INSIGHT INTO THE FIDELITY OF *ESCHERICHIA COLI* RNA POLYMERASE: INVESTIGATION OF MISINCORPORATION DURING TRANSCRIPTION ELONGATION UTILIZING TRANSIENT STATE KINETICS

by
Candice Kermitta Cunningham

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

Reader: Dr. Linda L. Spremulli

Reader: Dr. Thomas A. Kunkel

ABSTRACT

CANDICE K. CUNNINGHAM: Insight into the Fidelity of *Escherichia coli* RNA Polymerase: Investigation of Misincorporation During Transcription Elongation Utilizing Transient State Kinetics
(Under the direction of Dr. Dorothy A. Erie)

Concentration-dependent pre-steady state kinetics of correct nucleotide incorporation led to a proposed mechanism for transcription involving multiple conformational states of RNA polymerase (RNAP). Specifically, RNAP can exist in an unactivated state or an activated state. Transition between the two states is driven by conformational changes in RNAP following templated NTP binding to an allosteric site. Further investigation led to a structural model, where the movement of the allosteric site upon NTP binding facilitates translocation of the enzyme via a ratchet motion.

In this work, I use transient state kinetics to investigate the NTP concentration-dependence of misincorporation (UMP for CMP). I demonstrate misincorporation occurs only in the activated state while a subset of complexes enters into a non-productive unactivated state. Complexes in the non-productive state are "trapped" by an incorrect NTP bound in the catalytic site. I demonstrate the non-productive and "irreversibly" bound NTP is removed from the catalytic site in the presence of the correct NTP. Combining these data with structural analyses, I present a structural model for misincorporation similar to the model for correct incorporation with several key differences. I also characterize the concentration-dependent misincorporation kinetics for Δ-loop RNAP with residues R542-F545 deleted from fork loop 2, the proposed allosteric site. Deletion of the four residues enhances the fidelity of RNAP, suggesting fork loop 2 is an allosteric site responsible for the

fast phase of synthesis during transcription elongation.

Correct and incorrect incorporation kinetic assays using RNAP with mutations in the secondary channel demonstrate that βD675Y (*E.coli*) RNAP is a low fidelity variant, significantly increasing the amount of misincorporation when initiated from the promoter. I demonstrate βD765Y RNAP exhibits a higher fidelity from purified complexes, suggesting that the experimental procedure affects the fidelity of this variant RNAP. I also reveal a zero-order dependence on the apparent rate of misincorporation with a continual increase in the extent of misincorporation for [UTP] < 75μM in βD675Y RNAP. Considering recent crystal structures of RNAP II and *T. Thermophilus* RNAP, I posit βD675Y affects the closing of the trigger loop over the active site, thereby changing the misincorporation kinetics of the βD675Y RNAP.

For Logan and Emily, who continually remind me that it's really the simple things in life that make it worth living…

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Dorothy Erie, whose enthusiasm for science continually reignites my own passion for the subject. Above all else, by her example she has taught me the importance of integrity as a scientist and in doing so, has taught me about the scientist I want to be. I would like to thank the rest of my committee, Thomas Kunkel, Linda Spremulli, Nancy Thompson, and Dick Wolfenden for continued support through the years. Many thanks go to Shannon Holmes who welcomed me and trained me when I first joined the lab. She has continued to offer her encouragement and insight even after leaving UNC. I thank the lab for always keeping things interesting and for making this experience one I will never forget. In particular, I would like to thank Scott Kennedy for answering all of my seemingly endless questions, for offering helpful feedback when I needed it most, for insightful discussions about life and transcription, and for making me laugh when I didn't think I could.

On a personal note, I want to thank all of my friends, from around the world. In ways great and small, they have offered continual love and support that made this work possible. I would especially like to thank my best friend, Tina, who has been with me in this adventure of life for seventeen years. She has been a constant presence in all things good in my life and has supported me through times when I thought I could not keep going. There are truly no words to express my gratitude for her friendship. I would like to thank my little brother, Chuck, who has shared an invisible cord between his brain and mine for the past twenty-

seven years. He has always been there to make me laugh when I needed it most. He is also the reason I am blessed to be an aunt to two wonderful children, Logan and Emily, whose laughter can brighten my day even from 500 miles away. Last, but certainly not least, I want to thank my mom and dad who by the example that is the life they live have taught me the importance of hard work and dedication. Through the life and love they share, they have taught me that it is never okay to give up on something you want just because at times it becomes difficult or frustrating. They have continually encouraged me in every endeavor I have undertaken, they were proud of me when I couldn't be proud of myself, and they have loved me. They are more than my parents. They are my best friends and for this I will forever be grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3′	three prime end
5′	five prime end
~	approximately
°	degrees
>	greater than
<	less than
%	percent
α	alpha
Å	angstrom
A	adenosine monophosphate
AMP	adenosine-5’-monophosphate
AMP-CPP	α,β-methyleneadenosine 5’-triphosphate
ATP	adenosine-5’-triphosphate
β	beta
β′	beta prime
BSA	bovine serum albumin
°C	degrees Celsius
Ci	curie
C	cysteine
C	cytidine monophosphate
CMP	cytidine-5’-monophosphate

CTP	cytidine-5’-triphosphate
Δ	delta (deletion)
D	aspartate
ddH_2O	distilled, deionized water
DE	dead end
DE13	a specific DNA sequence described in detail in the text
DNA	deoxyribonucleic acid
DNAP	deoxyribonucleic acid polymerase
DTT	dithiothreitol
EC	elongation complex
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ES	enzyme-substrate complex
et al.	and others
F	phenylalanine
G	guanosine monophosphate
GTP	guanosine-5’-triphosphate
H	histidine
HEPES	4-(2-hydroxyethyl)-1 piperazineethanesulfonic acid
his	histidine
H. sapiens	*Homo sapiens*
I	isoleucine
in vitro	outside a living system

in vivo	inside a living system
k	rate constant
K	dissociation constant
K^+	potassium cation
λ	lambda
L	leucine
L	liter
M	molar
m	milli
μ	micro
μM^{-1}	inverse micromolar
Mg^{2+}	magnesium cation
m	milli
min	minutes
min^{-1}	inverse minutes
mol	mole
n	nano
N	asparagine
Ni^{+2}	nickel cation
NMP	nucleoside 5’-monophosphate
nt	nucleotide
NTP	nucleoside 5’-triphosphate
ω	omega

OH	hydroxyl group
p	plasmid designation
^{32}P	radioactive phosphorus-32
P_R	“rightward” promoter
PCR	polymerase chain reaction
pDE13	plasmid of the DE13 sequence
Pol II	RNA polymerase II
PP_i	pyrophosphate ion
Q	antitermination protein factor
ν	rate
R	arginine
RNA	ribonucleic acid
RNAP	ribonucleic acid polymerase
RNAP II	RNA polymerase II
S	substrate
s^{-1}	per second
S. cerevisiae	*Saccharomyces cerevisiae*
SEC	stalled elongation complex
sec	second
sec^{-1}	per second
σ	sigma
t	time

T. aquaticus	*Thermus aquaticus*
TEC	ternary elongation complex
T. thermophilus	*Thermus thermophilus*
UMP	uridine-5’-monophosphate
UTP	uridine-5’-triphosphate
V	valine
Y	tyrosine

CHAPTER 1: TRANSCRIPTION ELONGATION

Introduction

The process of synthesizing RNA from double-stranded DNA is known as transcription. Transcription is first in a series of events that leads to expression of the genetic information encoded within DNA. The enzyme responsible for carrying out transcription at reasonable rates and with high fidelity is known as RNA polymerase (RNAP). All cellular organisms make use of the multi-subunit RNAP to synthesize nearly all of the RNA in the cell. *Escherichia coli* RNAP has been well characterized and is similar to other prokaryotic and eukaryotic systems, making *E. coli* RNAP an ideal enzyme to study (Sweetser *et al.* 1987; Zhang *et al.* 1999; Cramer *et al.* 2001). RNAP core enzyme is approximately 450kDa in size and consists of 5 subunits: 2 α, β, β', and ω. Another subunit, σ, is required for promoter recognition to initiate transcription. The σ subunit with the core enzyme constitutes the holoenzyme (Sweetser *et al.* 1987). Within the active site of the enzyme, there is a Mg^{2+} cation that directly participates in phosphodiester bond formation, making this ion essential for transcription (Suh *et al.* 1992).

The process of transcription consists of four phases: open promoter formation, initiation, elongation, and termination. These phases are represented schematically in Figure 1.1. During open promoter formation, RNAP holoenzyme (core plus σ) binds to the promoter sequence found on the double stranded DNA. This binding of RNAP to the promoter leads to the melting of the double stranded DNA and subsequent formation of the

transcription bubble. Initiation is characterized by the binding of the first nucleotide to the RNAP and pairing with its complement on the template strand of the DNA. The enzyme remains at the promoter during initiation until approximately 6-9 nucleotides have been added to the growing RNA chain. Following successful synthesis of these 6-9 nucleotides, the σ subunit is released and the RNAP core enzyme escapes the promoter region such that transcription enters the elongation phase (Lewin 2000). During elongation, the ternary complex (RNAP, DNA template, nascent RNA chain) is kinetically stable and does not dissociate. As the enzyme moves along the DNA, the DNA is unwound at the front end of the transcription bubble, while the duplex is simultaneously rewound at the back. RNA is also displaced as a free polynucleotide chain. This process is totally processive, meaning that if the RNAP dissociates from the DNA at any time during transcription, RNAP core enzyme must rebind the σ initiation factor to rebind the promoter region of the DNA and begin synthesis anew (Landick 1999). Eventually, RNAP will come to the end of the gene being transcribed and enter the termination phase of transcription. During termination, the transcription bubble collapses as the RNA-DNA hybrid is disrupted. The DNA reforms the duplex state and the core enzyme and RNA are released. The studies presented here are focused on the elongation phase of transcription.

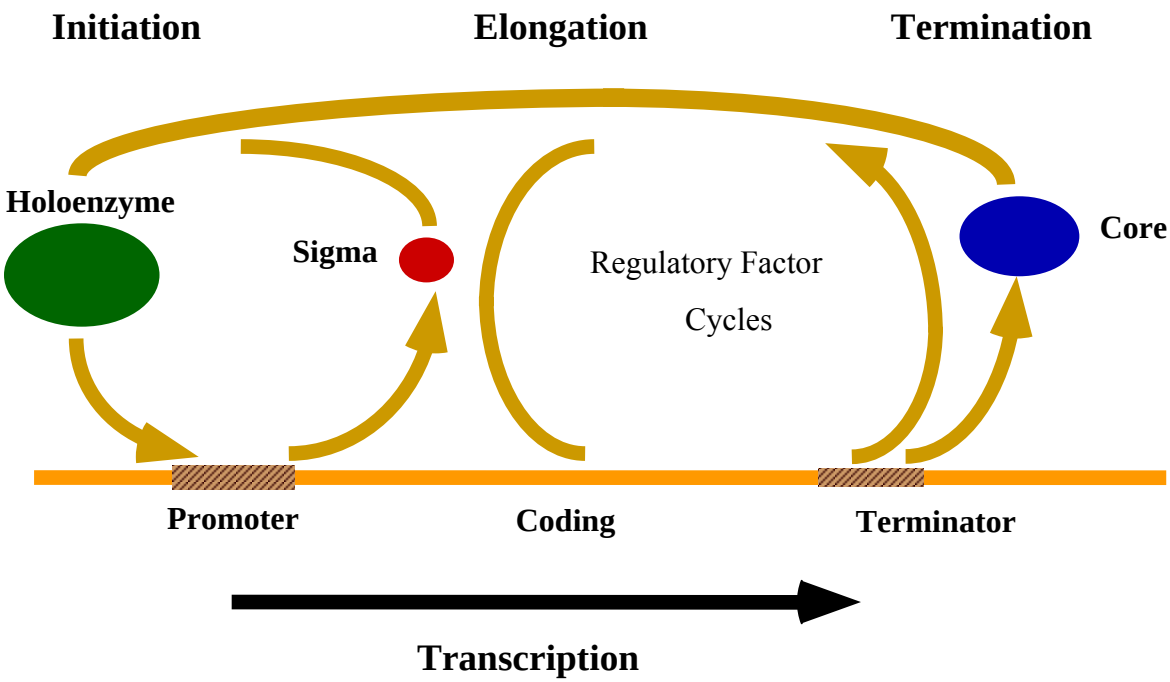


Figure 1.1: ***The transcription cycle.*** The holoenzyme (core plus subunit σ) is shown in green, the core (subunits $\alpha_2\beta\beta'\omega$) is shown in blue, and the sigma subunit (σ) is shown in red. The four phases of transcription (promoter binding, initiation, elongation, and termination) are illustrated.

Each nucleotide addition to the RNA chain involves a series of steps. First, the NTP to be added to the growing chain binds to the RNA polymerase. After binding, there is a phosphotransfer reaction at the α/β phosphodiester bond of the incoming NTP and the 3' hydroxyl of the nucleotide at the end of the RNA chain. Following the phosphotransfer reaction, pyrophosphate is released. Finally, the RNAP active site is translocated relative to the 3' end of the growing chain (Erie *et al.* 1992). The state prior to translocation is known as the pre-translocated state. Similarly, after translocation, the enzyme is said to be in the post-translocated state.

The entire process of melting, synthesizing, annealing, and displacing must occur at a reasonable rate. In *E. coli*, this rate on average is approximately 30-100 nucleotides per second during transcription elongation (Mooney *et al.* 1999). The process must also be carried out with a high fidelity (a low occurrence of incorrect nucleotide incorporation). Therefore, transcription by *E. coli* RNAP is one of the most highly regulated systems in the cell. Regulation occurs through the use of accessory proteins that bind the ternary complex, the DNA template, or the RNA transcript. Elongation is also regulated by particular sequence elements in the DNA or the RNA (e.g. pause sites, etc.) that interact with the RNAP to modulate the rate and/or fidelity of the enzyme during RNA synthesis.

RNA Polymerase Structure

Understanding of the structure of RNA polymerase has been significantly expanded in the past 10 years with the advent of high-resolution three-dimensional crystal structures from both prokaryotic and eukaryotic RNAPs. Specifically, several crystal structures of prokaryotic *Thermus aquacitcus* core (Zhang *et al.* 1999) and *Thermus thermophilus*

holoenzyme (Vassylyev *et al.* 2002) and eukaryotic *Saccharomyces cerevisiae* RNAP II core (Cramer *et al.* 2001) and *S. cerevisiae* RNAP II elongation complexes (Gnatt *et al.* 2001) have been published. More recently, RNAP II and *T. thermophilus* RNAP have been solved with the DNA, the RNA-DNA hybrid, and various NTPs and NTP analogs bound in the catalytic site (Cramer *et al.* 2001; Gnatt *et al.* 2001; Kettenberger *et al.* 2004; Westover *et al.* 2004; Wang *et al.* 2006; Vassylyev *et al.* 2007). These structures have provided tremendous insight into the structure/function relationship of transcription elongation complexes.

The overall structure of RNA polymerase resembles the shape of a crab claw with the two "jaws" of the claw formed by the two largest subunits of the enzyme, β and β' (Figures 1.2 and 1.3) (Zhang *et al.* 1999; Cramer *et al.* 2001). The catalytic site is located at the base of the cleft formed between the β and β' subunits. A magnesium ion that is required for synthesis is located in the catalytic site where the ion is chelated by a catalytic triad of three invariant aspartic acid residues (Zhang *et al.* 1999). The main channel of the enzyme, spanning the length of the crab claw, is 27Å in width. This channel houses the RNA-DNA hybrid (Figures 1.2 and 1.3) (Zhang *et al*. 1999; Gnatt *et al.* 2001). There are several other important structural elements of the RNAP that are located in the main channel. The bridge helix (F-helix) located in the β' subunit spans the main channel, abutting the RNA-DNA hybrid, and is thought to play a role in translocation (Figure 1.3) (Epshtein *et al.* 2002; Artsimovitch *et al.* 2003; Temiakov *et al.* 2005; Bar-Nahum *et al.* 2005; Tuske *et al.* 2005). Another structural element located between the main channel and the secondary channel is the trigger loop. The trigger loop is located under the bridge helix and is required for proper catalysis (Temiakov *et al.* 2005).

Another structural feature found in all structures of RNA polymerases is the secondary channel which leads directly into the catalytic site of the enzyme (Figures 1.2 and 1.3). The secondary channel is located on the β' subunit and is approximately 10-12Å in diameter at its narrowest point and 45Å in length, which makes the channel large enough to accommodate one diffusing NTP at a time (Zhang *et al.* 1999; Korzheva *et al.* 2000). This pore has been considered the primary means of NTP entry into the catalytic site for nucleotide binding and incorporation during transcription (Zhang *et al.* 1999; Korzheva *et al.* 2000; Cramer *et al.* 2000; Cramer *et al.* 2001; Gnatt *et al.* 2001; Vassylyev *et al.* 2002; Batada *et al.* 2004; Kettenberger *et al.* 2004; Westover *et al.* 2004; Armache *et al.* 2005; Temiakov *et al*. 2005). However, the size of the pore could potentially lead to a trafficking problem if all four NTPs must enter the catalytic site through this secondary channel. As such, other researchers have proposed that the primary pathway for NTP entry into the catalytic site is through the main channel (Nedialkov *et al.* 2003; Burton *et al.* 2005; Gong *et al.* 2005; Zhang *et al.* 2005). The secondary channel is also believed to play a role in regulation during transcription elongation. Specifically, the secondary channel is believed to function as an extrusion point of RNA during backtracking (Zhang *et al.* 1999; Artsimovitch & Landick 2000; Toulme *et al.* 2000). Backtracking is the process in which RNAP translocates backwards along the DNA template displacing the 3' end of the RNA transcript from the catalytic site (Reeder & Hawley 1996; Komissarova & Kashlev 1997; Nudler *et al.* 1997). The extrusion of the RNA through the secondary channel provides the substrate for GreA and GreB factor induced cleavage and thereby plays a role in the regulation of RNAP during transcription elongation (Komissarova & Kashlev 1997; Artsimovitch & Landick 2000; Toulme *et al.* 2000).

The RNA-DNA hybrid and melted DNA bubble lie in the main channel of the RNA polymerase between the rudder (upstream end) and the bridge helix and catalytic magnesium (downstream end) (Figure 1.3) (Korzheva *et al.* 2000). The bend angle between the downstream DNA and upstream duplex DNA is 90° (Figure 1.3) (Korzheva *et al.* 2000; Gnatt *et al.* 2001). The rudder is thought to maintain the upstream edge of the RNA-DNA hybrid by separating exiting RNA from the DNA, while fork loop 2 (βD loop I), another important structural element in the RNAP, has been suggested to maintain the downstream edge of the DNA bubble through stabilizing interactions with the DNA (Kettenberger *et al.* 2004; Vassylyev *et al.* 2007).

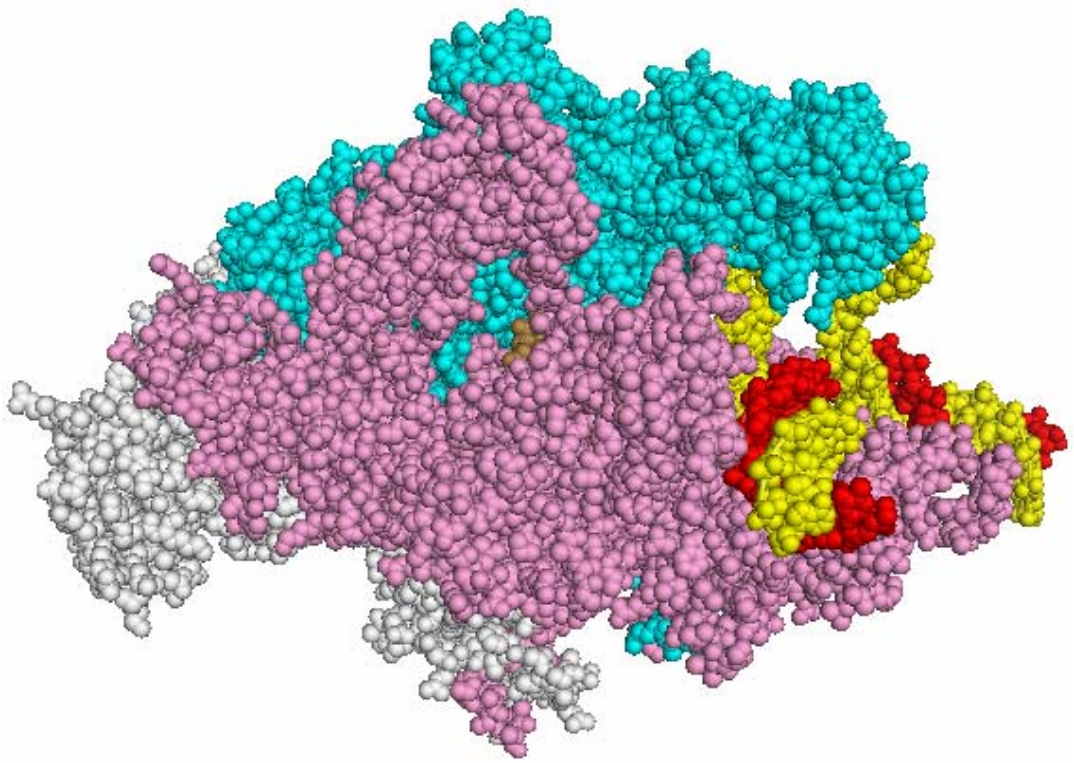


Figure 1.2: ***Space-filled model of the overall structure of T. aquaticus RNAP elongation complex.*** The overall structure of RNA polymerase resembles a crab claw shape with the RNA-DNA hybrid resting in the main channel formed between the two "jaws" of the claw. The β-subunit is shown in cyan, β'-subunit is pink, the two α-subunits and the ω-subunit is shown in white. The non-template strand of the DNA is shown in yellow, with the template strand of the DNA in red. The RNA is shown in brown and can be seen via the secondary channel located in the β'-subunit (PDB 1I6V).

