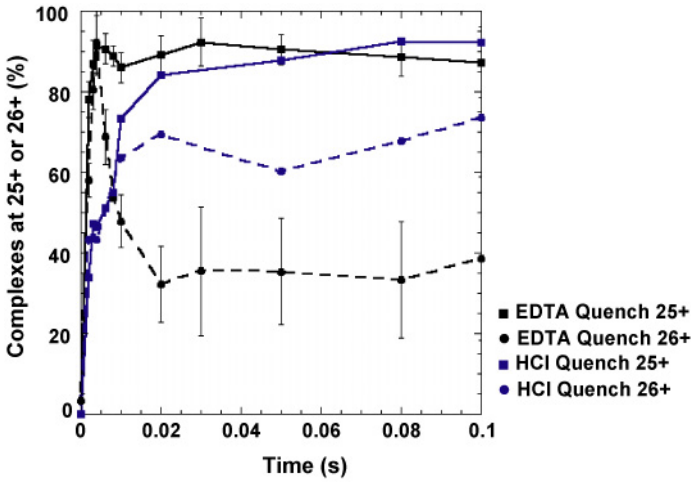


Figure 4.3C

Figure 4.3 Pre-incubation with 10μM ATP and 1mM dTTP Induces Isomerization

Reversal and Inhibition of Pyrophosphate Release.

Figure 4.3 Pre-incubation with 10μM ATP and 1mM dTTP Induces Isomerization Reversal and Inhibition of Pyrophosphate Release. The

A. Representative gel of an experiment in which elongation complexes were pre-incubated with 10μM ATP and 1mM dTTP and quenched with EDTA. The gels show the time course for CMP incorporation at position +25 and AMP incorporation at position +26. The positions are indicated on the left of the gel. The times are presented in seconds and denoted at the bottom of the gel.

B. Plots showing the rate and extents of CMP (+25) and AMP (+26) when reactions are quenched with 0.5M EDTA or 1M HCl. EDTA quench reactions are colored in black while HCl quench reactions are blue. The 25+ data are plotted with a solid line and the 26+ data are plotted with a dashed line.

C. The same as **B.** with a time scale of 0.1s

Previous work by Burton and coworkers has shown that quenching elongation reactions with EDTA and HCl alters the kinetics of nucleotide addition in RNAPII (Gong, Zhang *et al*. 2005; Zhang, Zobeck *et al*. 2005). We were interesting in examining if quenching with HCl would alter the reversal of AMP incorporation. Examination of Figure 4.3 clearly shows that the reversal reaction for AMP incorporation is no longer present when reactions are quenched with HCl (dashed blue line). The data for the HCl quench experiments, agree with similar experiments in which complexes were only pre-incubated with ATP (Figure 4.2A). This result indicates that the presence of dTTP in the reaction does not affect the rate of bond formation.

Closer inspection of Figure 4.3 reveals several attributes of both the EDTA and HCl plots that give important clues as to what is happening. The first interesting observation is the initial burst of AMP incorporation seen in the EDTA quench data. AMP appears to incorporate very rapidly, which is consistent with the results presented in Chapter 2. As previously noted, quenching with EDTA reveals the rate of NTP binding to the i+1 site (active site) while HCl quenching reveals the rate of phosphodiester bond formation. Additionally, it is important to note that, upon EDTA addition, phosphodiester bond formation has progressed to the extent revealed by the HCl curve of the same time point. Any deviation from the HCl curve indicates events that happen after EDTA addition.

Comparing the EDTA quench data to the HCl quench data reveals that, at the point of maximal AMP incorporation in the EDTA data, only ~50% of complexes have undergone bond formation. After the maximum burst height, there is a rapid reversal in the extent of AMP incorporation. This reversal is indicative of isomerization reversal of

the catalytic nucleotide. The initial burst in AMP incorporation seen at early time points (≤~0.004s) is due to the dTTP not having enough time to affect AMP incorporation before being inactivated by EDTA; however, the ATP is not inactivated because it has already been isomerized into the active site and is immune to the EDTA quench. Since dTTP is inactivated, isomerization reversal is not seen at early time points. Furthermore, these results indicated that ATP isomerization is much faster than either dTTP binding or its mode of action.

Another important feature of Figure 4.2 is the fact that the HCl curve rises above the EDTA quench curve. This feature indicates that the phosphodiester bond formation seen in the HCl quench curve undergoes bond reversal and subsequent isomerization reversal. If bond reversal did not occur, then the HCl and EDTA curves would be coincident with one another. This result indicates that dTTP binds to the enzyme, causes a conformational change, and induces bond reversal or inhibits pyrophosphate release.

The last interesting attribute of this plot is the slow reincorporation of AMP seen at longer time points in the EDTA curve. This observation can be explained by a transient blocking of pyrophosphate release. Assuming dTTP binding is transient, then reactions quenched with EDTA are able to go to completion because at long time points, AMP can be incorporated and pyrophosphate released.

4.2.2 $Na_5(PO_4)_3$ Induces Isomerization Reversal and Inhibits Pyrophosphate Release

We were interested in determining if the effects of dTTP on AMP incorporation were specific to dTTP if it was a more general effect. To that end, we decided to use $Na_5(PO_4)_3$. The results are presented in Figure 4.4. The plots clearly show that a similar

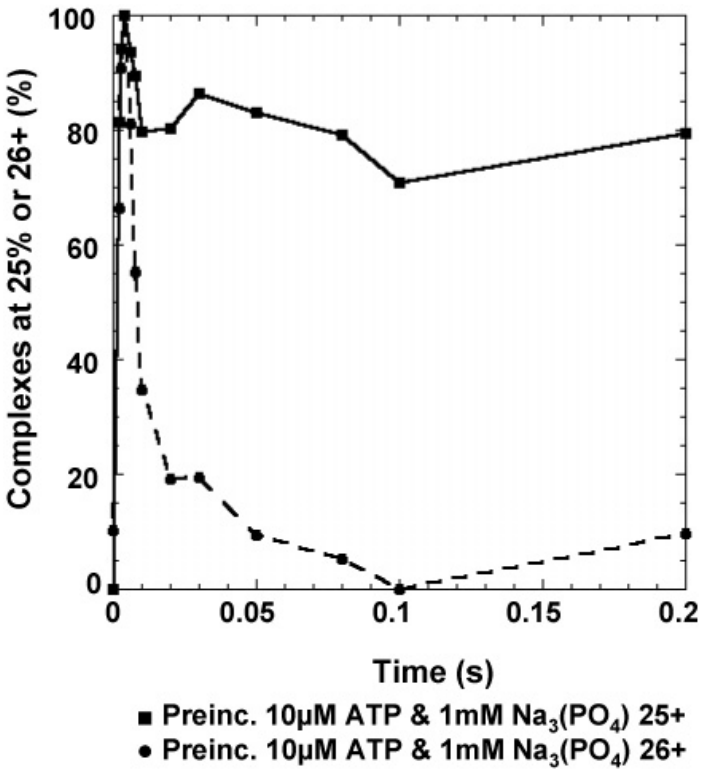


Figure 4.4 Pre-incubation of Elongation Complexes with ATP and $Na_5(PO_4)_3$ Induces Isomerization Reversal. Incorporation at position 25+ (CMP) is represented by a solid line and filled squares. Incorporation at position 26+ (AMP) is represented by a dashed line and filled circles.