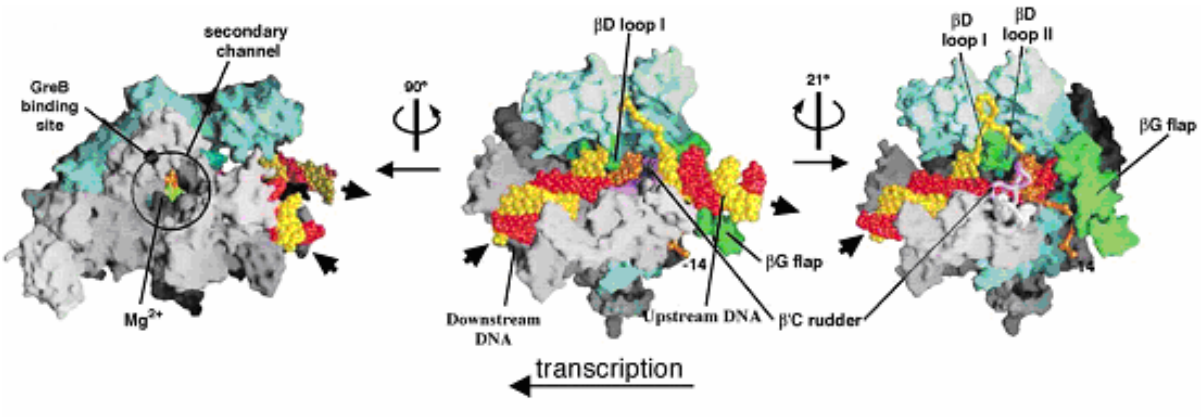


Figure 1.3: ***Model of the bacterial elongation complex based on T. aquaticus structure (Korzheva et al. 2000).*** The β-subunit is shown in cyan, β'-subunit is pink, the two α-subunits and the ω-subunit is shown in white. The non-template strand of the DNA is shown in yellow, with the template strand of the DNA in red. The RNA is shown in gold and can be seen via the secondary channel located in the β'-subunit. Numerous important structural features mentioned previously in the text are labeled.

Conformational States of the Elongation Complex

RNAP has been shown to exist in different states during transcription elongation (Erie *et al.* 1992; Erie *et al.* 1993; Matsuzaki *et al.* 1994; Kubori & Shimamoto 1996; Coulombe & Burton 1999; Yin *et al.* 1999; Davenport *et al.* 2000; Foster *et al.* 2001; Guthold & Erie 2001; Erie 2002; Tolic-Norrelykke *et al.* 2004). These different states result from conformational changes in the enzyme during transcription. Many of the important structural elements mentioned previously have been shown to exist in different conformations. The F-helix is seen in *S. cerevisiae* in a straight conformation (Cramer *et al.* 2000; Gnatt *et al.* 2001) while the F-helix is bent in the *T. thermophilus* holoenzyme (Vassylyev *et al.* 2002). The motions of the F-helix between the bent and straight conformations have been suggested to play a key role in translocation during transcription elongation (Epshtein *et al.* 2002; Artsimovitch *et al.* 2003; Temiakov *et al.* 2005; Bar-Nahum *et al.* 2005; Tuske *et al.* 2005). The trigger loop, required for proper catalysis, has also been seen in different conformations in both *S. cerevisiae* and *T. thermophilus* (Figure 1.4). The trigger loop is seen in an open conformation and a closed conformation. In the closed conformation, the trigger loop rests over the catalytic site and blocks access to the catalytic site via the secondary channel (Vassylyev *et al.* 2007). Additionally, fork loop 2 is seen in different conformations including an open, partially open and closed configuration (Figure 1.4).

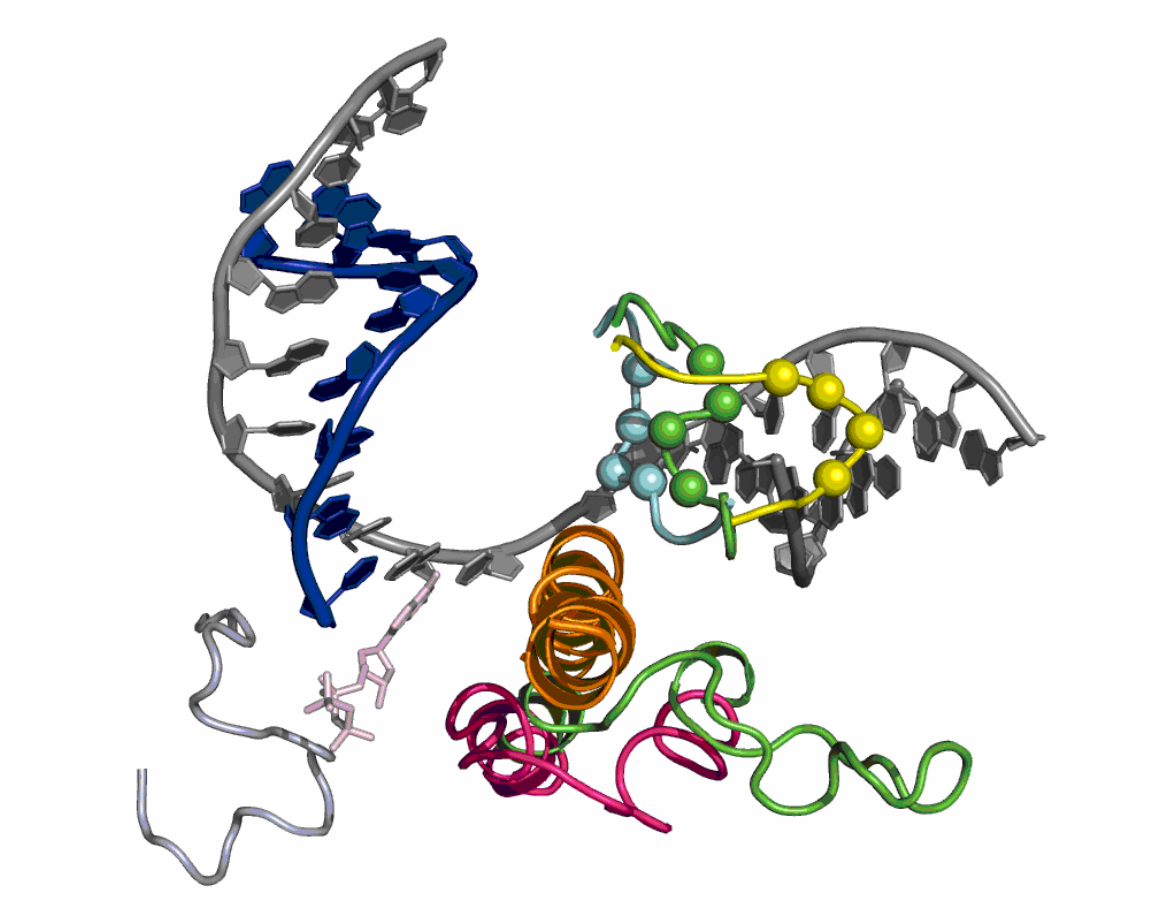


Figure 1.4: ***Different conformations of key structural elements in S. cerevisiae RNAP II.*** DNA template strand (grey), non-template strand (dark blue), and RNA (red) are from PDB 1Y77. The bound GTP (light purple) is from 2E2H. The bridge helix (orange) is from 1Y1V. The trigger loop exists in an "open" conformation (green, 1Y1V) and a closed conformation (magenta, 2E2H). Fork loop 2 is shown in three conformations: "open" (yellow, 2E2I), partially "closed" (green, 1Y1V), and "closed" (light blue, 1Y77) (adapted from Kennedy & Erie, manuscript in preparation).

RNAP is required to processively synthesize RNA at reasonable rates and with high fidelity. These stringent requirements make transcription by RNA polymerase the most highly regulated processes in gene expression. The different conformational states observed for RNAP elongation complexes play different roles in this regulation of transcription (Erie 2002). The different states of RNAP during transcription elongation are shown in Figure 1.5.

Primarily, RNAP exists in a long-lived activated state (n*). This activated state is characterized by rapid synthesis and low fidelity (Erie *et al.* 1993). RNAP can also exist in an unactivated state (n). RNAP in this state is capable of incorporating NTPs but synthesis is much slower than activated state synthesis. As such, the unactivated state is a higher fidelity state and is susceptible to regulation. From the unactivated state, RNAP can decay into states that are not capable of synthesis. These states, however, function in the regulation of transcription. In the hypertranslocated state (n_{hyper}), RNAP slips forward along the DNA template and the 3'-end of the RNA transcript becomes displaced from the catalytic site. In the backtracked state (n_{B1}, n_{B2}, and n_{B3}), RNAP translocates backwards along the DNA displacing the 3' end of the RNA. From the backtracked state, the enzyme can decay into cleavage states (n_{C1}, and n_{C2}) in which RNAP hydrolyzes the RNA transcript, creating a new 3'-end, or into arrest states (dead-end states) (n_{arrest}) in which elongation cannot be resumed even in the presence of high concentrations of all four NTPs. Table 1.1 summarizes the different conformational states and the accessory proteins that are capable of recognizing and acting on each of these states. The distribution of complexes between these states is regulated by many different factors, including the DNA, RNA, and the accessory proteins (Erie 2002).

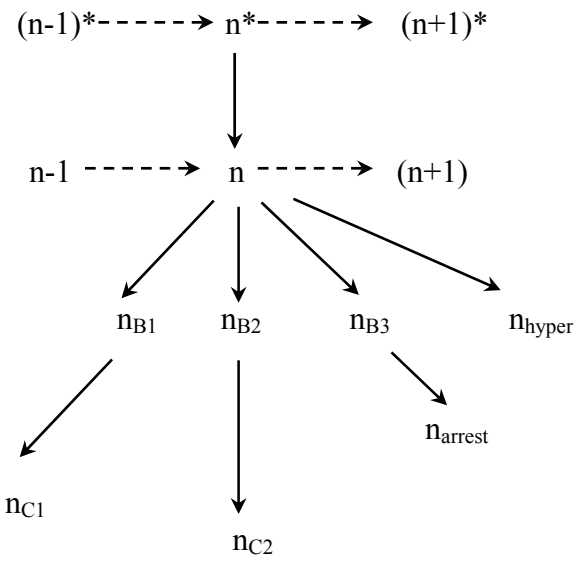


Figure 1.5: ***Different conformational states of the elongation complex (Erie 2002).*** n* is the activated state. After entering the unactivated state (n), complexes can undergo further conformational changes to backtracked states (n_B), cleavage states (n_C), arrest states (n_{arrest}), or hypertranslocated states (n_{hyper}). n* and n are synthesis states. n, n_B, and n_{arrest} are all regulatory states. n_C is a rescue state. The transitions between states are shown with single arrows for simplicity, however, each transition is essentially reversible.

State	Conformation	Synthesis/ Cleavage	Accessory proteins acting on state
Activated [fast] (n*)	Poised for catalysis	Synthesis	T4 alc termination protein
Unactivated [slow] (n)	Suboptimal conformation for catalysis	Synthesis	GreA, GreB, rho NusA, NusG, Gfh1
Backtracked (n_B)	RNAP reverse translocated on DNA, 3' end of RNA extruded from 2° channel	Cleavage (n_C)	GreA, GreB, Gfh1, NusG
Hypertranslocated (n_{hyper})	Active site slipped forward relative to 3' end	No synthesis/ no cleavage	NusA
Arrested (dead end) (n_{arrest})	Similar to back tracked state but no synthesis	No synthesis/ no cleavage	Reactivated by GreB cleavage

Table 1.1: ***Different conformational states of the elongation complex subject to synthesis and regulation (Erie 2002).*** The table summarizes the specific conformation of each state and whether or not it is a synthesis-competent state. Accessory proteins recognizing the different conformational states of the enzyme are also displayed.

A significant amount of the biochemical and structural information available for transcription elongation by RNAP was obtained through experiments examining the correct incorporation of a nucleotide into the growing RNA chain. This study returns the focus to misincorporation. Misincorporation, incorporation of an incorrect nucleotide, has been shown to occur at rates slower than that of correct incorporation (Erie *et al.* 1993). Capitalizing upon these slower rates may make it possible to investigate conformational states of elongation complexes that cannot be observed during correct incorporation. These conformational states are believed to be poorly populated during rapid synthesis but may still be physiologically important in regulation (Erie *et al.* 1993). Determining the rate limiting steps through the use of misincorporation studies should allow for the determination of the steps subjected to regulation. Specifically, we can use transient state kinetics to determine the steps in the incorrect nucleotide addition cycle. Determining the steps in misincorporation taken together with correct incorporation data and recently published crystal structures should provide insight into the regulation of transcription, specifically the fidelity of RNAP, and further our understanding of the process of transcription elongation.

Bibliography

Armache, K.J., Mitterweger, S., Meinhart, A. and Cramer, P. (2005). "Structures of complete RNA polymerase II and its subcomplex, Rpb4/7." J.Biol.Chem. **280**(8): 7131-7134.

Artsimovitch, I., Chu, C., Lynch, A.S. and Landick, R. (2003). "A new class of bacterial RNA polymerase inhibitor affects nucleotide addition." Science **302**(5645): 650-654.

Artsimovitch, I. and Landick, R. (2000). "Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals." Proc.Natl.Acad.Sci.U.S.A. **97**(13): 7090-7095.

Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A. and Nudler, E. (2005). "A ratchet mechanism of transcription elongation and its control." Cell **120**(2): 183-193.

Batada, N.N., Westover, K.D., Bushnell, D.A., Levitt, M. and Kornberg, R.D. (2004). "Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center." *Proc.Natl.Acad.Sci.U.S.A.* **101**(50): 17361-17364.

Burton, Z.F., Feig, M., Gong, X.Q., Zhang, C., Nedialkov, Y.A. and Xiong, Y. (2005). "NTP-driven translocation and regulation of downstream template opening by multi-subunit RNA polymerases." Biochem.Cell.Biol. **83**(4): 486-496.

Coulombe, B. and Burton, Z.F. (1999). "DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms." Microbiol.Mol.Biol.Rev. **63**(2): 457-478.

Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001). "Structural basis of transcription: RNA polymerase II at 2.8Å resolution." Science **292**(5523): 1863-1876.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R. and Kornberg, R.D. (2000). "Architecture of RNA polymerase II and implications for the transcription mechanism." Science **288**(5466): 640-649.

Davenport, R.J., Wuite, G.J., Landick, R. and Bustamante, C. (2000). "Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase." Science **87**(5462): 2497-2500.

Epshtein, V., Mustaev, A., Markovtsov, V., Bereshchenko, O., Nikiforov, V. and Goldfarb, A. (2002). "Swing-gate model of nucleotide entry into the RNA polymerase active center." Mol.Cell **10**(3): 623-634.

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): 224-39.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C. and von Hippel, P.H. (1993). "Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription." Science **262**(5135): 867-873.

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

Foster, J.E., Holmes, S.F. and Erie, D.A. (2001). "Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation." Cell **106**(2): 243-252.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A. and Kornberg, R.D. (2001). "Structural basis of transcription: an RNA polymerase II elongation complex at 3.3Å resolution." Science **292**(5523): 1876-1882.

Gong, X.Q., Zhang, C., Feig, M. and Burton, Z.F. (2005). "Dynamic error correction and regulation of downstream bubble opening by human RNA polymerase II." Mol.Cell. **18**(4): 461-470.

Guthold, M. and Erie, D.A. (2001). "Single-molecule study reveals a complex *E. coli* RNA polymerase." Chembiochem. **2**(3): 167-70.

Kettenberger, H., Armache, K.J. and Cramer, P. (2004). "Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS." Mol.Cell. **16**(6): 955-965.

Komissarova, N. and Kashlev, M. (1997). "RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA." J.Biol.Chem. **272**(24): 15329-15338.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A. and Darst, S.A. (2000). "A structural model of transcription elongation." Science **289**(5479): 619-625.

Kubori, T. and Shimamoto, N. (1996). "A branched pathway in the early stage of transcription by *Escherichia coli* RNA polymerase." J.Mol.Biol. **256**(3): 449-57.

Landick, R. (1999). "Shifting RNA polymerase into overdrive." Science **584**(5414): 598-599.

Lewin, B. (2000). Genes VII. New York, Oxford University Press Inc.

Matsuzaki, H., Kassavetis, G.A. and Geiduschek, E.P. (1994). "Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III." J.Mol.Biol. **235**(4): 1173-1192.

Mooney, R.A., Artsimovitch, I. and Landick, R. (1998). "Information processing by RNA polymerase: Recognition of regulatory signals during RNA chain elongation." J.Bacteriol. **180**(13):3265-75.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H., Yan, H. and Burton, Z.F. (2003). "NTP-driven translocation by human RNA polymerase II." J.Biol.Chem. **278**(20): 18303-18312.

Nudler, E., Mustaev, A., Lukhtanov, E. and Goldfarb, A. (1997). "The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase." Cell **89**(1): 33-41.

Reeder, T.C. and Hawley, D.K. (1996). "Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism." Cell **87**(4): 767-777.

Suh, W.C., Leirmo, S. and Record, M.T. Jr. (1992). "Roles of Mg2+ in the mechanism of formation and dissociation of open complexes between *Escherichia coli* RNA polymerase and the λPR promoter: Kinetic evidence for a second open complex requiring Mg2+." Biochemistry **31**: 7815-7825.

Sweetser, D., Nonet, M. and Young, R.A. (1987). "Prokaryotic and eukaryotic RNA polymerases have homologous core subunits." Proc.Natl.Acad.Sci.U.S.A. **84**(5): 1192-6.

Temiakov, D., Zenkin, N., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., Kashkina, E., Savkina, M., Zorov, S., Nikiforov, V., Igarashi, N., Matsugaki, N., Wakatsuki, S., Severinov, K. and Vassylyev, D.G. (2005). "Structural basis of transcription inhibition by antibiotic streptolydigin." Mol.Cell. **19**(5): 655-666.

Tolić-Nørrelykke, S.F., Engh, A.M., Landick, R. and Gelles, J. (2004). "Diversity in the rates of transcript elongation by single RNA polymerase molecules." J.Biol.Chem. **279**(5): 3292-3299.

Toulme, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M. and Rahmouni, A.R. (2000). "GreA and GreB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming." EMBO.J. **19**(24): 6853-9.

Tuske, S., Sarafianos, S.G., Wang, X., Hudson, B., Sineva, E., Mukhopadhyay, J., Birktoft, J.J., Leroy, O., Ismail, S., Clark, A.D. Jr., Dharia, C., Napoli, A., Laptenko, O., Lee, J., Borukhov, S., Ebright, R.H. and Arnold, E. (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., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I. and Landick, R. (2007). "Structural basis for substrate loading in bacterial RNA polymerase." Nature **448**(7150): 163-168.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H. and Artsimovitch, I. (2007). "Structural basis for transcription elongation by bacterial RNA polymerase." Nature **448**(7150): 157.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S. and Yokoyama, S. (2002). "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6Å resolution." Nature **417**(6890): 712-9.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D. and Kornberg, R.D. (2006). "Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis." Cell **127**(5): 941-954.

Westover, K.D., Bushnell, D.A. and Kornberg, R.D. (2004). "Structural basis of transcription: separation of RNA from DNA by RNA polymerase II." Science **303**(5660): 1014-1016.

Westover, K.D., Bushnell, D.A. and Kornberg, R.D. (2004). "Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center." Cell **119**(4): 481-489.

Yin, H., Artsimovitch, I., Landick, R. and Gelles, J. (1999). "Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules." Proc.Natl.Acad.Sci.U.S.A **96**(23): 13124-13129.

Zhang, C., Zobeck, K.L. and Burton, Z.F. (2005). "Human RNA polymerase II elongation in slow motion: role of the TFIIF RAP74 alpha1 helix in nucleoside triphosphate-driven translocation." Mol.Cell.Biol. **25**(9): 3583-3595.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K. and Darst, S.A. (1999). "Crystal structure of Thermus aquaticus core RNA polymerase at 3.3Å resolution." Cell **98**(6): 811-824.

CHAPTER 2: ACTIVE DISPLACEMENT OF NTPS DURING TRANSCRIPTION ELONGATION: A STRUCTURAL MODEL FOR MISINCORPORATION AND RESCUE

Introduction

Transcription, the processive DNA-directed synthesis of RNA, is catalyzed by RNA polymerase (RNAP) and is the first step in a chain of events that leads to gene expression in the cell. RNAP must catalyze the incorporation of an NMP into the growing RNA chain at reasonable rates and with high fidelity (low occurrence of incorrect NTP addition) to maintain the requirements of the cell. RNAP has been shown to exist in different states or conformations. These varying conformations are thought to play a role in the regulation of transcription, and hence the regulation of gene expression. Specifically, in both prokaryotes and eukaryotes, there is evidence that the enzyme can exist in an activated state of transcription, characterized by a faster rate of synthesis with low fidelity and an unactivated state that synthesizes RNA at a slower rate with higher fidelity of incorporation (Erie *et al.* 1992; Erie *et al.* 1993; Matsuzaki *et al.* 1994; Kubori & Shimamoto 1996; Coulombe & Burton 1999; Yin *et al.* 1999; Davenport *et al.* 2000; Foster *et al.* 2001; Guthold & Erie 2001; Erie 2002; Tolic-Norrelykke *et al.* 2004). In addition to these states, regulation of transcription has also been shown to occur through regulatory proteins (e.g. GreA, GreB, Nus factors, etc. in *E.coli*) or through interactions of the RNAP and specific sequences in the DNA or RNA. These regulatory factors all dictate the rate of incorporation as well as the

fidelity of the transcription reaction (Mooney *et al*. 1998).

The reported error rate for transcription in both prokaryotes and eukaryotes is 1 to 10 errors per 10^5 synthesized nucleotides (Libby & Gallant, 1991; Shaw *et al.* 2002). This rate is significantly higher than the genomic mutation rate of 1 in 10^9 errors reported for DNA polymerases (Echols & Goodman 1991; Kunkel & Bebenek 2000). Increased error during transcription can lead to deterioration of translation products, which can ultimately lead to functional instability and cell death (Taddei *et al.* 1997). As such, it is important to understand the machinery used in transcription to better understand the fidelity of RNAP. Despite its importance to the survival of the cell, the fidelity (ratio of correct incorporation to incorrect incorporation) of RNAP has not been well characterized (Alic *et al*. 2007). Here, we focus on the NTP concentration-dependent kinetics of misincorporation in an effort to better understand the fidelity of *E. coli* RNAP.

Addition of an incorrect nucleotide into the growing RNA chain during transcription is known as misincorporation. Erie *et al.* (1993) used *in vitro* misincorporation experiments to determine the initial branched mechanism of transcription elongation (Figure 2.1). This mechanism suggests that synthesis involves an activated (n*) and unactivated (n) enzyme complex. The transition from the unactivated to the activated state is characterized by conformational changes in the RNAP. The activated state (n*) is long-lived with synthesis occurring rapidly following NTP binding. Complexes can also decay off of the activated pathway and be trapped in the non-productive unactivated state (n). A fraction of these non-productive complexes can undergo further conformational changes into a dead-end state (n_{DE}) in which the complexes can not be elongated even in the presence of high

concentrations of NTPs (Erie *et al.* 1993). Similar results have been seen with eukaryotic RNA polymerase II during misincorporation (Thomas *et al.* 1998).

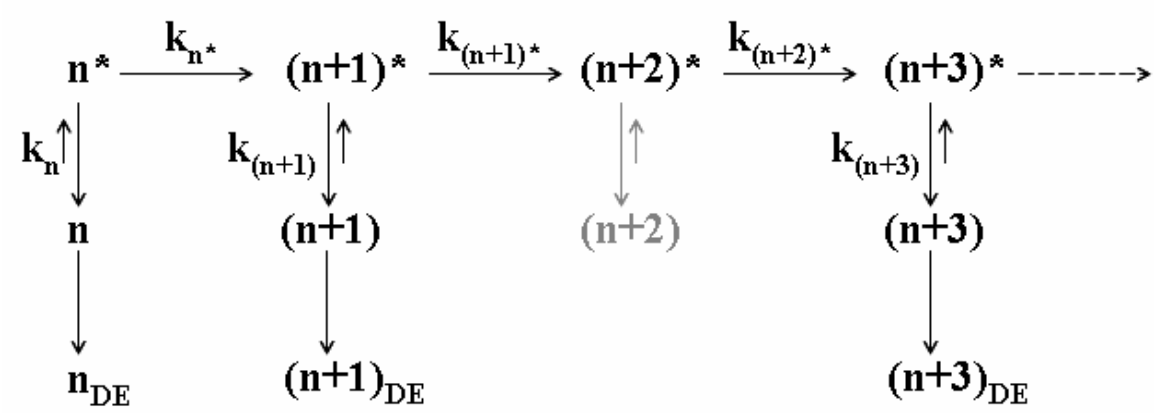


Figure 2.1: ***Original kinetic mechanism deduced from misincorporation experiments carried out by Erie et al.*** The n indicates the transcript position with the unactivated and activated state designated by n and n* respectively. The subscript "DE" represents complexes in the dead-end state of synthesis. The mechanism demonstrates the conformations of RNAP including a non-productive state of synthesis (n) to explain the observation of an incomplete reaction during misincorporation (Erie *et al.* 1993).

Further characterization of the kinetic mechanism using correct incorporation studies suggests that the transition from the unactivated to the activated state is achieved by nucleotide binding to a separate allosteric site on the RNA polymerase. Foster *et al.* (2001) performed experiments where a downstream templated, non-incorporatable NTP analog was added to the transcription reaction in the presence of the correct NTP for the n+1 (+26, pDE13) and n+ 2 (+27, pDE13) template positions. An increase in the rate of synthesis at the downstream position was observed in the presence of the non-incorporatable analog. This result suggested that there is an allosteric binding site on the enzyme. In the absence of an allosteric site, an inhibition of downstream NTP addition would have been observed in the presence of the analog (Foster *et al.* 2001). Synthesizing together kinetic data obtained for correct single nucleotide addition and the presence of an allosteric site, two kinetically indistinguishable non-essential activation mechanisms for correct incorporation are proposed (Holmes & Erie 2003).

Mechanism one (Figure 2.2A) assumes that the pre- and post-translocated states of the enzyme are in rapid equilibrium. In this model, an NTP binding first to the allosteric site causes a conformational change such that the enzyme enters the activated state of transcription [(n-1)*:NTP_A]. After activation, the NTP to be incorporated enters the catalytic site [(n-1)*:NTP_A:NTP] and pyrophosphate is released (n*:NTP_A:PPi) as synthesis occurs. However, if an NTP binds first to the catalytic site [(n-1):NTP], the enzyme remains in the unactivated state where catalysis can also occur (n:PPi) but at a slower rate (Holmes & Erie 2003).

The second proposed mechanism (Figure 2.2B) is similar to the first proposed mechanism with the exception of one key difference: the pre- and post-translocated states are

no longer assumed to be in rapid equilibrium. The system exists with the equilibrium favoring the pre-translocated state [(n-1)] until binding of an NTP to the allosteric site facilitates translocation [(n-1):NTP_A]. The NTP in the allosteric site can transfer into the catalytic site [(n-1):NTP_C], in which case the enzyme enters the same unactivated state of synthesis shown in Figure 2.2A. Alternatively, with an NTP bound to the putative allosteric site, a second NTP can bind to the catalytic site [(n-1):NTP_A:NTP_C] allowing for rapid synthesis along the activated pathway (Holmes & Erie 2003).

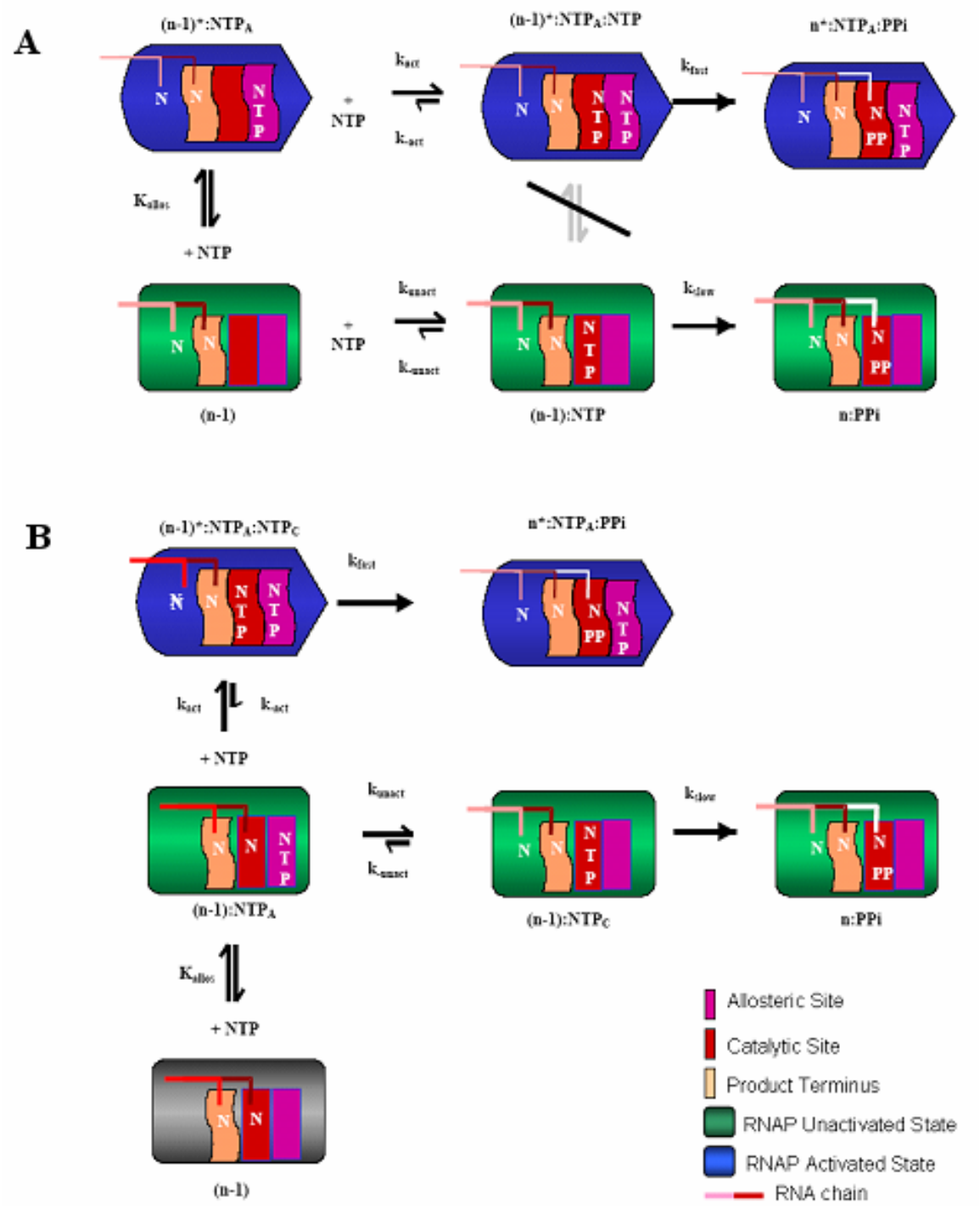

A
(n-1)*:NTP_A
(n-1)*:NTP_A:NTP
n*:NTP_A:PPi
+ NTP
K_allos
+ NTP
(n-1)
+ NTP
(n-1):NTP
n:PPi
B
(n-1)*:NTP_A:NTP_C
n*:NTP_A:PPi
+ NTP
(n-1):NTP_A
(n-1):NTP_C
n:PPi
+ NTP
(n-1)
Allosteric Site
Catalytic Site
Product Terminus
RNAP Unactivated State
RNAP Activated State
RNA chain

Figure 2.2: ***The proposed mechanisms of nucleotide addition during transcription elongation.*** (A) Mechanism one assuming that the pre- and post-translocated states of the enzyme are in rapid equilibrium. (B) A second kinetically indistinguishable mechanism in which the system exists pre-dominantly in the pre-translocated state of synthesis. The grey box represents the pre-translocated state (B). The green and blue boxes are the RNAP in the unactivated and activated states respectively. The magenta box represents the allosteric site, while the red box indicates the catalytic site. The product-terminus binding site is shown in the peach box. The red and pink lines indicate the growing RNA chain. NTP_C represents substrate bound to the catalytic site while NTP_A represents substrate bound to the allosteric site (Holmes & Erie 2003).

After examining the crystal structures of RNAPs from *T. thermophilus*, *T. aquaticus* and *S. cerevisiae*, Holmes and Erie (2003) proposed a location for the allosteric site (Figure 2.3). This site has many of the characteristics of a NTP binding site. Specifically, the region contains the flexible fork loop 2 (βD-loop I) that is surrounded by a β-sheet on one side and α-helices on the other. In both prokaryotes and eukaryotes, this loop is glycine-rich, a characteristic of "P-loops" which are responsible for binding NTPs (Walker *et al.* 1982; Kull *et al.* 1998; Via *et al.* 2000; Leipe *et al.* 2002). Furthermore, a totally conserved Walker B motif, an amino acid sequence that indirectly interacts with the γ-phosphate of NTPs through chelation of a Mg^{2+} ion, is located at the rear of the loop (Walker *et al.* 1982; Via *et al.* 2000). Taken together with the current non-essential activation mechanism (Figure 2.2), Holmes and Erie (2003) proposed a ratchet model for translocation (Figure 2.3). The model suggests that an NTP binds to fork loop 2, allosterically changing the conformation of the loop. This change in conformation of fork loop 2 begins a concerted movement which shifts the area of the protein directly contacting the DNA-RNA hybrid, moving the DNA-RNA hybrid via a ratchet motion, and thereby facilitating translocation (Holmes & Erie 2003).

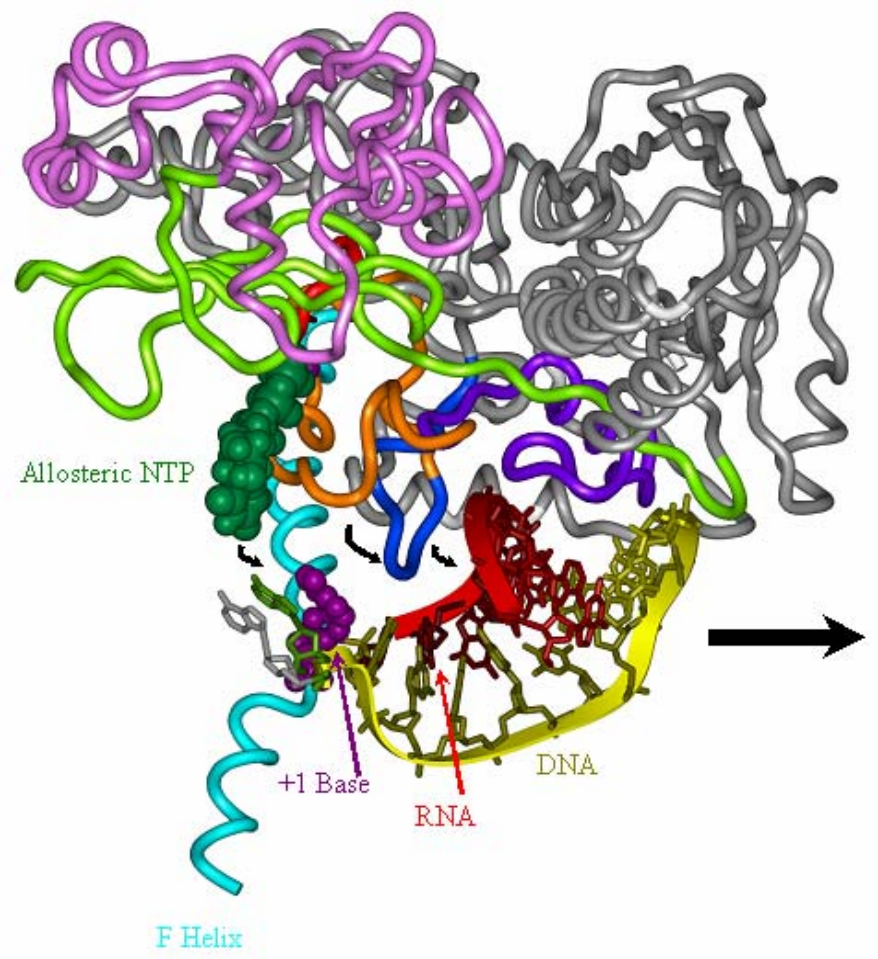


Figure 2.3: ***Model of the proposed RNAP allosteric site and the ratchet motion facilitating translocation (Holmes & Erie 2003).*** DNA template strand is shown in yellow with the nascent RNA chain shown in red. The F-helix is cyan, while fork loop 2 (βD loop I, the proposed allosteric binding site) is shown in orange. The flanking β-sheet and α-helices are shown in light green and pink, respectively. The rifampicin binding regions directly interacting with the DNA-RNA hybrid are purple and blue. A modeled in UTP molecule is shown bound to the allosteric site (green, space filled) and at 5-6Å distance from the DNA, the allosteric NTP can interact with the downstream DNA base (purple, space filled) (Holmes & Erie 2003).

To further support the presence of an allosteric site at fork loop 2, Kennedy studied the effect of pre-incubation and simultaneous addition of the downstream NTP (n+2) on correct incorporation for wild type RNAP as well as a Δ-loop RNAP where residues R542-F545 in the fork loop 2 region of the enzyme were removed. In wild type RNAP, simultaneous addition of n+2 (ATP) with n+1 (CTP) had little effect on the incorporation of n+1; however, pre-incubation of n+2 followed by addition of n+1 increased the rate in which n+1 incorporated into the RNA. Significantly, in both simultaneous and pre-incubation experiments there was a dramatic increase in the rate of n+2 being incorporated when n+2 was present in the wild type enzyme experiments. In fact, the rate of n+2 incorporation was limited only by the rate at which n+1 incorporated even at 10 fold higher concentrations than n+2. The enhanced rate of incorporation at n+2 suggests that there is, in fact, a second NTP binding site in RNAP that is acting allosterically (Kennedy & Erie, manuscript in preparation). The data for wild type RNAP exhibited biphasic kinetics, consisting of a slow phase and a fast phase of synthesis as seen previously (Foster *et al.* 2001; Holmes & Erie 2003).

Pre-incubation and simultaneous addition of n+2 with n+1 in the Δ-loop RNAP had little effect on the rate of n+1 incorporation compared to wild type. However, the rate of n+2 in the presence of n+2 for both pre-incubation and simultaneous addition was dramatically decreased in the Δ-loop mutant. This result is in stark contrast to the results seen for wild type RNAP where the rate of n+2 was enhanced by the presence of n+2, suggesting that fork loop 2 is in fact acting as the second NTP binding site in RNAP. Also in contrast to wild type RNAP, during simultaneous addition of n+2 with n+1 in the Δ-loop RNAP only a single slow phase of synthesis was observed. This phase was relatively

unaffected compared to wild type enzyme but the previously observed fast phase of synthesis was completely eliminated. This result suggests that there is a fast phase of synthesis that is utilizing the fork loop 2 (a.k.a., the allosteric site) while there is also a slow phase of synthesis that is independent of fork loop 2 (Kennedy & Erie, manuscript in preparation).