reversal in AMP incorporation takes place. In fact, the time scales in which rapid incorporation and subsequent isomerization reversal takes place are similar to the experiments in which 1mM dTTP was present. There are two primary differences between the reactions in which 1mM dTTP and 1mM $Na_5(PO_4)_3$ are present. The first difference is that the extent of reversal is significantly greater in complexes treated with $Na_5(PO_4)_3$. The second major difference is that complexes fail to recover from treatment with $Na_5(PO_4)_3$. The two results are likely caused by high concentrations of contaminating pyrophosphate found. High concentrations of pyrophosphate would cause AMP incorporation to compete with pyrophosphorolysis of the transcript.

4.2.3 Comparison of Results Obtained for Human RNAPII

Gong *et al.*(2005) observe similar results those results presented here. There are some minor differences between the two results. The authors use the translocation inhibitor α-amanitin and a high concentration of the i+2 nucleotide to induce isomerization reversal. Interestingly, they observed reversal when the nucleotide used was the i+2 templated nucleotide. The authors attribute isomerization reversal on the i+2 nucleotide trying to drive translocation in the presence of a translocation inhibitor. This observation is in stark contrast to the results presented here. In the results presented in Figure 4.3, after AMP incorporation, there is no i+2 NTP to attribute to isomerization. According to their model, the isomerization reversal would be focused on CTP in our experiments.

Another major difference between the two experiments is that, in the experiments done by Gong *et al.*(2005), the reactions never go to completion. In contrast, the

experiments in which 1mM dTTP is introduced, the reactions are able to slowly go to completion. The cause for this discrepancy is due to the use of α-amanitin in the experiments done by the authors. Since α-amanitin permanently disables translocation in RNAPII, complexes will never be able to go to completion. In experiments using dTTP, however, RNAP only transiently binds dTTP, therefore when dTTP is not bound to the complex, productive elongation can occur. It would be enlightening to perform these same experiments in the presence of a much more robust translocation blocker.

4.3 Experimental Procedures

4.3.1 Sources of Protein and DNA

Wild-type His_6-tagged RNAP was purified from log phase RL916 strain (gift of R. Landick) using the protocol described previously (Burgess and Jendrisak 1975; Uptain and Chamberlin 1997). All DNA templates used were derived from pDE13 using a QuickChange Site-Directed Mutagenesis Kit (Stratagene) (Erie, Hajiseyedjavadi *et al.* 1993). All mutant templates were purified by a Qiagen Miniprep Kit (Qiagen) or by phenol-chloroform extraction and the sequences confirmed by the DNA sequencing facility at UNC. All DNA transcription templates were amplified by PCR. The 540bp PCR product is biotinylated on one end and contains the λP_R promoter. The sequence of the first 30nt of DE13-A27g transcript is as follows:

DE13-A27g: pppAUGUAGUAAGGAGGUUGUAUGGAAC<u>C</u>AGCGC

The first cytosine to be incorporated is located at +25 (underlined).

4.3.2 In Vitro Transcription Reactions

To form OPCs, 80nM RNAP was combined with 80nM template DNA bound to streptavidin magnetic beads in 1x transcription buffer [30mM HEPES (pH 8.0), 10mM Mg^{+2} glutamate, 200mM K^{+} glutamate, 25μg/mL BSA, and 1mM DTT] and incubated at 37°C for 10 min. After incubation, SECs stalled at position +24 were formed by the addition of a mixture of 20μM ATP, 15μM UTP, 16μM GTP, and 4μM [α-^{32}P]-GTP (800Ci/mmol) to the OPCs at room temperature. The reaction was allowed to proceed for 1 min. SECs were purified from free NTPs by holding the reaction tube next to a magnet and washing with cold 1x transcription buffer a total of seven to ten times. The complexes were then resuspended in 1x transcription buffer to a final SEC concentration of ~14nM. The purified complexes were kept on ice until used in rapid quench kinetic experiments. All rapid quench experiments were conducted on a Kintek Rapid Quench 3 device at room temperature (~23°C). For each reaction time point, ~20μL of purified SECs were loaded into a sample injection loop and ~20μL of the required aqueous NTP solution was loaded on the other sample injection loop. Reactions were quenched with 0.5M EDTA-NaOH (ph 8.0) or 1M HCl after the reaction was allowed to proceed for the desired time. To make sure that the interval between time points while conducting the experiments did not affect the results, the order in which the time points were collected was altered. In addition, to ensure that complexes were still active throughout the time course of the experiment, a small aliquot of purified SECs were reacted with 1mM of all four NTPs and allowed to extend to full length transcript. The EDTA quench solution was removed from the samples by centrifuging the samples and holding them next to a strong magnet as described in Chapter 2. To remove the HCl quench solution from the samples, 10μg of yeast tRNA, along with a 2X reaction volume of 100% ethanol, were

added to each time point. The reactions were incubated at -20°C for a minimum of 15 min. The reactions were then centrifuged and the HCl solution removed. Complexes were resuspended in 8μL of 100% formamide, boiled for two min, and loaded onto an 8M urea, 20% polyacrylamide gel.

4.3.3 Data Quantification and Normalization of Transcription Rate Data

The amount of radiation in each sample lane was measured by a Molecular Dynamics Phosphoimager and analyzed using ImageQuant v5.2 software. The percentage of complexes at each position was calculated by dividing the amount of radiation at each template position by the total amount of radiation in bands ≥+24 for each lane. The data were normalized to 100% by dividing the amount of radiation at each time point by the highest percentage of complexes that incorporated past position +25 or +26. All experiments were conducted a minimum three times per NTP concentration. For the allosteric binding experiments, the data were normalized to 100% by dividing the amount of radiation at position +26 in each lane by the sample with the highest amount of radiation.

BIBLIOGRAPHY

Arnold, J. J. and C. E. Cameron (2004). "Poliovirus RNA-Dependent RNA Polymerase ($3D^{pol}$): Pre-Steady-State Kinetic Analysis of Ribonucleotide Incorporation in the Presence of Mg^{2+}." Biochem. **43**(18): 5126-5137.