These data, taken together with recent crystal structures where another important structural element known as the trigger loop was shown to interact with fork loop 2 and an NTP in the catalytic site, led to the model of nucleotide incorporation shown in Figure 2.4 (Toulokhonov *et al.* 2007; Vassylyev *et al.* 2007; Kennedy & Erie, manuscript in preparation). This model expands upon the model previously proposed by Holmes and Erie (2003) shown in Figure 2.2B. RNAP is shown first in the pre-translocated state (n). The previously added nucleotide is locked into the catalytic site by the trigger loop, shown to close down over the catalytic site for synthesis during transcription and preventing NTP entry into the catalytic site via the secondary channel (Toulokhonov *et al.* 2007; Vassylyev *et al.* 2007). An allosteric NTP binds to fork loop 2 via the main channel (1), and this binding of an NTP to the allosteric site weakens the trigger loop's affinity for the catalytic site. The trigger loop adapts an open conformation which opens access to the catalytic site (2). In the open conformation, the trigger loop interacts with the allosteric NTP and the NTP acts as a latch to hold the trigger loop in the open conformation. Following translocation, a second NTP can then enter into the catalytic site through the secondary channel (3a). When the NTP binds to the catalytic site, the trigger loop loses its affinity for the allosteric NTP and the trigger loop is again able to close down on the catalytic NTP for synthesis (4). A second possibility for NTP entry into the catalytic site is that the trigger loop interacting with the allosteric NTP carries the allosteric NTP over or under the bridge helix via a hand off

mechanism to transfer the allosteric NTP into the catalytic site. The allosteric NTP becomes the catalytic NTP and the trigger loop is able to close down over the catalytic site for chemistry (3b, 4) (Kennedy & Erie, manuscript in preparation).

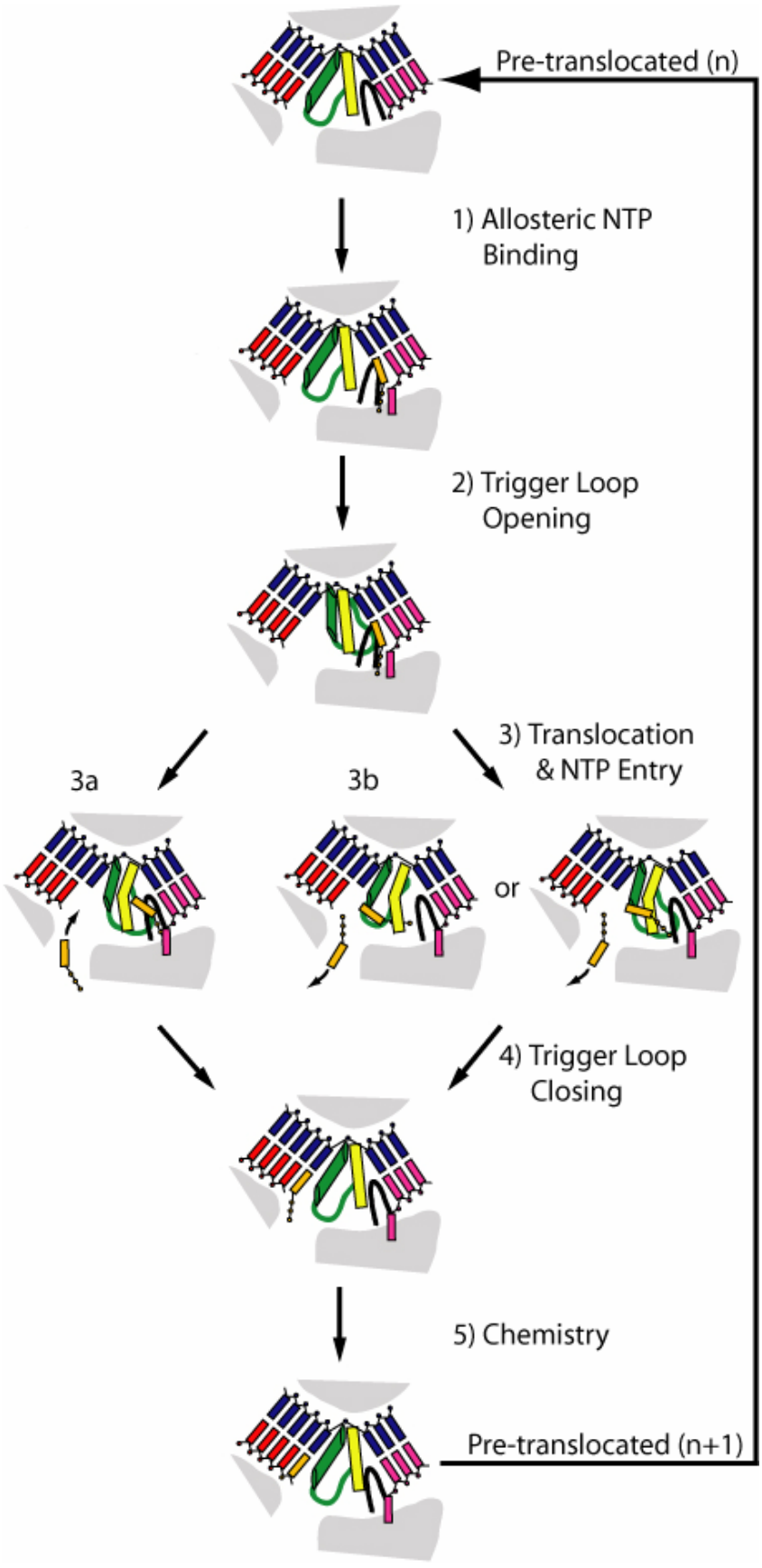

Pre-translocated (n)
1) Allosteric NTP Binding
2) Trigger Loop Opening
3) Translocation & NTP Entry
3a
3b
or
4) Trigger Loop Closing
5) Chemistry
Pre-translocated (n+1)

Figure 2.4: ***Model for NTP addition.*** The DNA template strand is shown in blue with the non-template strand shown in pink. The RNA chain is red. The allosteric NTP is orange. Fork loop 2 is black, the bridge helix (F-helix) is yellow, and the trigger loop is shown in green (Kennedy & Erie, manuscript in preparation). This model structurally expands upon the model previously proposed by Holmes & Erie (2003).

In this study, we focus on determining the rate of misincorporation and compare these data to the available data on correct incorporation to gain a greater understanding of the fidelity of *E. coli* RNAP. We have studied the NTP concentration-dependent kinetics of misincorporation for wild type *E. coli* RNAP and determined that misincorporation can be described by a non-essential activation mechanism where synthesis can only occur in the activated state while a subset of complexes are "trapped" in the unactivated state. Furthermore, NTP concentration-dependent kinetic studies of misincorporation were performed utilizing the Δ-loop RNAP. The results indicate that this mutant is a high fidelity mutant with a decreased rate and extent of misincorporation. We also reveal an active displacement of the incorrect NTP in the presence of the correct nucleotide in both wild type and Δ-loop RNAP and propose a structural model for misincorporation similar to the model proposed by Kennedy & Erie (manuscript in preparation).

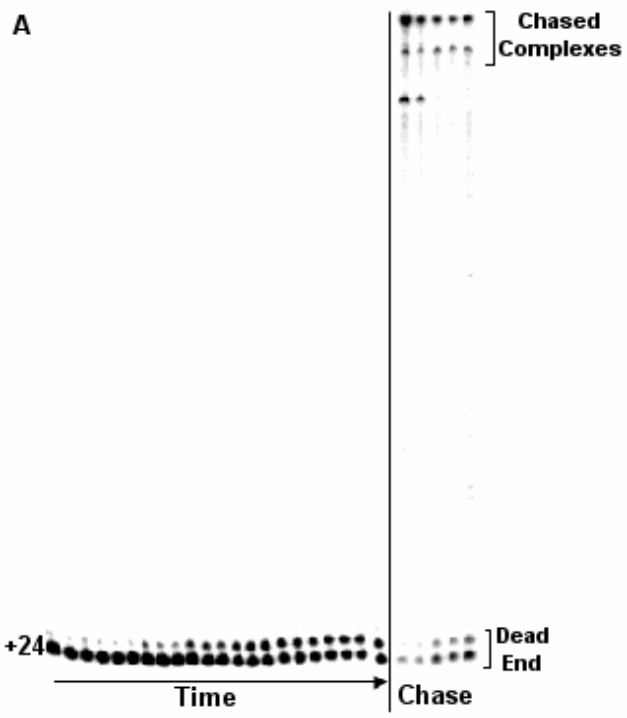
Results

All experiments were performed using the pDE13 DNA template where the first CMP to be incorporated in the RNA chain is at position +25 (Erie *et al.* 1993). This template is biotinylated at the 5' end and attached to streptavidin-coated magnetic beads. Elongation complexes were formed by adding *E. coli* RNAP, DNA, UTP, ATP, and [α-^{32}P] GTP and stalled at position +24 by omitting CTP. The complexes were then placed next to a magnet and purified by washing with buffer (See Methods.) We then monitored the misincorporation of UMP for CMP at position +25 as a function of time. Because misincorporation happens at a rate much slower than that of correct incorporation, reactions were carried out by hand as opposed to using rapid quench techniques (Erie *et al*. 1993).

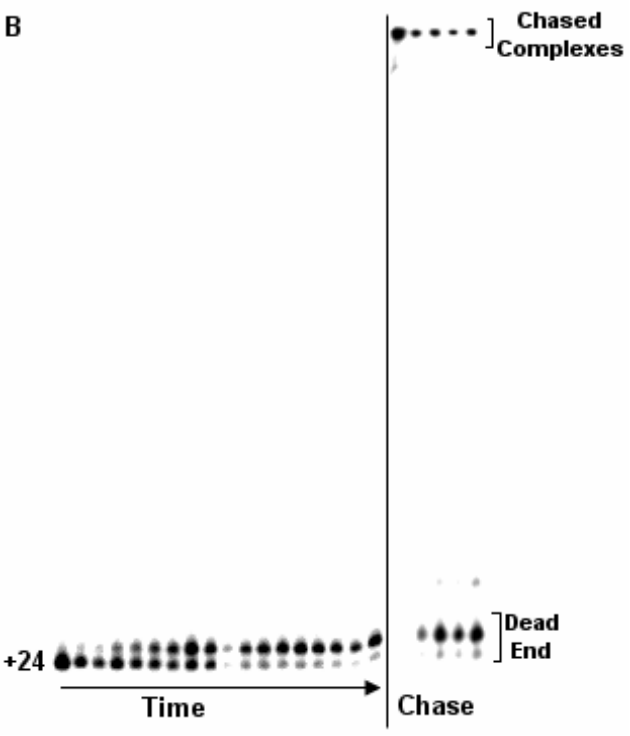
Concentration-Dependence of UMP Incorporation Using Wild Type RNAP

Figure 2.5A and 2.5B shows gels of misincorporation of UMP for CMP at position +25 in the nascent RNA chain as a function of time represented at a low (5μM) and high concentration (600μM) of UTP, respectively. Inspection of these images reveals an increase in the rate (appearance of complexes at position +25) and an increase in the extent (total percent of complexes at position +25) of misincorporation with increasing concentration of UTP. Specifically, there is a burst of misincorporation and then no further misincorporation is observed. Notably, at lower UTP concentrations, less than 100% of the complexes misincorporate UTP in place of CTP even after 40 minutes (Figure 2.5A). A subset of these complexes is still competent for correct incorporation (chased complexes – complexes added to the presence of all four NTPs) and the remaining complexes have entered an inactive dead-end state (Figure 2.5A, chases).

Unlike DNAP, which will misincorporate most bases to 100% given sufficient time (Wong *et al.* 1991), only a subset of the RNA complexes misincorporate UMP at lower concentrations. The fact that we do not see 100% misincorporation with RNAP at all concentrations of UTP indicates that a subset of complexes are decaying from productive synthesis into a non-productive synthesis path in an NTP concentration dependent manner. This observation is consistent with the original misincorporation experiments where misincorporation was thought to occur only along an activated path with complexes falling off pathway into a non-productive unactivated state (Erie *et al.* 1993).



pppAUGUAGUAAGGAGGUUGUAUGGAA<u>C</u>AACGCAUAACCCUGA...
+25



pppAUGUAGUAAGGAGGUUGUAUGGAA<u>C</u>AACGCAUAACCCUGA...
+25

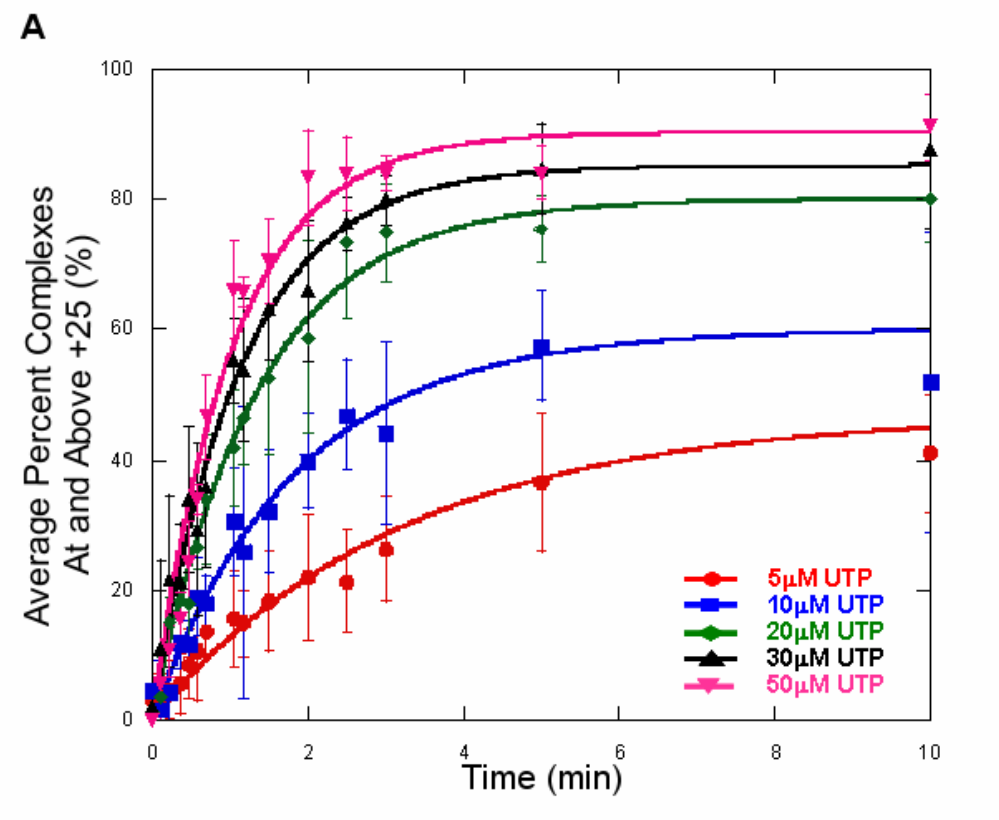
Figure 2.5: ***Representative denaturing gels showing UMP misincorporation at position +25 at (A) 5μM UTP and (B) 600μM UTP added to purified complexes stalled at position +24 in the nascent RNA chain.*** The rate of misincorporation at position +25 increases with an increase in UTP concentration. Also, the percent of complexes misincorporated at position +25 increases with increasing UTP concentration. Time = 0 (prior to NTP addition), 0.12, 0.24, 0.35, 0.47, 0.59, 0.7, 0.82, 0.94, 1.05, 1.17, 1.5, 2, 2.5, 3, 5, 10, 20, and 40 minutes. Chase reactions = 0, 4, 14, 24, and 39 minutes.

To quantitatively analyze these data, the percent of total complexes 25 nucleotides and longer were quantified and plotted as a function of time. Figure 2.6 shows kinetic data for misincorporation of UMP for CMP at eight different UTP concentrations. Inspection of the data in Figure 2.6 reveals quantitatively that both the rates and extents of misincorporation increase with increasing UTP concentration. To determine the pseudo-first-order-rate constant (k_{app}) and the maximum extents of misincorporation (plateau values) the data for all UTP concentrations were fit to single exponentials [$y=Aexp(-k_{app}t)+C$] (Figure 2.6).

To examine the concentration dependence of the rate of UMP misincorporation, the apparent rate constants obtained from the single exponential fits were plotted versus UTP concentration (Figure 2.7). Given our current model for NTP addition, we would expect to see a sigmoidal plot of k_{app} versus [UTP]. A sigmoidal substrate saturation curve would indicate a quadratic dependence of the rate of UMP incorporation on UTP concentration and suggest that there are two NTP binding sites serving the purpose of misincorporation in RNAP (Segel 1975; Schulz 1994). However, the rate of misincorporation versus UTP concentration increases approximately linearly with UTP concentration (Figure 2.7). This linear result is similar to that seen with DNAPs and indicates NTP binding sites have not reached saturation at 600μM UTP (Wong *et al.* 1991). This lack of saturation has also been observed for RNA polymerase III for concentrations up to 600μM (Alic *et al.* 2007).

To further examine the misincorporation reaction, the maximum extents of misincorporation (total percent complexes at position +25) were plotted versus UTP concentration (Figure 2.8). The plot of percent extent versus [UTP] shows the increase in extent of misincorporation with increasing UTP concentration. Interestingly, the plot of

extent of misincorporation versus UTP concentration fits well to Michaelis-Menten kinetics and a K_m of 6μM was obtained. Typically for Michaelis-Menten kinetics, the plot of initial velocity (ν) versus substrate concentration ([S]) is used to determine the kinetic variables where K_m represents the concentration of substrate at which the enzyme reaches half-maximum velocity ($½V_{max}$) (Segel 1975; Schulz 1994). However, Figure 2.8 demonstrates an extent of misincorporation that is dependent on substrate concentration and binding, and therefore K_m is similar to the constant of half saturation of an NTP binding site on the RNAP.


A
Average Percent Complexes
At and Above +25 (%)
Time (min)
5μM UTP
10μM UTP
20μM UTP
30μM UTP
50μM UTP

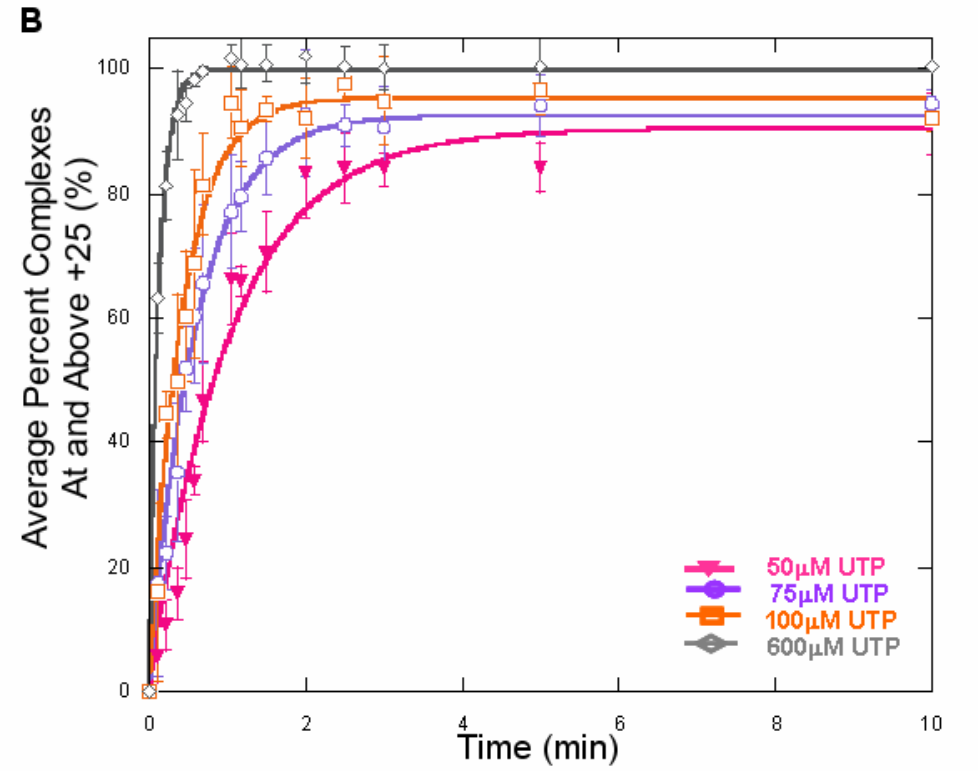

B
Average Percent Complexes
At and Above +25 (%)
Time (min)
50μM UTP
75μM UTP
100μM UTP
600μM UTP

Figure 2.6: ***Plots of percent misincorporated complexes at position +25 versus time at (A) 5 – 50μM UTP and (B) 50-600μM UTP.*** These data are fit to single exponentials to obtain the apparent rate constant (k_{app}) and maximum extent of misincorporation (plateau value) for each concentration of UTP. Error bars represent the standard deviation for three to five sets of data for each concentration of UTP.