Arnold, J. J., D. W. Gohara, et al. (2004). "Poliovirus RNA-Dependent RNA Polymerase ($3D^{pol}$): Pre-Steady-State Kinetic Analysis of Ribonucleotide Incorporation in the Presence of Mn^{2+}." Biochem. **43**(18): 5138-5148.

Bar-Nahum, G., V. Epshtein, et al. (2005). "A Ratchet Mechanism of Transcription Elongation and Its Control." Cell **120**(2): 183-193.

Brautigam, C. A. and T. A. Steitz (1998). "Structural and Functional Insights Provided by Crystal Structures of DNA Polymerases and Their Substrate Complexes." Current Opin. Struct. Bio. **8**(1): 54-63.

Burgess, R. R. and J. J. Jendrisak (1975). "Procedure for the Rapid, Large-Scale Purification of *Escherichia coli* DNA-Dependent RNA Polymerase Involving Polymin P Precipitation and DNA-Cellulose Chromatography." Biochem. **14**(21): 4634 - 4638.

Erie, D. A., O. Hajiseyedjavadi, et al. (1993). "Multiple RNA Polymerase Conformations and GreA: Control of the Fidelity of Transcription." Science **262**(5135): 867-873.

Foster, J. E., S. F. Holmes, et al. (2001). "Allosteric Binding of Nucleoside Triphosphates to RNA Polymerase Regulates Transcription Elongation." Cell **106**: 243-252.

Gong, X. Q., C. Zhang, et al. (2005). "Dynamic Error Correction and Regulation of Downstream Bubble Opening by Human RNA Polymerase II." Mol. Cell **18**(4): 461-470.

Nedialkov, Y. A., X. Q. Gong, et al. (2003). "NTP-driven Translocation by Human RNA Polymerase II." J. Bio. Chem. **278**(20): 18303-18312.

Sousa, R. (2005). "Machinations of a Maxwellian Demon." Cell **120**(155-156).

Uptain, S. M. and M. J. Chamberlin (1997). "*Escherichia coli* RNA Polymerase Terminates Transcription Efficiently at Rho-Independent Terminators on Single-Stranded DNA Templates." Proc. Nat. Acad. Sci. **94**(25): 13548-13553.

Zhang, C. and Z. F. Burton (2004). "Transcription Factors IIF and IIS and Nucleoside Triphosphate Substrates as Dynamic Probes of the Human RNA Polymerase II Mechanism." J. Mol. Bio. **342**(4): 1085-1099.

Zhang, C., K. L. Zobeck, et al. (2005). "Human RNA Polymerase II Elongation in Slow Motion: Role of the TFIIF RAP74 α1 Helix in Nucleoside Triphosphate-Driven Translocation." Mol. Cell. Bio. **25**(9): 3583-3595.

CHAPTER 5:

CLONING, MUTAGENESIS, AND EXPRESSION OF RECOMBINANT RNA POLYMERASE

5.1 Cloning of *Thermus thermophilus* RNA Polymerase

We were interested in developing an expression system for RNAP for the purpose of determining which residues were important in transcription elongation using mutational analysis. *Thermus thermophilus* RNAP was chosen for the following reasons: 1) an expression system was not available 2) the crystal structure was available so structure-function relationships could be easily inferred (Vassylyev, Sekine *et al*. 2002) 3) purification could be easily performed (Xue, Hogan *et al*. 2000).

All *T. thermophilus* genes encoding for the RNAP subunits (*rpoA*=α-subunit, *rpoB*=β-subunit, *rpoC*=β'-subunit, *rpoD*=σ-subunit, *rpoZ*=ω-subunit) were cloned out of genomic DNA. All PCR reactions were conducted in the presence of 10% DMSO (v/v) (Shen and Hohn 1992). The general cloning strategy is presented in Figure 5.1. The primer names and their corresponding sequences are listed in Table 5.1. The *rpoB* and *rpoC* genes were too large to clone directly out of genomic DNA. To solve this problem, *rpoB* was cloned in two fragments. The 1366bp *rpoB'* fragment contained an *Nde*I site 5' of the start codon and an *Eco*RI site engineered 1366bp into the gene using primers "*rpoB*'-frag for" and "*rpoB'*-frag rev". These engineered restriction sites did not alter the underlying protein sequence. This fragment was ligated into pET-21c and termed pSK3. The 1994bp *rpoB* fragment contained an *Eco*RI site that overlaps with the *Eco*RI site of