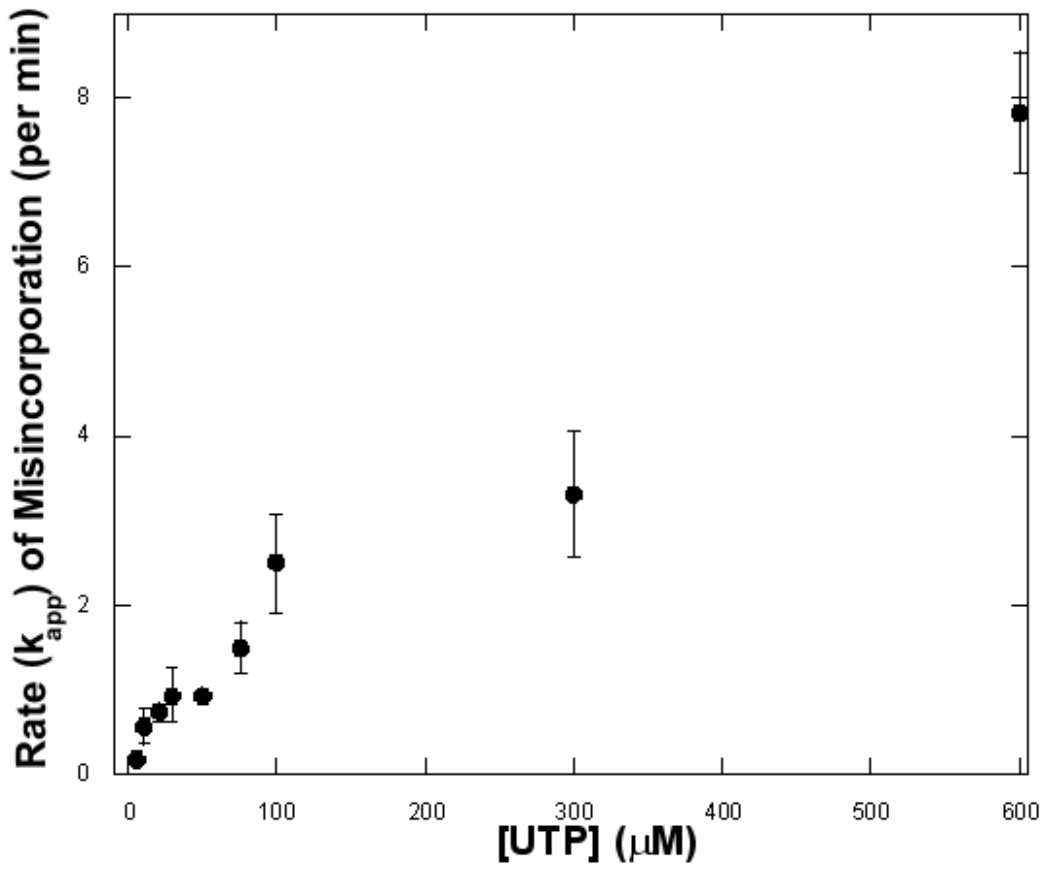


Figure 2.7: ***Plot of rate*** **(*k_{app}*, *min^{-1}*) *versus [UTP] (μM).*** The rate of misincorporation increases approximately linearly with increasing UTP concentration, indicating that RNAP has not reached substrate saturation at 600μM UTP. Error bars represent standard deviation for three to five sets of data for each concentration of UTP.

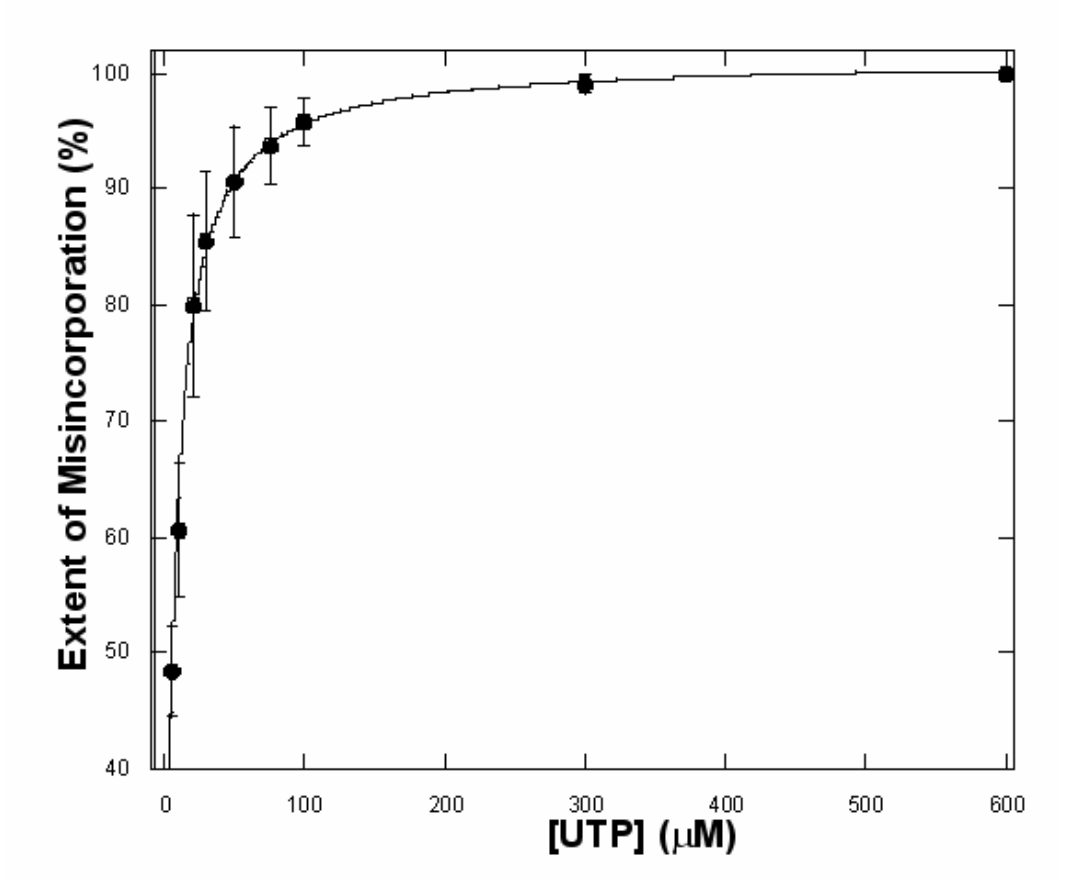


Figure 2.8: ***Plot of maximum extent of misincorporation (%) versus [UTP] (μM).*** Data were fit to Michaelis-Menten kinetics (extent misincorporation $= \frac{V_{\max} * [UTP]}{K_m + [UTP]}$) with a K_m value of 6μM.

Chase Reactions from Wild Type RNAP

At designated times during the *in vitro* misincorporation reactions, a sample of the purified elongation complexes was added to the presence of all four NTPs (1mM) to extend the transcript to full length and ensure that the complexes were still active (Chase Reaction, See Methods). As previously noted, misincorporation does not go to 100% completion at lower concentrations of UTP (<75μM) suggesting that a subset of the complexes are falling off pathway into a non-productive synthesis state. The chase reactions show that a significant portion of the complexes that do not misincorporate U for C at +25 are still competent for synthesis and are not dead-end (inactive) complexes (represented in Figure 2.5A and 2.5B, final five lanes). These data suggest that, in the presence of the correct NTP, the previously non-productively bound incorrect nucleotide could be displaced so that the complexes could incorporate the correct nucleotide and continue on to a complete (100%) reaction. The complexes that do not chase are dead-end complexes and remain inactive regardless of the concentration of all four NTPs present.

To further investigate the observation of non-productively bound nucleotide being displaced in the presence of all four NTPs, we posited that the presence of the correct NTP alone would be sufficient to displace the non-productively bound nucleotide and allow for continued synthesis of the RNA chain. To test this hypothesis, the misincorporation reaction with purified elongation complexes was performed with the addition of 20μM UTP. The misincorporation reaction was carried out for 10 minutes before CTP was added to the reaction (See Methods). Following misincorporation, reactions go to completion within the first 7 seconds of CTP addition (all concentrations: 5μM, 50μM, and 1mM) (Figure 2.9).

To determine the exact time scale of the addition of the correct NTP following misincorporation, similar chase experiments were performed using rapid quench kinetic techniques (See Methods). These rapid quench chase reactions show that CTP displaces the previously non-productively bound NTP, and incorporates CMP on the same time scale as correct single nucleotide addition with no prior misincorporation (0.1-0.2 seconds: Foster *et al.* 2001; Holmes & Erie 2003; Kennedy & Erie, manuscript in preparation). This rapid incorporation demonstrates that these non-productive complexes are not dead-end and that the complexes are not in a backtracked state. In a backtracked state, the RNAP has translocated backwards along the DNA while extruding the 3' end of the RNA such that regulatory proteins can act on the misincorporated base (Reeder & Hawley 1996, Komissarova & Kashlev 1997; Nudler *et al.* 1997; Zhang *et al.* 1999; Artsimovitch & Landick 2000; Toulme *et al.* 2000). It has been shown that recovery from the backtracked state is slow; therefore, if the complexes were in a backtracked state, the addition of the correct nucleotide would cause the complexes to progress slowly out of this state.

We also investigated the possibility that the incorrect base added after misincorporation could chase the complexes out of the non-productive state. The misincorporation reaction with purified elongation complexes was performed with the addition of 20μM UTP. The misincorporation reaction was carried out for 10 minutes before complexes were washed again to remove any unreacted UTP. We attempted to restart the misincorporation reaction by adding 20μM UTP back to these complexes (See Methods.) The addition of low concentrations of UTP after the misincorporation reaction yielded no change in the extent of complexes at position +25 over the course of 10 minutes and

demonstrates that the correct NTP or high concentration of the incorrect NTP is necessary to chase the complexes out of the non-productive state of synthesis.

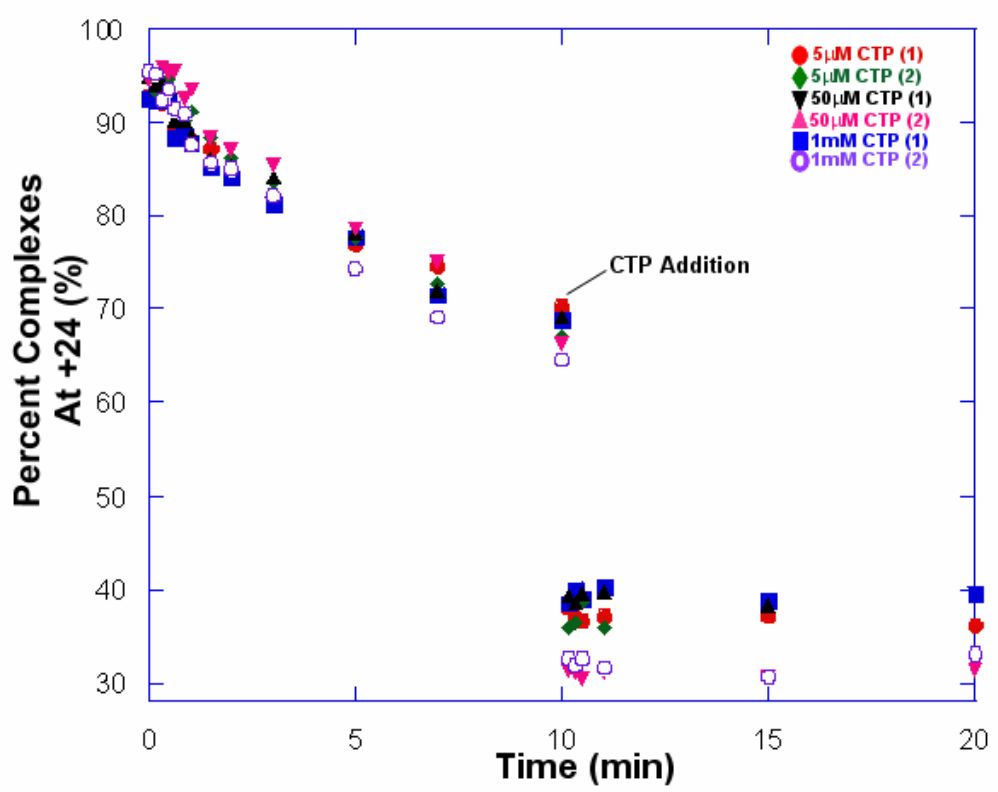


Figure 2.9: ***Plot of the disappearance of complexes out of position +24 following the addition of CTP after 10 minutes of misincorporation.*** The chase reaction of CTP into position +25 occurs within the first seven seconds of CTP addition following the misincorporation reaction.

Concentration-Dependence of UMP Incorporation using Δ-loop RNAP

Pre-incubation and simultaneous addition experiments performed by Kennedy with wild type RNAP and Δ-loop RNAP, where four residues (R542-F545) of fork loop 2 were deleted, support the existence of a second NTP binding site in RNAP and demonstrate that fork loop 2 is likely to be the area of the protein where the second NTP is binding to act allosterically during transcription elongation (manuscript in preparation). To better understand the role of the allosteric site in misincorporation, we continued our misincorporation studies using the Δ-loop RNAP.

Initially, we performed running start experiments with Δ-loop RNAP where misincorporation was initiated from the promoter with 15μM UTP in the presence of 20μM ATP and 20μM GTP (See Methods). Inspection of the gels in Figure 2.10 reveals that the Δ-loop mutant (Figure 2.10B) exhibits a decreased rate (appearance of complexes at position +25) and decreased extent (total percent of complexes at position +25) of misincorporation compared to that of wild type (Figure 2.10A). Kennedy and Erie demonstrated that correct incorporation of CTP at position +25 with Δ-loop mutant is only modestly affected by the deletion of the four residues in fork loop 2 (manuscript in preparation). The decreased rate and extent of misincorporation observed in the Δ-loop RNAP, therefore, suggest that the deletion mutant enzyme is a higher fidelity enzyme compared to wild type RNAP.

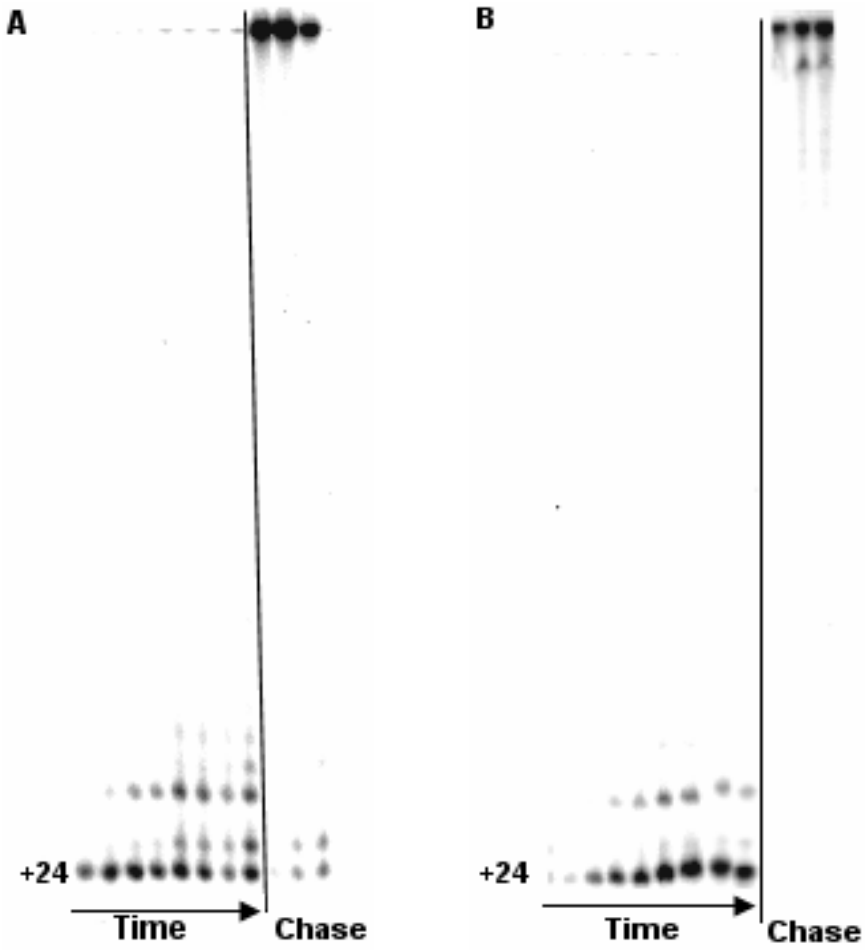


Figure 2.10: ***Representative denaturing gels showing UMP misincorporation in Δ-loop RNAP at position +25 initiated from the promoter with 15μM UTP.*** The running start reaction of (A) wild type RNAP and (B) Δ-loop RNAP are shown. The rate and extent of misincorporation is decreased in the Δ-loop RNAP compared to wild type enzyme. Time = 0.5, 1, 1.5, 2, 5, 7, 10, and 20 minutes. Chase reactions = 0, 5, and 20 minutes.

We further characterized the Δ-loop RNAP by performing the same concentration-dependent kinetic series described for wild type RNAP where specific concentrations of UTP were added to purified elongation complexes and the misincorporation reaction was monitored over time. Δ-loop RNAP exhibited the same decrease in rate and extent of misincorporation with purified elongation complexes as was seen in the running start reaction (Figure 2.10). Images of representative gels for misincorporation of UMP for CMP at position +25 in the nascent RNA chain as a function of time represented at a low (20μM) and high concentration (600μM) of UTP are shown in Figure 2.11A and 2.11B, respectively. As with wild type RNAP (Figure 2.5), we observe an increase in the rate and extent of misincorporation with increasing UTP concentration in the Δ-loop RNAP. Also, at lower UTP concentrations we still observe that less than 100% of the complexes misincorporate UTP in place of CTP even after 40 minutes (Figure 2.11A). A subset of these complexes is still competent for correct incorporation (chased complexes) and the remaining complexes have entered an inactive dead-end state (Figure 2.11A, chases). Interestingly, the Δ-loop RNAP enters into the dead-end state after the misincorporation event at +25 (Figure 2.11A and B) while in the wild type enzyme, the dead-end state occurs prior to misincorporation at +24.

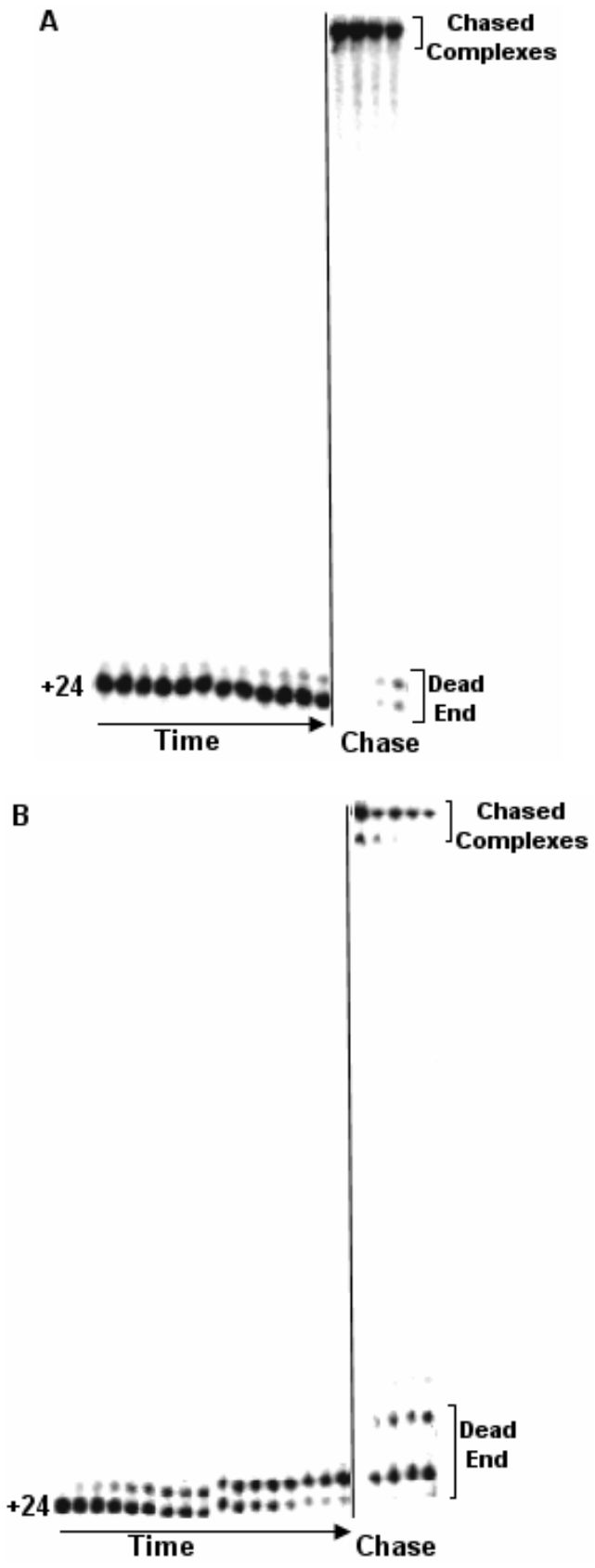

A
Chased
Complexes
+24
Dead
End
Time
Chase
B
Chased
Complexes
Dead
End
+24
Time
Chase

Figure 2.11: ***Representative denaturing gels showing UMP misincorporation with Δ-loop RNAP at position +25 at (A) 20μM UTP and (B) 600μM UTP added to purified complexes stalled at position +24 in the nascent RNA chain.*** The rate of misincorporation at position +25 increases with an increase in UTP concentration. Also, the percent of complexes misincorporated at position +25 increases with increasing UTP concentration. Time = 0 (prior to NTP addition), 0.12, 0.24, 0.35, 0.47, 0.59, 0.7, 0.82, 0.94, 1.05, 1.17, 1.5, 2, 2.5, 3, 5, 10, 20, and 40 minutes. Chase reactions = 0, 4, 14, 24, and 39 minutes.