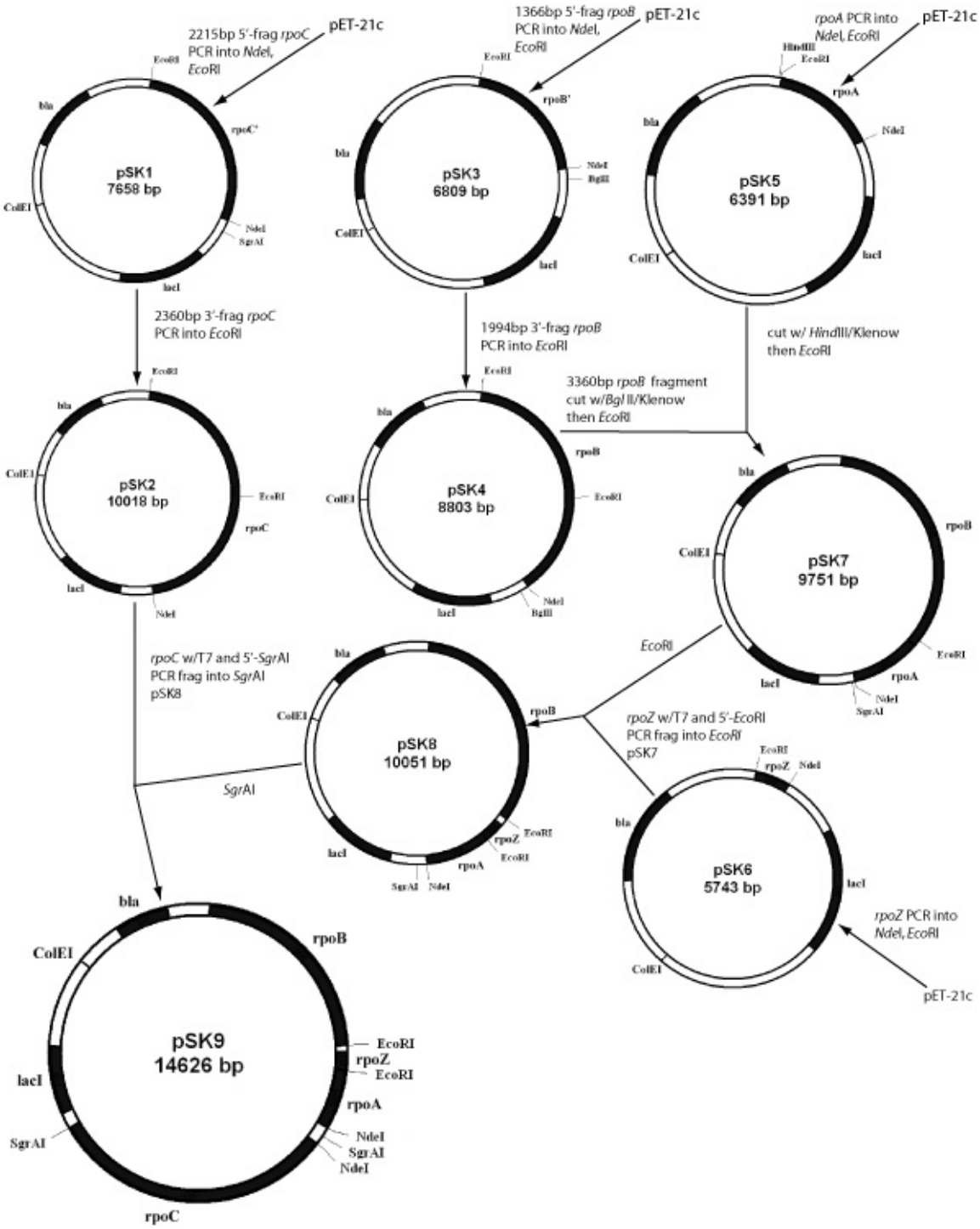


Figure 5.1 Schematic for the Construction of pSK9. The names of each construct, along with their respective size in base-pairs, are located in the inside of each schematic. Open reading frames are designated by black filled regions. Important restriction sites are labeled next to a tick mark.

Primer Name	Primer Sequence
rpoA-NdeI for	5'-GGG GAT TCC ATA TGT TGG ATT CCA AGC TCA AGG C
rpoA-EcoRI rev	5'-GGA ATT CTC ACT CCT TCA GGG TGA AGC C
rpoB'-frag for	5'-ATT CGC CCA TAT GGA GAT CAA GCG GTT C
rpoB'-frag rev	5'-GGA ATT CAC GGA GAA GAC CTG CTT GG
rpoB-3'-frag for	5'-CCC ATA TGG AGA TCA AGC GGT TC
rpoB-3'-frag rev	5'-GGA ATT CAC GGA GAA GAC CTG CTT
rpoC'-frag for	5'-ATT CGC CCA TAT GAA AAA AGA GGT TCG TAA G
rpoC'-frag rev	5'-GGA ATT CGG CGT TGA AGG CCT C GC
rpoC-3'-frag for	5'-GGA ATT CGA CGG GGA CCA GAT G
rpoC-3'-frag rev	5'-GGA ATT CGC CCT TTT AAG CCT GGT T
rpoC-T7-SgrAI for	5'-GCCACAGGTGCGGTTGCTGG
rpoC-T7-SgrAI rev	5'-AGCGGCTCTTAAGCCTGCTTGCCGGGCTG
rpoD-NdeI for	5'-GGG GAT TCC ATA TGA AGA AGA GCA AGC GCA AG
rpoD-EcoRI rev	5'-GGA ATT CTT AGT CCA GGA AGT CCC TGA G
rpoZ-NdeI for	5'-GGG GAT TCC ATA TGG CGG AAC CGG GCA TT
rpoZ-EcoRI rev	5'-GGA ATT CCT ACT CCT CCC GCT CCA C
rpoZ-T7-EcoRI for	5'-GGA ATT CGC GAA ATT AAT ACG ACT CAC TAT

Table 5.1 PCR Primers Used in the Construction of pSK9. Underline bases constitute the location of the designated restriction site.

rpoB' and an *Eco*RI site 3' of the stop codon, respectively. pSK3 was cut with *Eco*RI and the *rpoB* PCR fragment was ligated in, yielding pSK4. The internal *Eco*RI site of the completed *rpoB* gene was removed by site-directed mutagenesis as described later in this chapter.

The cloning of *rpoC* followed a similar strategy as the one followed for *rpoB*. The 2215bp *rpoC'* fragment was cloned using primers "*rpoC'*-frag for" and "*rpoC'*-frag rev", which engineered an *Nde*I site 5' of the start codon and an *Eco*RI site engineered 2215bp into the gene. This fragment was ligated into pET-21c and termed pSK1. The 2360bp 3' fragment of *rpoC* was cloned using primers "*rpoC*-3'-frag for" and "*rpoC*-3'-frag rev", which contain an *Eco*RI site that overlapped with the *Eco*RI site of *rpoC'* and an *Eco*RI site 3' of the stop codon, respectively. This PCR fragment was ligated into pSK1 cut with *Eco*RI to yield pSK2.

The *rpoA*, *rpoD*, and *rpoZ* genes were cloned by PCR with an *Nde*I site 5' of the start codon and an *Eco*RI site 3' of the stop codon and ligated into pET-21c. The primers used were "*rpoA*-NdeI for" and "*rpoA*-EcoRI rev" for *rpoA*, "*rpoD*-NdeI for" and "*rpoD*-EcoRI rev" for *rpoD*, and "*rpoZ*-NdeI for" and "*rpoZ*-EcoRI rev" for *rpoZ*. The plasmids containing *rpoA*, *rpoD*, and *rpoZ* were termed pSK5, pSK6, and pSK10, repectively.

To introduce *rpoB* and its associated T7 promoter into pSK5 (containing *rpoA*), pSK4 (containing *rpoB*) was treated with *Bgl*II, which cuts upstream of the T7 promoter, and then treated with Klenow to form a blunt end. *RpoB* was then excised from pSK4 by treating with *Eco*RI. The resulting *T7::rpoB* fragment had a blunt end 5' of the T7 promoter and an *Eco*RI sticky end 3' of the stop codon. *T7::rpoB* was ligated into pSK5 that had been cut with *Hind*III, blunted with Klenow, and then further treated with *Eco*RI.

The resulting plasmid, pSK7, contains *rpoA* and *rpoB*, each under the control of a separate T7 promoter with the ORFs in opposite directions.

To introduce *rpoZ* with its own T7 promoter into pSK7, it was necessary to engineer an *Eco*RI site upstream of the T7 promoter. To do this, *rpoZ* was amplified from pSK6 using primers "*rpoZ*-T7-EcoRI for", which engineers an *Eco*RI site upstream of the T7 promoter, and "*rpoZ*-EcoRI rev". The *T7::rpoZ* PCR fragment was treated with *Eco*RI and ligated into pSK7 cut with *Eco*RI, yielding pSK8. *T7::rpoZ* is able to ligate into pSK7 in two possible orientations. It was determined through sequencing that the ORF was in the same orientation as *rpoA*.

Introduction of *rpoC* under control of its own T7 promotor into pSK8 required the PCR amplification *rpoC* from pSK2. PCR primers were designed that contained the *sgr*AI found upstream of the T7 promoter and an engineered *sgr*AI site downstream of the *rpoC* stop codon. The primers used were "*rpoC*-T7-SgrAI for" and "*rpoC*-T7-SgrAI rev". The *T7::rpoC* PCR fragment was treated with *sgr*AI and ligated into pSK8 treated with the same enzyme, yielding pSK9. *T7::rpoC* is able to ligate into pSK7 in two possible orientations. It was determined through sequencing that the ORF was in the same orientation as *rpoA* and *rpoZ*. pSK9 contains all the essential genes required for expression of each subunit under control of its respective T7 promoter.

5.2 Mutagenesis of RNA Polymerase

Advancements in site-directed mutagenesis over the past 20 years have made it relatively easy to form mutant forms of most target proteins. Unfortunately, not all targets are equally easy to mutate. The fact that RNAPs found in all cellular life are

found to be multi-subunit presents a unique challenge in performing mutagenesis on them. Other additional challenges involved in the mutagenesis of RNAP involve the percent G/C content of target genes and the large sizes of *rpoB* and *rpoC* (which encode for the β and β' subunit, respectively).

The proper design of primers used in site-directed mutagenesis is critically important for successful results. Primers should be, on average, 22-27nt in length with a minimum length of 19nt. The %G/C-content is another important factor to take into account when designing primers. The G/C content of the primer should be no more than 50% of the overall sequence. Furthermore, it is recommended that both the 5' and 3'-ends of the primer have at least one, but no more than three guanosine or cytidine bases in a row.

Two important aspects of primer design involve melting temperature (T_m) and secondary structure formation. The T_m of annealing for each primer was calculated by the following equation: $T_m = -16.6 \log [S] + 41.5(\chi_{GC}) + 81$ where T_m is the melting temperature, S is the concentration of monovalent salt, and χ_{GC} is the molar fraction of GC base pairs in the DNA (Rees, Yager *et al*. 1993). Secondary structure of single stranded DNA can significantly affect how well mutagenesis works. Secondary structure stability is dependent on sequence, primer length, %GC content, and buffer salt concentration, therefore it is advisable to determine the secondary structure of any primers that are >30nt in length or have a GC content of >50%. Secondary structure calculation were using the mFold server v3.2 (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/dna-form1.cgi) (Zuker 2003). Primers that were designed so that the T_m for the calculated secondary structure was

~10ºC less than the annealing T_m of the primer were found to be the most successful.

5.2.1 Site-directed Mutagenesis of pRL706 to Form Δ-loop RNAP Mutant

pRL706 (Gift of R. Landick) is a pTrc99c derived expression plasmid which contains the *rpoB* gene (Amann, Ochs *et al*. 1988; Severinov, Mooney *et al*. 1997). The total size of pRL706 is ~7.8kb. Due to the relatively large size of pRL706, mutagenesis, especially multi-residue deletions, becomes difficult. Several attempts were made to make the β-(ΔR542-F545) RNAP mutant (Δ-loop RNAP) in one step, but were unsuccessful. Therefore, in order to form the four residue deletion, it was required to break the process into four separate mutagenesis reactions in which one residue was deleted at a time. The residues were deleted in the following order: 1)ΔF545 2)ΔG544 3)ΔR542 4)ΔA543. The deletions were performed in this order to alleviate the need to design mutagenesis primers with excessive secondary structure. All mutagenesis reactions were performed with a QuikChange XL Site-directed Mutagenesis Kit (Stratagene). The sequence of the primers used in these mutagenesis reactions are presented in Table 5.2. The optimal T_m range for primers used in the mutagenesis of *E. coli rpoB* was found to be 55-58ºC. For each reaction, ~20ng of pRL706 DNA and ~125ng of each primer were used. Each mutagenesis reaction was digested with *Dpn*I and transformed into the *E. coli* cloning strain DH5α using standard biochemical techniques.

5.2.2 Site-directed mutagenesis of RpoB in pIA509 to form β-(D446A/G449A)

pIA509 (Gift of I. Artsimovitch) is a pET21 derived expression plasmid that

Primer Name	Primer Sequence
rpoB-ΔF545 for	5’-GAA CGT GCA GGC GAA GTT CGA GAC
rpoB-ΔF545 rev	5’-GTC TCG AAC TTC GCC TGC ACG TTC
rpoB-ΔG544-F545 for	5’-CGT GAA CGT GCA GAA GTT CGA GAC G
rpoB-ΔG544-F545 rev	5’-CGT CTC GAA CTT CTG CAC GTT CAC G
rpoB-ΔR542, G544-F545 for	5’-GTC TGA CCC GTG AAG CAG AAG TTC G
rpoB-ΔR542, G544-F545 rev	5’-CGA ACT TCT GCT TCA CGG GTC AGA C
rpoB-ΔR542-F545 for	5’-GTC TGA CCC GTG AAG AAG TTC GAG AC
rpoB-ΔR542-F545 rev	5’-GTC TCG AAC TTC TTC ACG GGT CAG AC

Table 5.2 Mutagenesis Primers Used in to Make Δ-loop RNAP

contains a polycistronic operon of *rpoA-rpoB-rpoC* under the control of a single T7 promoter and terminator and each ORF is preceded by its own ribosomal binding site (Artsimovitch, Svetlov *et al*. 2003; Artsimovitch, Vassylyeva *et al*. 2005). Because this plasmid contains each gene coding for core RNAP, it has a size of ~14.8kb.

We were interested in performing site-directed mutagenesis on *rpoB* to make the β-(D446A/G449A) mutant RNAP (Walker-RNAP). Due to the large size of this plasmid site-directed mutagenesis would be non-trivial to perform. To overcome this problem, *rpoB* was excised from pIA509 with the *Nco*I and *Sbf*I endonucleases and subcloned into pCDFDuet (Novagen) treated with the same enzymes and, additionally, CIP. The resulting plasmid, pCDF-Duet-*rpoB*, is ~7.8kb in size. Site-directed mutagenesis was able to be performed as previously described using standard techniques. The sequence of the primers are as follows: forward: 5'-GATGATATCGCACACCTCGCCAACCG-3', reverse: 5'CGGTTGGCGAGGTGTGCGATATCATC-3'. After confirming the mutations, *rpoB* was excised from pCDFDuet-*rpoB* with NcoI and SbfI and ligated back into pIA509 from which is it originated.