The percent of total complexes 25 nucleotides and longer were quantified and plotted as a function of time. Figure 2.12 shows kinetic data for misincorporation by Δ-loop RNAP of UMP for CMP at 3 different UTP concentrations (representing a low, intermediate, and high concentration of UTP). Inspection of the data in Figure 2.12 reveals quantitatively that both the rates and extents of misincorporation increase with increasing UTP concentration in the Δ-loop RNAP. As with wild type enzyme, the Δ-loop RNAP data were fit to single exponentials to determine the pseudo-first-order-rate constants (k_{app}) and the maximum extents of misincorporation (plateau values) for all UTP concentrations (Figure 2.12).

To determine the concentration dependence of the rate of UMP misincorporation with the Δ-loop RNAP, the apparent rate constants obtained from the single exponential fits were plotted versus UTP concentration (Figure 2.13). These data were plotted with wild type data for comparison. The concentration-dependent rate for Δ-loop RNAP exhibits a 100-fold decrease compared to the rate of wild type RNAP. Previous work by Erie *et al.* (1993) suggested that misincorporation can occur only in an activated (fast) state of synthesis. Kennedy and Erie demonstrated that deletion of the four fork loop 2 residues eliminates the fast phase of synthesis suggesting that the fork loop 2 is responsible for the fast (activated) synthesis (manuscript in preparation). As such, the decreased rate of misincorporation is expected in the Δ-loop mutant and further supports that misincorporation can only occur in an activated state of synthesis.

To further examine the misincorporation reaction with Δ-loop RNAP, the maximum extents of misincorporation (total percent complexes at position +25) were plotted versus UTP concentration and compared to wild type enzyme (Figure 2.14). As in wild type RNAP, the plot of percent extent versus [UTP] for Δ-loop RNAP demonstrates an increase in extent

of misincorporation with increasing UTP concentration. The Michaelis-Menten kinetic fit to the maximum extent data shows a 10- fold decrease in the binding affinity for the NTP to the Δ-loop mutant. The experimental K_m for wild type is 6μM and for Δ-loop RNAP, K_m is 50μM. This decreased affinity of the NTP for the mutant RNAP suggests that the NTP is binding to fork loop 2. Therefore, the experimentally determined value for K_m may be considered a binding constant for the allosteric site of the RNAP (K_{allos}, Figure 2.2A and 2.2B).

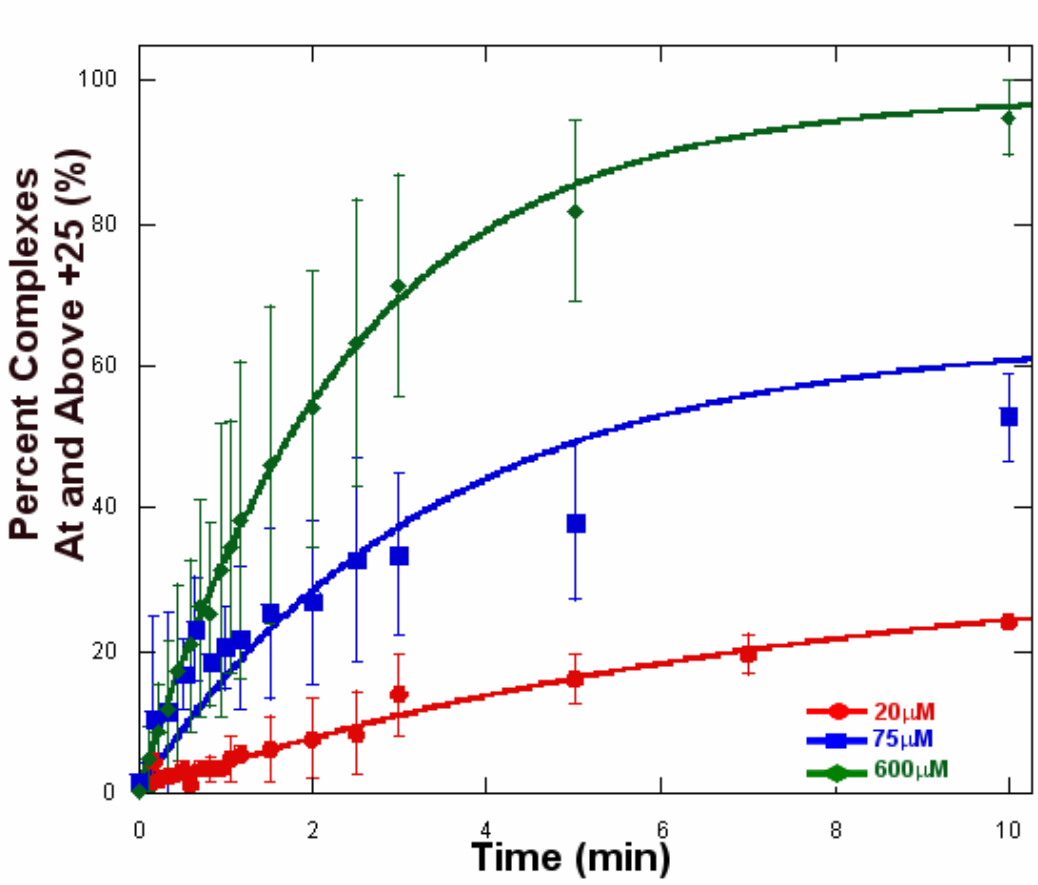


Figure 2.12: ***Plot of percent misincorporated complexes at position +25 versus time at a low (20μM), intermediate (75μM), and high (600μM) concentration of UTP utilizing the Δ-loop RNAP.*** These data are fit to single exponentials to obtain the apparent rate constant (k_{app}) and maximum extent of misincorporation (plateau value) for each concentration of UTP. Error bars represent the standard deviation for three to four sets of data for each concentration of UTP.

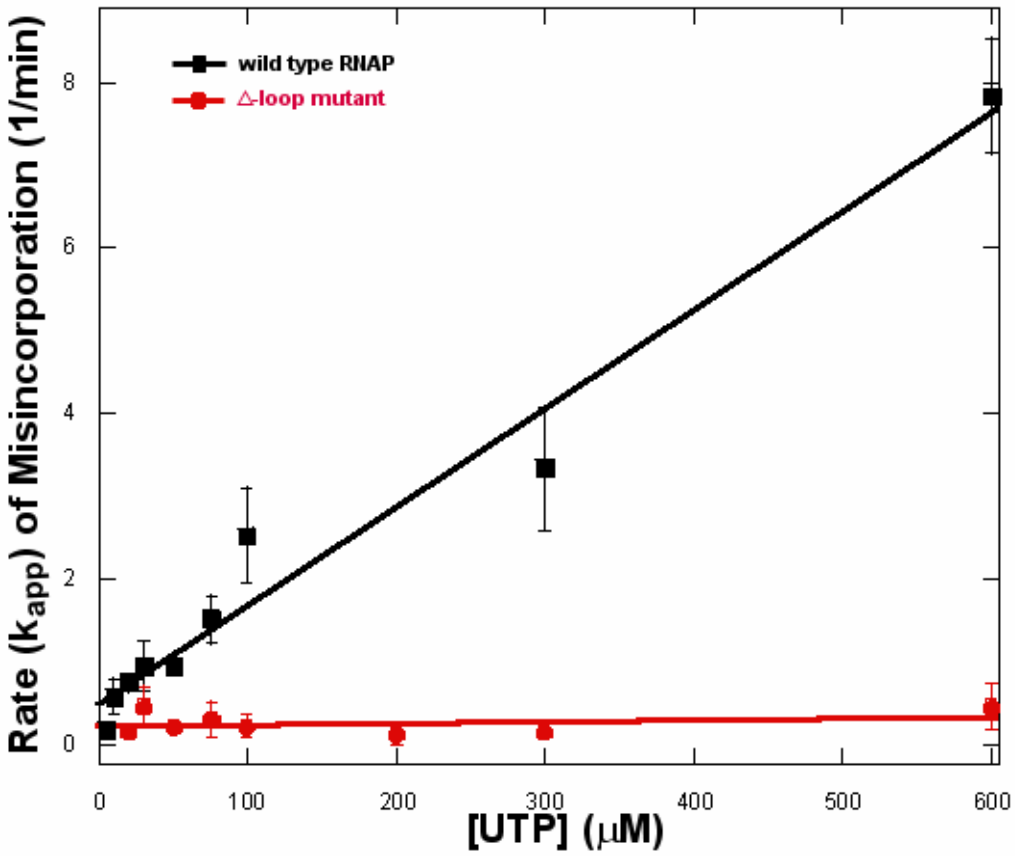


Figure 2.13: ***Rate (k_{app}, min^{-1}) versus [UTP] (μM) for wild type (black squares) and Δ-loop mutant RNAP (red circles).*** Linear fits [y = m*x + b] to the data reveal an approximate 100-fold decrease in the rate of misincorporation in the mutant RNAP. $\mathbf{Rate_{wild\ type}} = 0.0119 \mu M^{-1} min^{-1}$; $\mathbf{Rate_{\Delta\text{-}loop}} = 0.000115\ \mu M^{-1} min^{-1}$.

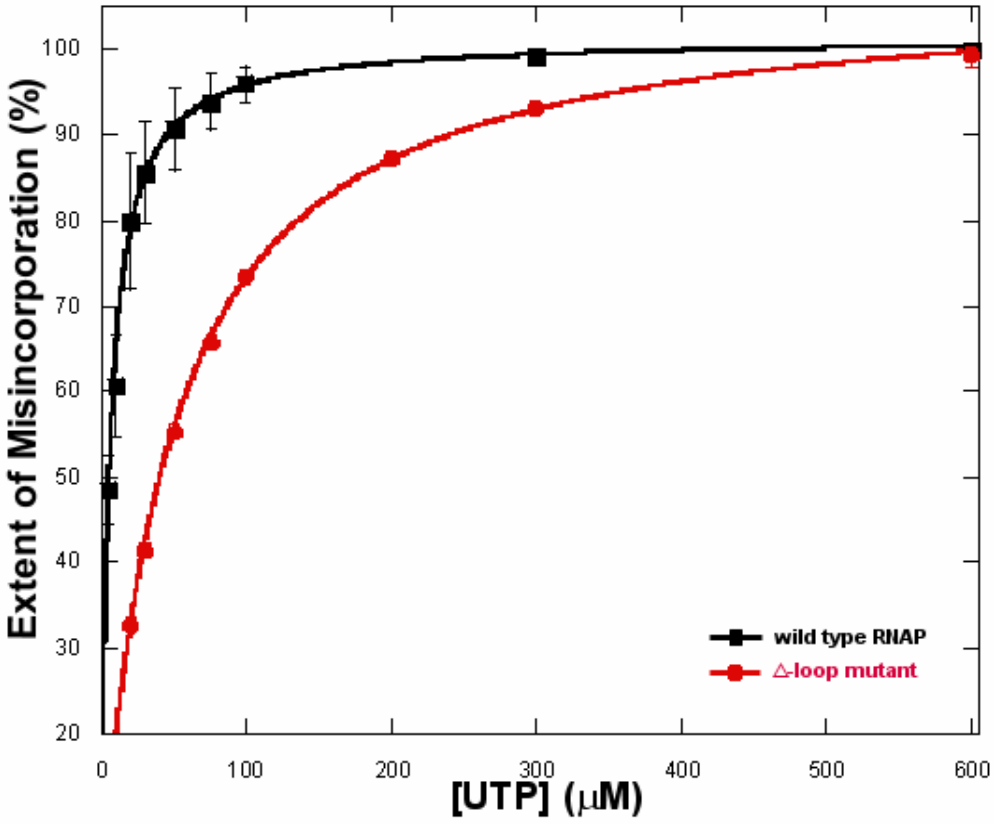


Figure 2.14: ***Plot of maximum extent of misincorporation (%) versus [UTP] (μM) for wild type (black squares) and Δ-loop mutant RNAP (red circles).*** Data were fit to Michaelis-Menten kinetics (extent misincorporation $= \frac{V_{max} * [UTP]}{K_m + [UTP]}$), revealing a 10-fold decrease in the affinity for NTP binding in the Δ-loop mutant (K_m = 50μM) compared to wild type enzyme (K_m = 6μM).

Chase Reactions from Δ-loop RNAP

At designated times during the *in vitro* misincorporation reactions with Δ-loop RNAP, a sample of the purified elongation complexes was added to the presence of all four NTPs (1mM) to extend the transcript to full length and ensure that the complexes were still active (Chase Reaction, See Methods). As previously noted for wild type and Δ-loop RNAP, misincorporation does not go to 100% completion at lower concentrations of UTP suggesting that a subset of the complexes are falling off pathway into a non-productive synthesis state. As in wild type, the chase reactions for Δ-loop RNAP show that a portion of the complexes not undergoing misincorporation are also not dead-end (inactive) complexes (represented in Figure 2.10, final three lanes; and, Figure 2.11, final five lanes). These data show that, in the presence of the correct NTP, the previously non-productively bound incorrect nucleotide could be displaced so that the complexes could incorporate the correct nucleotide and continue on to a complete (100%) reaction in the Δ-loop RNAP as well as in wild type enzyme.

We further investigated the chase reaction of the Δ-loop RNAP by performing the same CTP addition experiment described for wild type RNAP. The misincorporation reaction with purified elongation complexes from Δ-loop RNAP was performed with the addition of 20μM UTP. The misincorporation reaction was carried out for 10 minutes before CTP was added to the reaction (See Methods). Following misincorporation, reactions go to completion within the first 7 seconds of 100μM CTP addition (Figure 2.15). This result is consistent with the minimal effect of Δ-loop RNAP on correct incorporation of CTP and suggests that the mechanism by which the correct NTP "rescues" the enzyme from a non-

productive state of synthesis is unaffected by the deletion of the four fork loop 2 residues (Kennedy & Erie, manuscript in preparation).

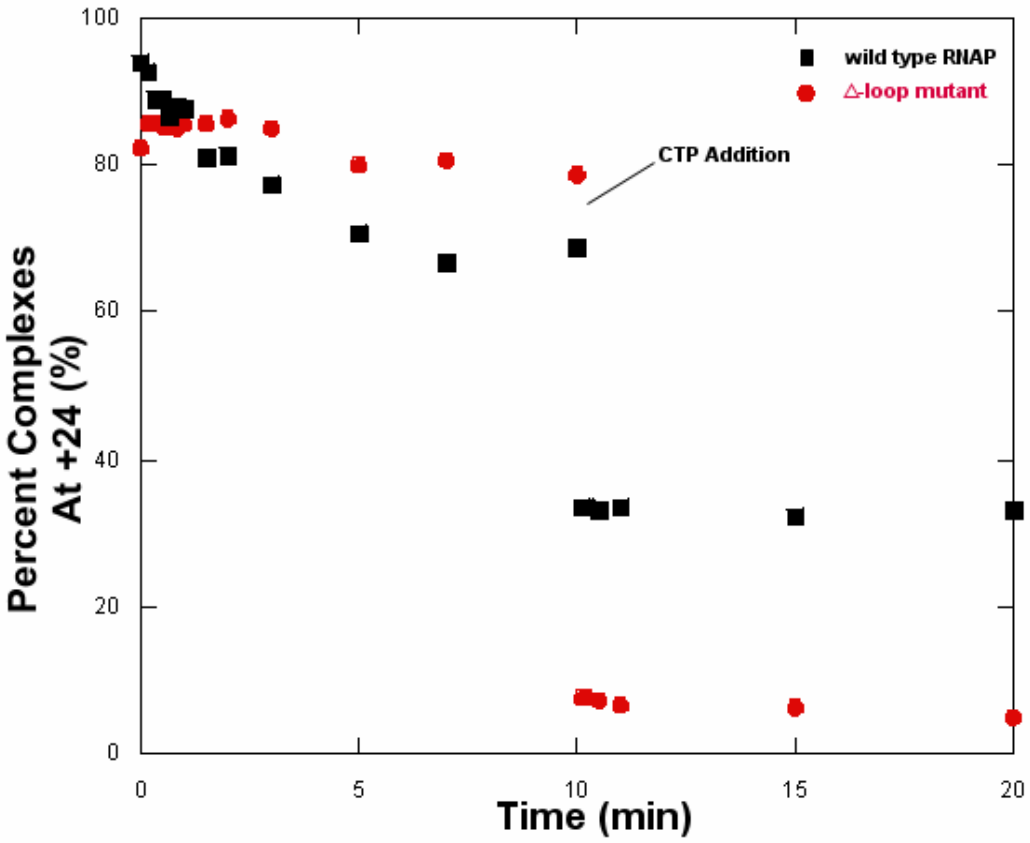


Figure 2.15: ***Plot of the disappearance of complexes out of position +24 following the addition of CTP after 10 minutes of misincorporation for wild type (black squares) and Δ-loop mutant (red circles).*** The chase reaction of CTP into position +25 occurs within the first seven seconds of CTP addition following the misincorporation reaction in wild type and in Δ-loop RNAP.

Discussion

Model of Incorrect Nucleotide Incorporation

Previous misincorporation studies observed that there is an activated synthesis, while a subset of complexes were falling off pathway into a non-productive state (Erie *et al.* 1993). These experiments were initiated from the promoter and used only 20μM and 1mM UTP in the presence of ATP and GTP. We have extended the misincorporation experiment by performing the misincorporation reaction using purified elongation complexes with a wide range of UTP concentrations. These experiments differ from the original misincorporation experiments in that the reactions are restarted after a stall as opposed to the running start afforded by the promoter initiated reactions. From these experiments, we demonstrate that there is an increasing rate and extent of misincorporation with increasing UTP concentration (Figures 2.6, 2.7, and 2.8). We also observe an incomplete misincorporation reaction at UTP concentrations less than 75μM (Figures 2.5 and 2.6). This result supports the original misincorporation data by Erie *et al.* (1993), suggesting that misincorporation can only occur during activated synthesis with a subset of complexes falling off the activated pathway into a non-productive synthesis pathway. We also demonstrate that the non-productive complexes are competent for elongation in the presence of the correct NTP and are therefore not dead-end complexes (Figures 2.5 and 2.9). Similar observations are seen for an RNAP mutant with four residues deleted from the fork loop 2 region of the protein (Δ-loop RNAP) (Figures 2.11 and 2.12). However, Δ-loop RNAP misincorporates at rates and extents less than that of wild type (Figures 2.13 and 2.14) with minimally affected correct incorporation kinetics which defines the Δ-loop RNAP as a higher fidelity enzyme and suggests that fork loop 2 has a significant role in the activated state of synthesis.

With the misincorporation data in hand, we attempted to determine the mechanism for the incorrect nucleotide incorporation into the growing RNA chain. Several models of nucleotide incorporation have been proposed and each model was used to fit our misincorporation data. We attempted first to fit our data to the original misincorporation mechanism proposed by Erie *et al.* (1993, Figure 2.1). Individually, we were able to fit each single exponential curve to a portion of the mechanism (Figure 2.16); however, no universal fit to the data could be obtained without the presence of a second NTP binding event.

We also tried to fit the data to the mechanism proposed for *H. sapiens* RNAP II for HDAg-stimulated elongation (Figure 2.17). This mechanism from Nedialkov *et al.* is similar to the original misincorporation mechanism proposed by Erie *et al.* (1993). The mechanism describes a conformational change in the enzyme (24 to 24*) equated to translocation of the enzyme along the DNA. This translocation is likely facilitated by the presence of the downstream NTP. The NTP can bind to the enzyme in state 24 or 24*. The state that has undergone the conformational transition (24*) is readily in a position for forward synthesis (activated state) with state 24 being slower to synthesize (unactivated state) (Nedialkov *et al.* 2003). We have presented evidence that misincorporation can occur only along an activated path of synthesis. Despite this similarity to the mechanism proposed by Nedialkov *et al.*, we were unable to fit our data to the mechanism shown in Figure 2.17. The HDAg-stimulated elongation mechanism is insufficient in describing the varying extents of misincorporation we observe. The mechanism does not describe any non-productive state of synthesis. All concentrations of UTP described in this work were simulated to 100% completion with the HDAg-stimulated elongation mechanism.