5.2.3 Introduction of a SECIS Element into E. coli rpoB

We were interested in labeling RNAP with a fluorescent tag for the purposes of performing single-molecule FRET. Previous work on labeling proteins has focused on the labeling of endogenous cysteine residues with fluorescent dyes functionalize with the thiol-reactive maleimide group. A drawback to this technique can only be used with proteins that contain only 1-2 solvent accessible cysteine residues. Unfortunately, RNAP contains 11 cysteines; most of which are solvent accessible, therefore this method cannot

be used and an alternate method must be found.

To overcome the multi-cysteine problem, a technique was developed that involves the selective incorporation of a non-native selenocysteine into the protein (Arnér, Sarioglu *et al*. 1999). Under acidic conditions and a reducing environment, seleno-cysteine has a much higher affinity for thiol-reactive groups than a native cysteine which can be selectively labeled with a fluorescent dye (Johansson, Chen *et al*. 2004). Therefore, under optimal conditions, it may be possible to selectively label RNAP and perform single-molecule FRET.

In order to selectively incorporate a selenocysteine into RNAP, a SECIS element (Figure 5.2) was introduced into the β-subunit dispensable region (residues 200-368) via mutagenesis. A minimal SECIS element was calculated from the coding sequence of *rpoB* using the SECISDesign server (Liu, Reches *et al*. 1998; Busch, Will *et al*. 2005). The amino acid sequence was changed from a.a. 202-RKLPATIILRALNY-215 to a.a. 202-RKLPVAGLLQALNY-215. The gene sequence was altered from 5'-CGCCGTAAACTGCCTGCGAACATCATTGTGCGCGCCC-3' to 5'-TGACGTAAACTGCCGGTAGCAGGTCTGCTCCAGGCCC-3', where the underlined bases are the ones mutagenized. The mutagenesis reactions were carried out on pRL706 as previously described for the Δ-loop mutagenesis. Introduction of the SECIS element in one step was not possible due to the large number of changes required for the gene sequence to be correct, therefore, the mutagenesis reactions were split into six sequential reactions with each step being dependent on the previous reaction. The sequences of the primers used for these mutagenesis reactions are listed in Table 5.3. The final plasmid was named pRL706-SECIS.

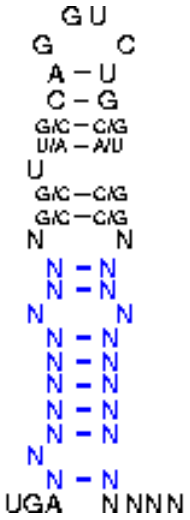


Figure 5.2 General Structure of a SECIS Element. When the SECIS element is transcribed, the RNA folds into this generalized structure. The SECIS element introduced into *rpoB* has a sequence that corresponds to the black labeled bases. 'N' represents any nucleotide. The blue region is not required for proper functioning of the SECIS element.

Primer Name	Primer Sequence
rpoB-SECIS-1 for	5’-CGT AAA CTG CCT GTA GCA ATC ATT GTG CG
rpoB-SECIS-1 rev	5’-CGC ACA ATG ATT GCT ACA GGC AGT TTA CG
rpoB-SECIS-2 for	5’-GTA GCA ATC ATT CTC CAG GCC CTG AAC TAC
rpoB-SECIS-2 rev	5’-GTA GTT CAG GGC CTG GAG AAT GAT TGC TAC
rpoB-SECIS-3 for	5’-CCT GTA GCA GGT ATT CTC CAG GCC CTG
rpoB-SECIS-3 rev	5’-CAG GGC CTG GAG AAT ACC TGC TAC AGG
rpoB-SECIS-4 for	5’-GCA GGT CTG CTC CAG GCC CTG AAC
rpoB-SECIS-4 rev	5’-GTT CAG GGC CTG GAG CAG ACC TGC
rpoB-SECIS-5 for	5’-GTA TCG ACC GTT GAC GTA AAC TGC CTG
rpoB-SECIS-5 rev	5’-CAG GCA GTT TAC GTC AAC GGT CGA TAC
rpoB-SECIS-6 for	5’-CCG TTG ACG TAA ACT GCC GGT AGC
rpoB-SECIS-6 rev	5’-GCT ACC GGC AGT TTA CGT CAA CGG

Table 5.3 Mutagenesis Primers Used to Insert a SECIS element *rpoB* (SECIS-RNAP)

5.2.4 Site-directed Mutagenesis of RpoB in T. thermophilus

Performing site-directed mutagenesis on DNA from a thermophilic organism provides unique challenges. Any nucleic acid is stabilized with an increased G/C content. *T. thermophilus*, which lives at 82ºC, has a genomic G/C content of ~69%, which is significantly higher than *E. coli* (Nagahari, Koshikawa *et al*. 1980). To increase the challenges of site-directed mutagenesis, there are localized DNA sequences in *rpoB* that have >80% G/C content.

Attempts were made to remove engineered internal *Eco*RI site in *rpoB* of pSK4 using a standard QuikChange XL Site Directed Mutagenesis kit (Stratagene) with the following primers: forward: 5'-GGGCCAGGAAGCCGTTCTCG-3', reverse: 5'-CGAGAACGGCTTCCTGGCCC-3', but were unsuccessful. The difficulty was presumably resulted from the high localized G/C content of the *rpoB*. It has been shown that the addition of betaine (N,N,N-trimethylglycine) to DNA eliminates the base pair composition dependence on the thermal denaturation of DNA; effectively, betaine makes the stability of a A:T and G:C base pairs equally stable (Rees, Yager *et al*. 1993; Henke, Herdel *et al*. 1997); therefore, 1.5M betaine was added to the site-directed mutagenesis reactions using the same QuikChange XL Site Directed Mutagenesis kit. The optimal T_m range for primers used in the mutagenesis of *T. thermophilus rpoB* and was found to be 58-62ºC, which is slightly higher than for *E. coli*. Presumably this is because DNA from thermophilic organisms is relatively more stable than non-thermophilic organisms at the same temperature; therefore, higher temperatures are required for proper primer annealing to the DNA.

5.3 Expression and Purification of Recombinant *E. coli* RNA Polymerase

The original purification method was worked out by (Burgess and Jendrisak 1975); however, this purification was for wild type genomic RNAP. This work was further extended to allow purification for his_6-tagged RNAP (Uptain and Chamberlin 1997; Santangelo, Mooney *et al*. 2003). The purification presented here is a modification of the published protocols and selects for RNAP with a mutant β-subunit. This purification method can be extended to any RNAP subunit with a his_6-tag or other affinity tag.

5.3.1 Expression of Mutant E. coli RNA Polymerase in a Wild-type Background

Expression of Δ-loop RNAP was done in the *E. coli* strain TOM100 (*Δara, Δlac, malP::lacI*^{q}, *recA*$^-$*[recA::Kan*^{R}*]*) (Gift of T. Santangelo) (Santangelo, Mooney *et al*. 2003). Strain TOM100 was used for two reasons: 1) expression of *rpoB* in pRL706 is under the control of the IPTG inducible *trc* promoter and not the often used T7 promoter, therefore a strain containing the λ DE3 lysogen is not needed 2) the strain is *recA*$^-$, which will reduce *rpoB* reverting back to wild-type by homologous recombination.

All bacterial cultures were done in super-broth growth media (tryptone=25g/L, yeast extract=15g/L, NaCl=5g/L) and carbenicillin (100μg/mL). A typical purification requires 6L of bacterial culture for enough cells to perform. Bacterial cultures were never grown in volumes exceeding 2L per 6L flask. Each flask of growth media was inoculated with 10mL of overnight culture and allowed to grow at 37ºC with agitation. When cultures reached an OD_{600}≈0.7-0.8, they were induced with 1mM IPTG. It is important to wait until late log-phase before inducing to reduce the possibility of cyto-

toxicity. After induction, the cultures were incubated with agitation for 4 h and then pelleted and stored at -80ºC until ready for use. A typical yield is 4-5g of wet cell paste per liter of culture.

Expression of Walker-RNAP was done in the *E. coli* strain HMS174(DE3) [*F*$^-$ *recA1 hsdR* ($r_{K12}^-m_{K12}^+$) *rif*^{r} *λDE3*]. Strain HMS174(DE3) (Novagen) was used for two reasons. The first is that pIA509 is a pET-21 derived plasmid and, therefore, under the control of the T7 promoter. Because of this, expression requires a strain that carries the λ DE3 lysogen. Secondly, HMS174(DE3) is a *recA*$^-$ strain, which will reduce rpoB reverting back to wild-type due to homologous recombination. Expression of Walker-RNAP was carried out as previously described for Δ-loop RNAP.

5.3.2 Expression of SECIS E. coli RNA Polymerase

Expression of SECIS RNAP was done in the *E. coli* strain TOM100 co-transformed with pRL706-SECIS and pSUABC, which contains *selA*, *selB*, and *selC* genes required for selenocysteine incorporation. A total of 8L of culture were grown for each purification. Each 2L of culture was inoculated with 10mL of overnight culture and allowed to grow at 37ºC with agitation. After ~2.5 h of incubation (a minimum of 1h before inoculation), sodium selenite and cysteine were added to the cultures at a final concentration of 5μM and 0.1mM, respectively. The cultures were incubated until OD_{600}≈1.0 was reached and then induced with 1mM IPTG. Cells were incubated for 5 h and harvested by centrifugation. Typical yields were found to be 3-4g of wet cell paste per liter of culture. Purification of SECIS-RNAP follow the same procedure as described for Δ-loop RNAP and Walker-RNAP.

5.3.3 Purification Mutant E. coli RNA Polymerase in a Wild-type Background

Pelleted cells are resuspended in 3mL of Lysis Buffer (50mM Tris-HCl pH 8.0, 10mM EDTA, 5% glycerol (v/v), 1mM DTT, 300mM NaCl) per gram of cell paste. It is critical that the resuspended cells are as homogeneous as possible. Lysozyme is added to the resuspended cells at a concentration of 300μg/mL and incubated at 4ºC with gentle agitation for 20 min. After 20 min sodium deoxycholate is added to a final concentration of 0.2% (w/v). Sodium deoxycholate is a biological detergent that helps to further lyse cells and solubilize cellular and membrane components. The cells are further incubated at 4ºC with gentle agitation for 20 min. After incubation, the resuspended cells will be very viscous.

Sonication was used to completely lyse the cells and help break the DNA into small fragments that will more likely to stay in solution. Proper sonication is critical for increasing the purification yield. Cell lysate was sonicated in small aliquots of no more than 40mL. To avoid overheating, samples were sonicated at on ice for 10s intervals followed by 10s of no sonication. The cycling was repeated for a total of 2 min per aliquot. After sonication, the cell lysate had a similar viscosity to water.

Cellular debris was pelleted by centrifugation in a Beckman GS-3 rotor at 10,000 rpm for 45 min. The supernatant should be yellow in color and semi-transparent. A 10% stock solution (v/v) of polyethylenimine (pH 7.9) was added to the cleared lysate to a final concentration of 0.8% (v/v). Polyethylenimine is used to precipitate DNA from solution, which will pull RNAP out of solution with it. After addition, the solution became flocculent and was allowed to incubate with slow stirring for 20 min.

After incubation, the precipitated DNA and associated protein was pelleted using

the GS-3 rotor at 10,000 rpm for 15 min. The supernatant was decanted and the resulting pellet was resuspended in 100mL of TGED+0.4M NaCl buffer (TGED= 10mM Tris-HCl (pH 8.0), 5% glycerol (v/v), 1mM EDTA, 0.5mM DTT) to remove contaminating DNA binding proteins. Complete resuspension of the pellet is required for efficient washing of the pellet. Homogeneous resuspension is best achieved using a mortar and pestle followed a tissue grinder. The solution was pelleted as previously described. The wash steps were repeated for a total of 4-5 times.

To elute the RNAP from the DNA, the pellet was washed with 100mL of TGED+1M NaCl. The polymin P and associated DNA was pelleted using the GS-3 rotor at 10,000 rpm for 45 min. The supernatant, which contains soluble RNAP, was decanted and the RNAP was precipitated out of solution by the addition of $(NH_4)_2SO_4$ to a final concentration of 35% (w/v). The solution was allowed to incubate for a minimum of 1h prior to pelleting using the SS-34 rotor at 15,000 rpm for 1h.

The supernatant from the $(NH_4)_2SO_4$ precipitation was removed and the remaining protein pellet was resuspended in a small volume of TGED buffer (typically 5-10mL) and allowed to dissolve into solution with gentle agitation. Additional TGED buffer was added to until the conductivity was less than the conductivity of TGED+0.2M NaCl (The total volume was typically 30-35mL and gave a conductivity of ~13 mS/cm at 25ºC). Having a conductivity that is higher than TGED+0.2M will keep reduce the affinity of RNAP to the heparin column and dramatically reduce the final yield of the purification.