Additionally, a second mechanism proposed by Nedialkov *et al.* for TFIIF-stimulated elongation in *H. sapiens* RNAP II was tested as a potential mechanism for the incorrect nucleotide incorporation in *E. coli* RNAP (Figure 2.18) (2003). This mechanism is the same as the HDAg-stimulated elongation mechanism shown in Figure 2.17 with the additional requirement of a second NTP binding event which allows for a second path of synthesis. This mechanism is the same mechanism proposed by Holmes & Erie (2003), with the exception of the absence of the allosteric site in the mechanism proposed by Nedialkov *et al.* (2003). Instead, Nedialkov *et al.* propose that the downstream NTP pre-loads in the main channel and facilitates translocation. The mechanism describes two synthesis pathways: a faster pathway supported by higher NTP concentrations (activated state) and a slower pathway that dominates at lower NTP concentrations (unactivated state) (Nedialkov *et al.* 2003). The experiments performed with *H. sapiens* RNAP II differ from the *E. coli* RNAP experiments. Experiments with RNAP II use complexes stalled at C40 in a template where ATP and GTP are omitted from the reaction. The sequence after C40 is …AAAGG…and to monitor incorporation after the stall, the reaction is given both ATP and GTP. In this way, NTPs can pre-load and synthesize as described previously (Nedialkov *et al.* 2003). Our misincorporation reactions were carried out in the absence of the downstream NTP; however, experiments performed in the presence of the downstream NTP revealed no effect on the rate of misincorporation at position +25 in the nascent RNA chain. This result suggests that for misincorporation, *E. coli* RNAP is not pre-loading downstream NTPs and therefore the TFIIF-stimulated elongation mechanism can not describe our misincorporation data. This mechanism also does not explain the varying extents of misincorporation observed in *E. coli*

RNAP as all complexes go to 100% completion when simulated using the TFIIF-stimulated elongation mechanism for RNAP II.

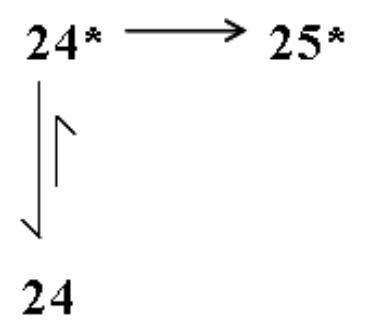


Figure 2.16: ***Mechanism adapted from the original misincorporation mechanism.*** This mechanism describes a non-productive unactivated state (24) shifting to an activated state (24*) where misincorporation can occur (25*) (Erie *et al.* 1993).

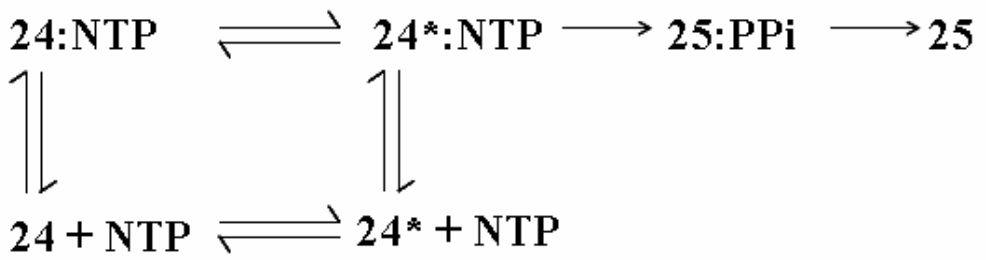


Figure 2.17: ***HDAg-stimulated elongation mechanism proposed by Nedialkov et al. (2003).*** The mechanism is similar to the original misincorporation mechanism by Erie *et al.* (1993). (24) represents an unactivated state of synthesis and (24*) represents the enzyme after a conformational change in the complex that facilitates activated synthesis.

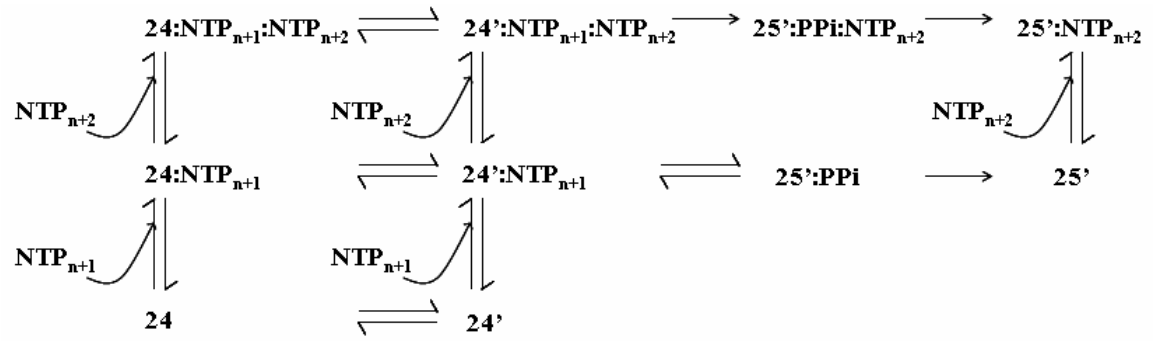


Figure 2.18: ***TFIIF-stimulated mechanism proposed by Nedialkov et al. (2003).*** The mechanism is similar to the mechanism proposed by Holmes and Erie (2003) with NTPs pre-loading in the main channel instead of binding into an allosteric site. The slower unactivated state (24':NTP_{n+1}) dominates at lower concentrations of NTP while the fast activated state (24':NTP_{n+1}:NTP_{n+2}) dominates at higher concentrations of NTP.

We attempted to simulate our misincorporation data using a fourth mechanism shown in Figure 2.19. In this mechanism, the motions of the F-bridge (F-helix) facilitate translocation and shift the RNAP into subsequent states for transcription elongation (Bar-Nahum *et al.* 2005). The pre-translocated state ($24_{PRE\text{-}TRANS}$) is in rapid equilibrium with the translocated state (24_{TRANS}) with the F-helix in a straight conformation. In the translocated state (24_{TRANS}), the enzyme is capable of substrate binding (24:S) where the complex can proceed to chemistry. However, bending of the F-helix can cause breaking of the 3' base pairing in the pre-translocated state ($24_{PRE\text{-}TRANS}$) shifting the complex into a trapped state (24_{TRAP}). This trapped state is susceptible to backtracking where the 3' end of RNA extrudes from secondary channel. Similar bending of the F-helix can facilitate translocation ($24_{PRE\text{-}TRANS}$ to 24_{FRAY}). The bent F-helix in this translocated state (24_{FRAY}) blocks access to the catalytic site through the secondary channel. Upon F-helix straightening, the RNAP enters into the translocated state capable of binding NTPs ($24_{PRE\text{-}TRANS}$). This mechanism proposed by Bar-Nahum *et al*. represents one productive path of synthesis and contains a built in non-productive synthesis state (24_{TRAP}). In this work, we have previously used a non-productive "trapped" state to describe our misincorporation data. However, we were unable to simulate reasonable fits to our data using the mechanism proposed by Bar-Nahum *et al.* (2005). The mechanism is set up for an equilibrium shifted to the post-translocated state where rapid equilibrium allows for 100% incorporation. As such, the mechanism could not simulate the varying extents of misincorporation with increasing UTP concentration we observe in our concentration-dependent kinetics.

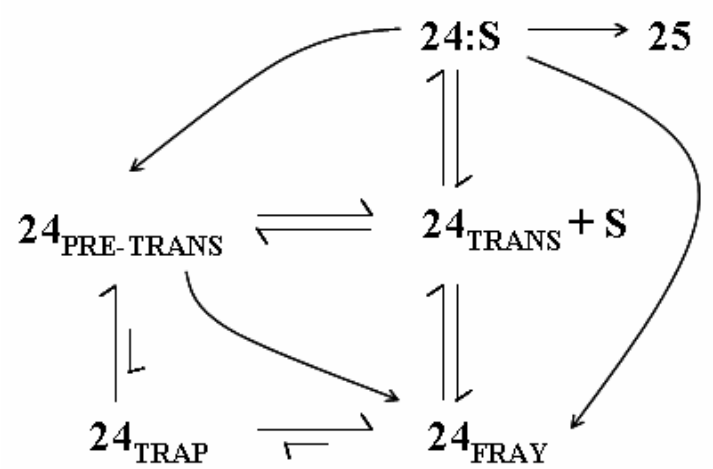
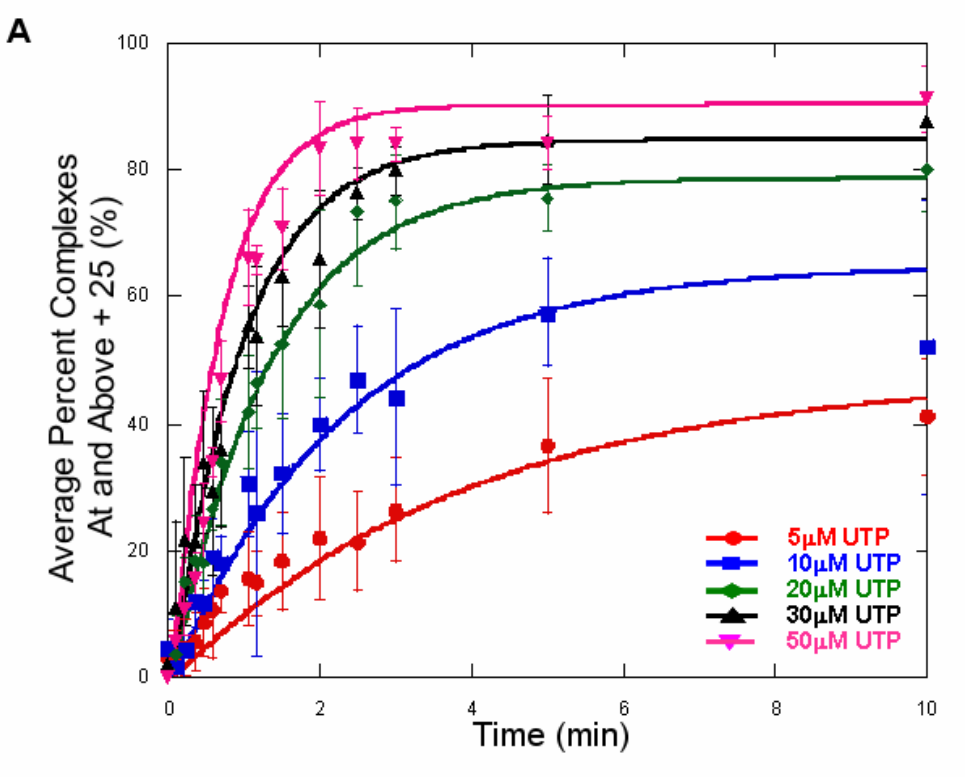


Figure 2.19: ***Ratchet mechanism proposed to describe correct and incorrect nucleotide incorporation in E. coli RNAP (Bar-Nahum et al. 2005).*** Motions of the F-helix play a central role in this mechanism of transcription elongation, facilitating translocation and shifting the RNAP into subsequent states.

We finally attempted to simulate the misincorporation data using our current model for transcription elongation (Figure 2.2A). This non-essential activation mechanism assumes a rapid equilibrium between the pre- and post-translocated states and utilizes binding of an NTP into the allosteric site to change the conformation of the RNAP such that the enzyme enters into the activated state of synthesis. Our misincorporation data suggests that misincorporation can only occur along this activated synthesis path; thus, we assumed zero synthesis in the unactivated state ($k_{slow} = 0$). Our data also suggests that NTPs may bind the catalytic site in the unactivated state; however, complexes with an NTP bound in the unactivated state are "trapped" in this state without entering into a dead end state of the enzyme. If the complexes were not being trapped, we would expect to see 100% misincorporation for all concentrations of UTP after a sufficient length of time. We do not observe 100% misincorporation until UTP concentrations > 75μM even after forty minutes of reaction. We allowed k_{-unact} to equal zero, thereby "trapping" non-productively bound NTPs in the unactivated state. Data were fit to the simulated non-essential activation mechanism with a single set of rate constants (Figures 2.20 and 2.21). Simulated rates of reaction for the mechanism of misincorporation are shown in Table 2.1.


A
Average Percent Complexes
At and Above + 25 (%)
Time (min)
5μM UTP
10μM UTP
20μM UTP
30μM UTP
50μM UTP

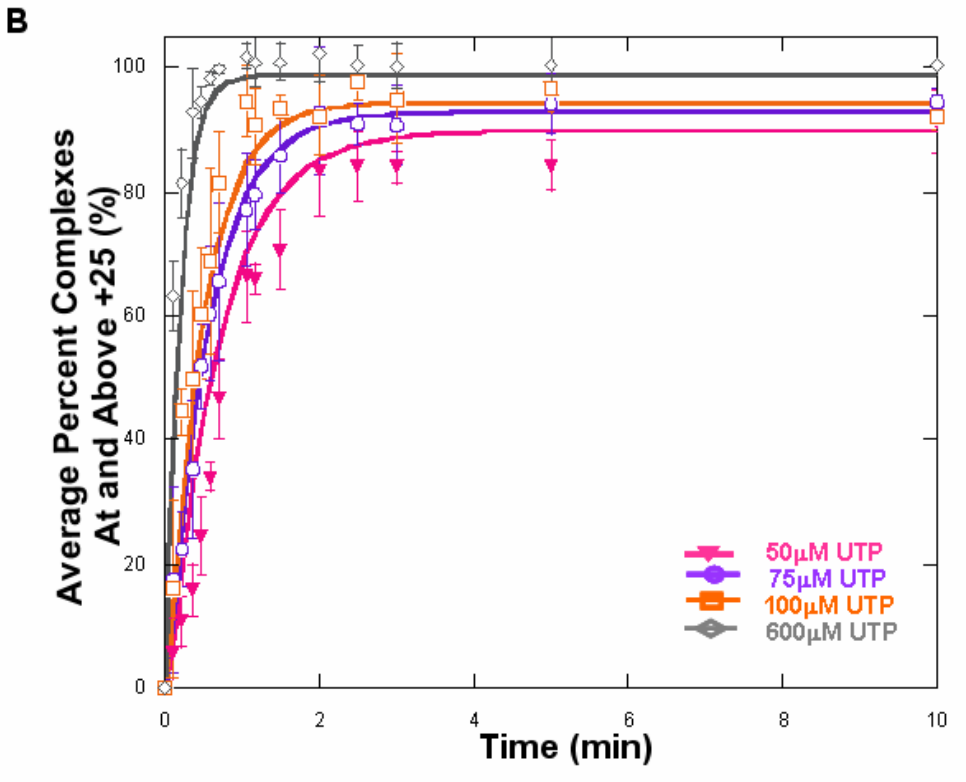

B
Average Percent Complexes
At and Above +25 (%)
Time (min)
50μM UTP
75μM UTP
100μM UTP
600μM UTP

Figure 2.20: ***Average kinetic data for UTP concentrations (A) 5-50μM and (B) 50-600μM fit by a single set of rate constants to the non-essential activation mechanism in Figure 2.21.*** Error bars indicate standard deviation for three to five trials for each concentration of UTP.

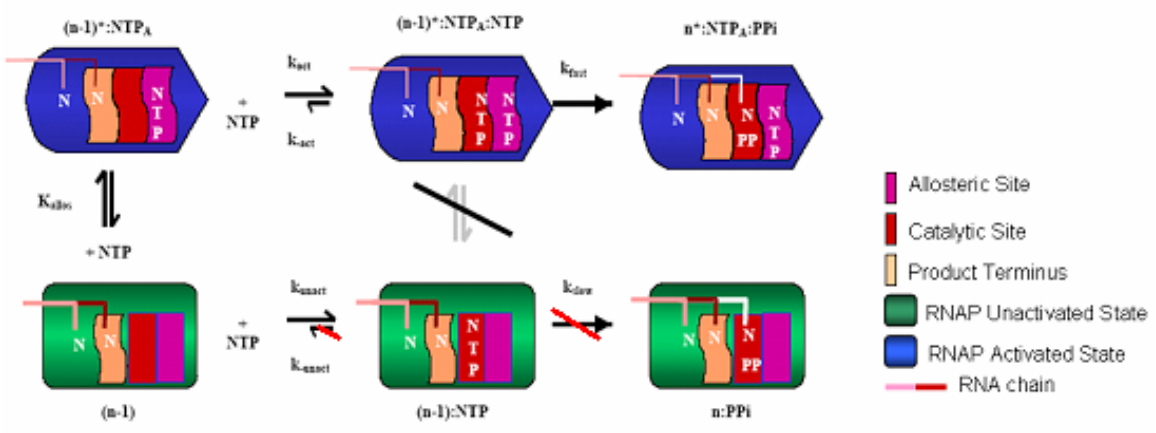


Figure 2.21: ***Non-essential activation mechanism describing misincorporation by E. coli RNAP.*** The green and blue boxes are the RNAP in the unactivated and activated states respectively. The magenta box represents the allosteric site, while the red box indicates the catalytic site. The product-terminus binding site is shown in the peach box. The red and pink lines indicate the growing RNA chain. NTP represents substrate bound to the catalytic site while NTP_A represents substrate bound to the allosteric site. As indicated by the red bars, NTP binding first into the catalytic site of RNAP represents non-productive binding where the rate of unactivated state synthesis (k_{slow}) is zero. The rate of escape from the unactivated state (k_{-unact}) is also zero, suggesting that the incorrect NTP is indeed "trapped" in the catalytic site for this non-productive state of synthesis.

System	K_{allos} (μM)	k_{unact} (μM^{-1}, s^{-1})	k_{-unact} (s^{-1})	k_{slow} (s^{-1})	k_{act} ($\mu M^{-1}s^{-1}$)	k_{-act} (s^{-1})	k_{fast} (s^{-1})
correct incorporation	100	5.8	33	2.7	4300	91000	730
incorrect incorporation	3.4	$1.1x10^{-3}$	0	0	$2.1x10^{-3}$	0.13	0.063

Table 2.1: ***Simulated rates of the non-essential activation mechanism described in Figure 2.21.*** Correct incorporation rates previously determined are shown for comparison (Holmes & Erie 2003). The rate constants refer to those shown in Figure 2.21.

Inspection of the rates from our misincorporation data to the mechanism shown in Figure 2.21 reveals that misincorporation occurs at rates that are orders of magnitude (10^3 – 10^6) slower than correct incorporation (Table 2.1). Also, for incorrect nucleotide addition, the rate of k_{slow} (unactivated synthesis) is zero. Perhaps the most interesting observation is the rate of k_{-unact} is also zero. A rate of zero for k_{slow} and k_{-unact} suggests that an NTP is binding to the catalytic site for unactivated synthesis, but the NTP binds non-productively and "irreversibly." This "irreversible" binding explains the observation that misincorporation does not go to completion (100%). There is a subset of complexes that is entering into the unactivated state and getting "trapped." These complexes are not, however, dead-end complexes (Chase Reactions). The observation that the "trapped" NTP can be chased in the presence of the correct NTP indicates that although the complexes do not catalyze incorporation of the incorrect base, they are fully functional to incorporate the correct base. These results also indicate that UTP has not caused the complexes to enter a backtracked state because backtracked states are slow to recover and a slow incorporation of the correct NTP is not observed (Figure 2.9).

Also notable, the binding constant for the allosteric site is 30 times tighter than that of correct incorporation. It has been suggested that there is a negative-cooperativity between the allosteric site and the catalytic site (Foster *et al.* 2001). In negative-cooperativity of allosteric enzymes, binding of each substrate molecule decreases the intrinsic affinities of the substrate for the vacant sites (Segel 1975). A tight binding of the incorrect nucleotide to the allosteric site decreases the affinity of the incorrect nucleotide for the catalytic site. This binding would be a way of modulating the fidelity of RNAP as tight binding of the incorrect nucleotide into the catalytic site would likely increase the amount of misincorporation

observed. This tight binding of the incorrect nucleotide to the allosteric site is seemingly a key factor in fidelity. The rates into the unactivated (k_{unact}) and activated (k_{act}) states are similar ($1.1x10^{-3}$ and $2.1x10^{-3}$, respectively.) However, comparing the rates of misincorporation for k_{unact} and k_{act} to correct incorporation, the rate into the activated state is reduced by a factor of 10^6 while the rate into the unactivated state experiences a 10^3-fold decrease. This dramatic reduction in the rate of NTP binding to the catalytic site in the activated state suggests that NTP binding of the catalytic site in this state is largely affecting the fidelity of RNAP.