All column purifications were performed on an ÄKTA-FPLC (Amersham Biosciences). The heparin column used was a 5mL HiTrap Heparin HP (Amersham

Biosciences). This step separates out DNA binding proteins from each other based on their binding affinity to heparin. The column was equilibrated in TGED+0.3M NaCl. The protein sample was loaded at a flow rate of 1ml/min and then washed with TGED+0.3M NaCl for a total of 10 column volumes (50mL) or until the absorbance reached close to baseline. After washing, RNAP was eluted from the column with an isocratic cut to TGED+0.6M NaCl and collected in 1mL fractions. RNAP typically elutes from the column at fractions 36-59. Interestingly, the peak corresponding to RNAP elution usually has a shoulder indicating that two species are eluting at slightly different points. It was found by SDS-PAGE that these two convoluted peaks correspond to core and holoenzyme; therefore, it may be possible to further separate these two peaks by forming an elution gradient from 0.3M-0.6M NaCl. It should also be noted that a peak usually elutes around fractions 19-26 but this peak does not contain protein and can safely be ignored.

The presence of protein in each fraction was confirmed by adding 10μL of sample to 100μL Bradford solution. Fractions containing protein were pooled and dialyzed overnight in TG+1M NaCl buffer (TG=20mM Tris-HCl (pH 8.0), 5% glycerol (v/v)) at 4ºC. After dialysis, the pooled fractions were loaded onto a 5mL HisTrap HP (Amersham Biosciences) nickel column equilibrated with TG+1M NaCl buffer. The column was washed with 10 column volumes (50mL) of TG+1M NaCl buffer to remove wt-RNAP and other extraneous proteins. His_6-tagged RNAP was eluted with an isocratic cut to TG+1M NaCl+150mM imidazole and collected in 1mL fractions. RNAP typically elutes from the column at fractions 26-36. The peak usually seen at fractions 19-23 is not protein and its composition is unknown and can be safely ignored. Each fraction was

tested for the presence of protein in the same method as previously described. Fractions that contained protein were pooled.

At this point, the pooled volume is typically about 10-12mL and, therefore, the RNAP concentration is very dilute. RNAP was concentrated by a 35% $(NH_4)_2SO_4$ precipitation (w/v). RNAP was pelleted using the SS-34 rotor at 15,000 rpm for 1h. After pelleting, the supernatant was carefully removed and the pellet is dissolved in 0.75-1mL TG+1M NaCl buffer. The protein sample was then dialyzed overnight into storage buffer (50mM Tris-HCl (pH 8.0), 50% glycerol (v/v), 1mM EDTA, 1mM DTT, 100mM NaCl). A typical yield is 500-700μL of ~1-2μM RNAP. RNAP should be stored at -80ºC until ready to use. Additionally, if desired, a 2-5-fold molar excess of σ-subunit can be added to the final preparation in order to saturate the formation of holoenzyme.

BIBLIOGRAPHY

Amann, E., B. Ochs, *et al*. (1988). "Tightly Regulated *tac* Promoter Vectors Useful for the Expression of Unfused and Fused Proteins in *Escherichia coli*." Gene **69**: 301-315.

Arnér, E. S. J., H. Sarioglu, *et al*. (1999). "High-level Expression in *Escherichia coli* of Selenocysteine-containing Rat Thioredoxin Reductase Utilizing Gene Fusions with Engineered Bacterial-type SECIS Elements and Co-expression with the *selA*, *selB* and *selC* Genes." J. Mol. Bio. **292**(5): 1003-1016.

Artsimovitch, I., V. Svetlov, *et al*. (2003). "Co-overexpression of Escherichia coli RNA Polymerase Subunits Allows Isolation and Analysis of Mutant Enzymes Lacking Lineage-specific Sequence Insertions." J. Bio. Chem. **278**(4): 12344-12355.

Artsimovitch, I., M. N. Vassylyeva, *et al*. (2005). "Allosteric Modulation of the RNA Polymerase Catalytic Reaction Is an Essential Component of Transcription Control by Rifamycins." Cell **122**(3): 351-363.

Burgess, R. R. and J. J. Jendrisak (1975). "A Procedure for the Rapid, Large-Scale Purification of Escherichia coli DNA-Dependent RNA Polymerase Involving Polymin P Precipitation and DNA-Cellulose Chromatography." Biochem. **14**(21): 4634-4638.

Busch, A., S. Will, *et al*. (2005). "SECISDesign: A Server to Design SECIS-elements Within the Coding Sequence." Bioinfo. **21**(15): 3312-3313.

Henke, W., K. Herdel, *et al*. (1997). "Betaine Improves the PCR Amplification of GC-rich DNA Sequences." Nuc. Acids Res. **25**(19): 3957-3958.

Johansson, L., C. Chen, *et al*. (2004). "Exploiting the 21st Amino Acid-Purifying and Labeling Proteins by Selenolate Targeting." Nature Meth. **1**(1): 61-66.

Liu, Z., M. Reches, *et al*. (1998). "The Nature of the Minimal 'Selenocysteine Insertion Sequence' (SECIS) in *Escherichia coli*." Nuc. Acids Res. **26**(4): 896-902.

Nagahari, K., T. Koshikawa, *et al*. (1980). "Cloning and Expression of the Leucine Gene from *Thermus thermopbilus* in *Escherichia coli*." Gene **10**(1): 137-145.

Rees, W. A., T. D. Yager, *et al*. (1993). "Betaine Can Eliminate the Base Pair Composition Dependence of DNA Melting." Biochem. **32**(1): 137-144.

Santangelo, T. J., R. A. Mooney, *et al*. (2003). "RNA Polymerase Mutations that Impair Conversion to a Termination-Resistant Complex by Q Antiterminator Proteins." Genes Dev. **17**(10): 1281-1292.

Severinov, K., R. Mooney, *et al*. (1997). "Tethering of the Large Subunits of *Escherichia coli* RNA Polymerase." J. Bio. Chem. **272**(39): 24137-24140.

Shen, W.-H. and B. Hohn (1992). "DMSO Improves PCR Amplification of DNA with Complex Secondary Structure." Trends Gene. **8**(7): 227.

Uptain, S. M. and M. J. Chamberlin (1997). "*Escherichia coli* RNA Polymerase Terminates Transcription Efficiently at Rho-independent Terminators on Single-stranded DNA Templates." Proc. Nat. Acad. Sci. **94**(25): 13548-13553.

Vassylyev, D. G., S. Sekine, *et al*. (2002). "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution." Nature **417**(6890): 712-719.

Xue, Y., B. P. Hogan, *et al*. (2000). "Purification and Initial Characterization of RNA Polymerase from *Thermus thermophilus* Strain HB8." Biochem. **39**(46): 14356-14362.

Zuker, M. (2003). "Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction." Nuc. Acids Res. **31**(13): 3406-3415.