Structural Model of Incorrect Nucleotide Incorporation – Activated State Synthesis

The structural model of nucleotide addition during transcription elongation recently proposed by Kennedy and Erie (manuscript in preparation) taken together with the misincorporation data presented here leads us to propose a structural model for misincorporation similar to that proposed for correct incorporation with several key differences (Figure 2.22). We have presented compelling evidence that misincorporation can only occur in the activated (fast) state of synthesis. This activated state is achieved by a change in the conformation of the RNAP facilitated by NTP binding to an allosteric site located on fork loop 2 in the main channel. For an incorrect nucleotide to be incorporated into the growing RNA chain, the incorrect NTP must bind first to the allosteric site. We suggest that this NTP is acting allosterically on both the RNAP and the DNA. Binding of the incorrect NTP to fork loop 2 serves to shift the conformation of the RNAP into the activated state but also facilitates proper alignment of the DNA by translocation which allows the n+1 base to align in the catalytic site. The NTP bound to the allosteric site is acting as a check for

incorporation. With the incorrect NTP bound, the alignment of the DNA in the catalytic site proceeds at rates slower than that of correct incorporation as the enzyme is attempting to properly align an incorrect nucleotide for incorporation. Once the DNA is properly aligned, however, a second incorrect NTP can enter into the catalytic site via the secondary channel. Binding of the NTP in the catalytic site loosens the trigger loop interacting with the allosteric NTP such that the trigger loop closes over the catalytic site, further aligning the base in a way that allows for chemistry to occur; thereby, incorporating the incorrect base.

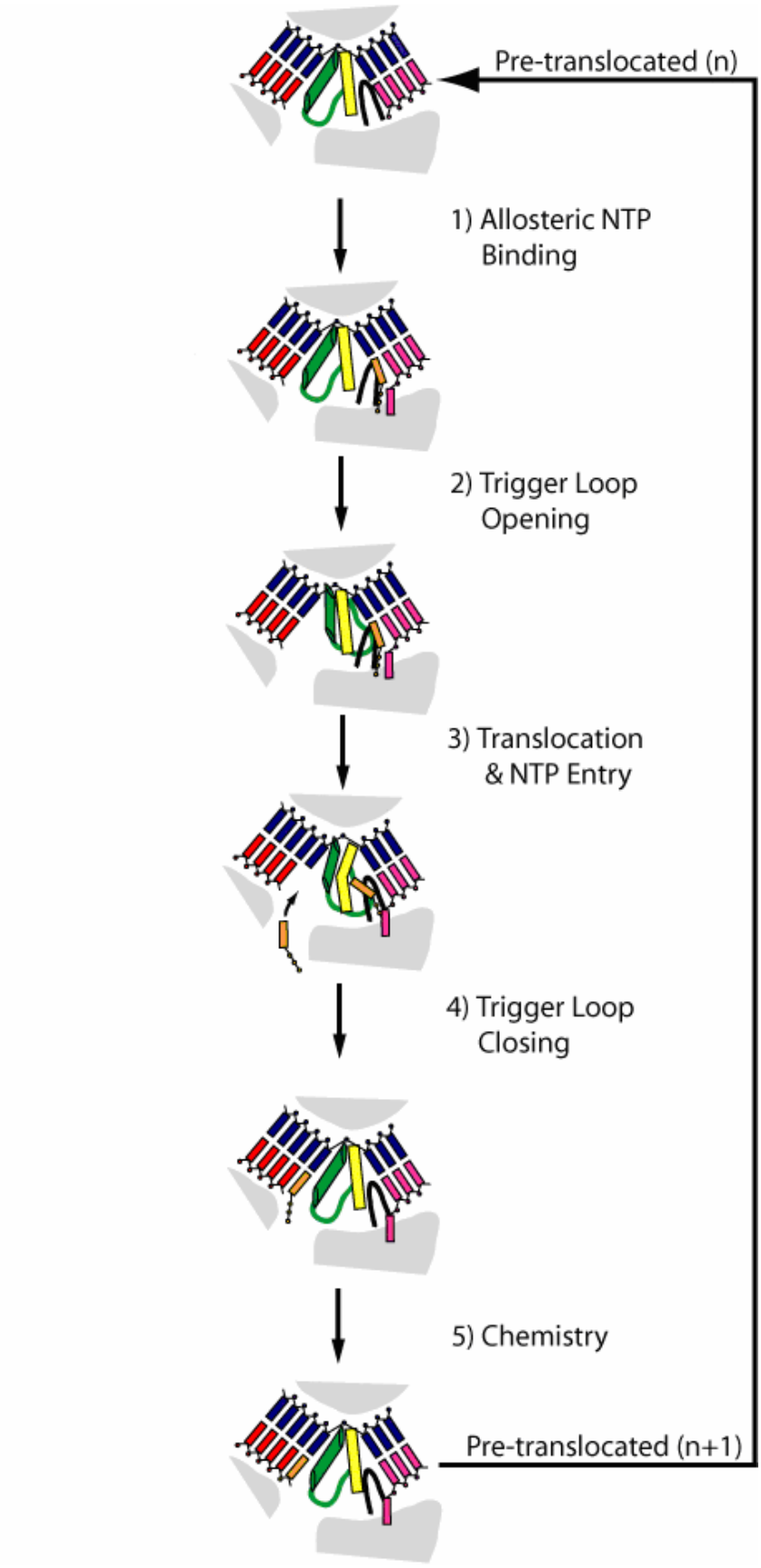

Pre-translocated (n)
1) Allosteric NTP Binding
2) Trigger Loop Opening
3) Translocation & NTP Entry
4) Trigger Loop Closing
5) Chemistry
Pre-translocated (n+1)

Figure 2.22: ***An approximate model for activated state incorporation of an incorrect nucleotide into the nascent RNA chain.*** The DNA template strand is shown in blue with the non-template strand shown in pink. The RNA chain is red. The incorrect NTP is orange. Fork loop 2 is black, the bridge helix (F-helix) is yellow, and the trigger loop is shown in green. This model is adapted from the correct nucleotide incorporation model during transcription elongation proposed by Kennedy and Erie (manuscript in preparation).

Structural Model of Incorrect Nucleotide Incorporation – Non-Productive Binding in the Unactivated State

Recent crystal structures have shown the previously mentioned trigger loop in a closed and open conformation (Figure 2.23) (Toulokhonov *et al.* 2007; Vassylyev *et al.* 2007). The trigger loop in the closed conformation interacts with the NTP in the catalytic site and closes over the NTP for synthesis. The closed trigger loop has also been shown to block access to the catalytic site via the secondary channel. When considering the non-productive binding of the NTP in the catalytic site during misincorporation we considered the trigger loop closing as part of our structural model (Figure 2.24). An incorrect NTP entering into the catalytic site binds non-productively in such a way that the NTP is not properly aligned with the +1 DNA base for synthesis. We propose that the incorrect NTP actually frays the DNA such that the +1 base is better aligned with fork loop 2, the putative allosteric site. With the NTP bound in the catalytic site, the trigger loop closes over the incorrectly bound NTP and interacts with the NTP in such a way that the trigger loop is "locked" into the closed conformation. This closing of the trigger loop would prevent the incorrect NTP in the catalytic site from escaping via the secondary channel thereby trapping the NTP in the non-productive state of synthesis.

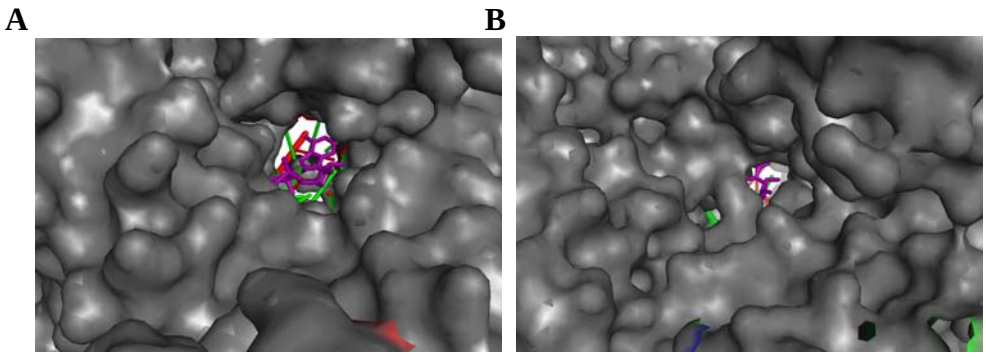


Figure 2.23: ***The open and closed conformations of the trigger loop affect the accessibility of the catalytic site through the secondary channel.*** The opened secondary channel (A, PDB 2PPB) allows an NTP (purple) access to the active site whereas a closed secondary channel (B, PDB 2O5J) would restrict access and consequently escape (adapted from Kennedy & Erie, manuscript in preparation).

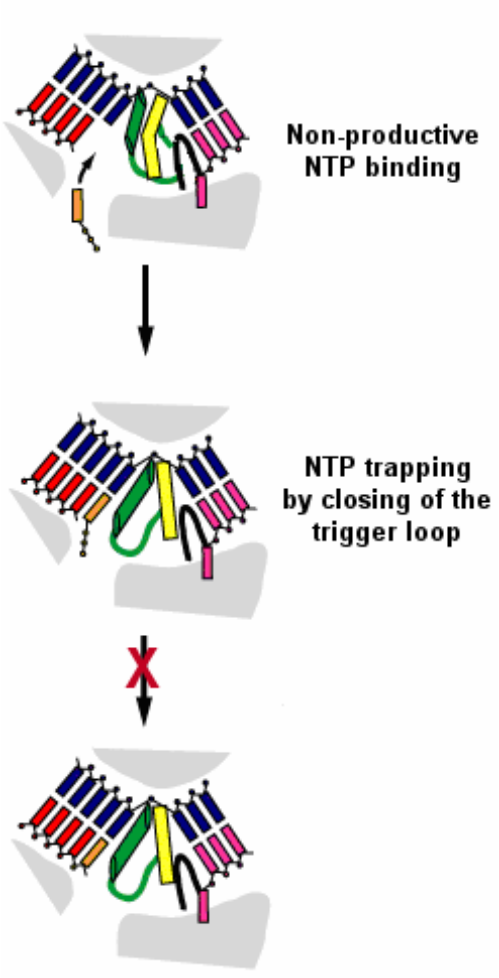


Figure 2.24: ***An approximate model for a non-productive and "irreversibly" bound NTP in the unactivated state.*** The DNA template strand is shown in blue with the non-template strand shown in pink. The RNA chain is red. The incorrect NTP is orange. Fork loop 2 is black, the bridge helix (F-helix) is yellow, and the trigger loop is shown in green. This model is adapted from the correct nucleotide incorporation model during transcription elongation proposed by Kennedy and Erie (manuscript in preparation).

Active Displacement of the Non-Productively Trapped NTP by the Correct Nucleotide

We have shown that the subset of complexes that are not undergoing misincorporation ("trapped" complexes) can be chased to completion in the presence of the correct NTP (Figures 2.5, 2.9, 2.10, 2.11, and 2.15). The complexes not undergoing misincorporation, according to the proposed mechanism, are "trapped" in the unactivated state by a non-productively bound NTP in the catalytic site. The evidence that these "trapped" complexes can be chased in the presence of the correct NTP suggests that there is an active displacement of the incorrect "irreversibly" bound NTP from the catalytic site in the presence of the correct NTP. Considering the structural model presented, we propose that for the chase reaction of the "trapped" NTP, the correct NTP comes in through the main channel and binds to the available allosteric site (fork loop 2). As mentioned, the incorrect NTP bound in the catalytic site may be fraying the DNA such that the +1 base is aligned with the allosteric site, affording us the specificity needed to explain the ability of the correct NTP to actively displace the incorrect NTP from the catalytic site. The correct NTP binding to fork loop 2 acts allosterically to change the conformation of the RNAP as well as the alignment of the DNA within the main channel of the enzyme. This alignment is likely responsible for shifting the incorrect NTP in the catalytic site in such a way that the NTP is freed from the catalytic site. The movement of the incorrect NTP affects the interaction of the NTP with the trigger loop, releasing the trigger loop from the closed conformation over the catalytic site. This release of the trigger loop opens up the secondary channel such that the previously non-productively bound NTP can exit. The correct NTP remains bound to the allosteric site throughout this active displacement of the correct NTP and is now readily available to incorporate. The correct NTP can incorporate via the slow unactivated state of

synthesis where the correct NTP is transferred via a hand off mechanism from fork loop 2 by the trigger loop into the catalytic site; or, the correct NTP can be incorporated via the fast activated phase of synthesis when a second correct NTP enters the catalytic site through the available secondary channel (Kennedy & Erie, manuscript in preparation).

In this work, we have successfully characterized the concentration-dependent kinetics of misincorporation by *E. coli* RNAP as well as *E. coli* Δ-loop RNAP. We have presented evidence for the presence of an allosteric site in the RNAP located on fork loop 2, which upon NTP binding is responsible for shifting the RNAP into a fast state (activated state) of synthesis. We have demonstrated that misincorporation can only occur in this activated state of synthesis while a subset of complexes during misincorporation enter into a non-productive "trapped" state where an NTP is bound "irreversibly" in the catalytic site. We have demonstrated that these "trapped" complexes are capable of synthesis in the presence of the correct NTP and therefore we propose the active displacement of NTPs during transcription elongation. The proposed model for misincorporation and subsequent active displacement of NTPs is likely to be further supported by continued work with the wild type *E. coli* RNAP and Δ-loop mutant RNAP.

Experimental Procedures

Sources of protein and DNA

His-tagged wild type RNAP was purified from log phase cells of strain RL916 (gift of R. Landick) as described previously (Burgess & Jendrisak 1975; Uptain & Chamberlain 1997). β-Δ(R542-F545) mutant RNA polymerase was made by standard molecular biology techniques on the pRL-706 plasmid. Expression was carried out in the *E. coli* strain

TOM100 (gift of T. Santangelo) and purified as previously described. (Santangelo *et al.* 2003). The DNA template was prepared from pDE13 and amplified by PCR. The biotinylated 540 nucleotide fragment contains the λPR promoter and codes for a transcript in which the first cytosine to be incorporated is at +25 as indicated below:

+25
pppAUGUAGUAAGGAGGUUGUAUGGAA**C**AACGCAUAACCCUGA…

In vitro transcription reactions – misincorporation from promoter initiation

RNAP (60nM) and 5'-biotinylated DNA template (60nM) were incubated for 10 minutes at 37°C in 1X transcription buffer (30mM HEPES (pH 8.0), 10mM Mg^{2+}-glutamate, 200mM K^{+}-glutamate, 25μg/mL BSA, and 1mM DTT) to form open promoter complexes. Transcription was initiated by adding 15μM UTP, 20μM ATP, and 20μM [α-^{32}P] GTP (160Ci/mmol). The reaction was monitored over time at room temperature (~23°C). Reactions were quenched using 100% formamide and products were separated on 20% acrylamide (19:1), 8M urea denaturing gels.

In vitro transcription reactions – purified stalled elongation complexes

RNAP (60nM) and 5'-biotinylated DNA template (60nM) bound to streptavidin-coated magnetic beads were incubated for 10 minutes at 37°C in 1X transcription buffer (30mM HEPES (pH 8.0), 10mM Mg^{2+}-glutamate, 200mM K^{+}-glutamate, 25μg/mL BSA, and 1mM DTT) to form open promoter complexes. Complexes stalled at position +24 were formed by adding 20μM UTP, 20μM ATP, and 20μM [α-^{32}P] GTP (160Ci/mmol) and incubating at room temperature for 35 seconds (1.5 minutes for Δ-loop RNAP). The complexes were washed ten to fifteen times using ice-cold 1X transcription buffer by holding the reaction tube next to a strong magnet to retain the complexes. Complexes were

resuspended in 1X transcription buffer, aliquoted for the different reactions, and stored on ice. To ensure that the results were not dependent on the time complexes remained on ice, reactions were carried out in a different order with each experiment. Kinetic experiments were carried out by hand at room temperature. Misincorporation reactions were initiated by the addition of the indicated concentration of UTP to the purified SECs. Reactions were quenched using 100% formamide and products were separated on 20% acrylamide (19:1), 8M urea denaturing gels. UTP concentrations reported are the final concentration in solution.

In order to test the effect on misincorporation of pre-incubation and simultaneous addition with the downstream DNA base, purified stalled elongation complexes were formed as described previously. A fraction of these complexes were then pre-incubated with 2.5mM ATP for one minute prior to addition of 20μM UTP. The remaining fractions were given 2.5mM ATP + 20μM UTP simultaneously. Reactions were monitored over time and quenched using 100% formamide before products were separated on 20% acrylamide (19:1), 8M urea denaturing gels. Concentrations reported are final concentrations in solution.

Chase reactions

At designated times during the *in vitro* transcription reactions, a sample of the reaction was added to the presence of all four NTPs (1mM) to extend the transcript to full length and ensure that the complexes were still active. Further chase reactions were carried out using only CTP. The *in vitro* transcription reaction with purified stalled complexes was performed as described, adding 20μM UTP to the purified SECs. The misincorporation reaction went for 10 minutes before addition of 5μM, 50μM, or 1mM CTP. The CTP addition reactions were monitored by hand for 10 minutes.

For the UTP wash experiment, to test if addition of the incorrect NTP after misincorporation could also restart the reaction, the *in vitro* transcription reaction was carried out as previously described with 20μM UTP. Complexes were allowed to misincorporate for 10 minutes before the reaction was stopped by placing the tube in a cold magnet. Unreacted UTP was washed away and the complexes were resuspended in the appropriate volume of 1X transcription buffer. UTP (20μM) was added back to the reaction and the transcription reaction was monitored for another 10 minutes. Reactions were quenched in 100% formamide and all aliquots were run on 20% acrylamide (19:1), 8M urea denaturing gels.

Rapid quench chase reactions

Purified stalled elongation complexes were made as described previously. The misincorporation reaction was initiated by adding 20μM UTP. This reacted for 10 minutes before the reaction was stopped by placing the tube on ice. Rapid quench experiments were carried out on a Kintek Rapid Quench Flow 3 apparatus at room temperature. For each time point, 20μL of complexes were injected into one reactant loop and 20μL of the designated concentration of CTP (10 and 100μM) was injected into the other reactant loop. Reactants were mixed for the desired amount of time and quenched with 0.5mM EDTA. Each time point represents a different experiment. To assure that the results were not dependent on the time complexes remained on ice, time points were carried out in a different order. All products were run on 20% acrylamide (19:1), 8M urea denaturing gels.

Data Analysis

Quantification and normalization of rate data

The amount of radioactivity in each lane of the gels was measured on an Amersham Biosciences PhosphorImager and analyzed with ImageQuant software. The percentage of

complexes at each position on the template was calculated by dividing the amount of radioactivity in the indicated band by the total amount of radioactivity in all the bands +24 nucleotides in length and longer. To compare data from different experiments, it was necessary to normalize the data such that at time 0, there was 0% incorporation. Due to the incomplete misincorporation reaction, the maximum extent of incorporation could not be normalized to 100%. To normalize these data, the maximum extent of incorporation determined by the single exponential fit to the data was used as the maximum for each concentration. The experiments were conducted three to five times for each concentration.

Fits of the kinetic data to the mechanism

For the wild type enzyme, each data set was fit to the single-exponential equation using Kaleidagraph v4.01. The data from the single-exponential fits of the individual rate curves were used as a starting point to obtain initial values for binding constants to the catalytic and allosteric sites and the rate constants for the unactivated and activated states as previously described (Foster *et al.* 2001). For the non-essential activation mechanism and all other mechanisms attempted, KinSim (Anderson *et al*. 1988) was used to fit the data "manually" – meaning the data were simulated using many combinations of rate and binding constants until the best fits were obtained. A second program designed by Cherie Lanyi (UNC) using MatLAB was used to verify the fits to the data given the single set of rate constants obtained from the manual fit in KinSim, fitting all UTP data simultaneously. An exhaustive combination of rates was tested, though we cannot say with absolute certainty that the final rates are the only set of numbers that fit the data.

The same procedure was used to fit the Δ-loop mutant RNAP to the current non-essential activation mechanism for misincorporation. The rate and binding data indicate a

100-fold decrease in the rate with a concomitant 10-fold decrease in the binding affinity of the NTP for the allosteric site. Simulating a 10-fold decrease in the rate of fast synthesis (k_{fast}) yields simulated curves that are a near fit to the experimental data, but at the time of this work an exact fit has not been obtained.

Bibliography

Alic, N., Ayoub, N., Landrieux, E., Favry, E., Baudouin-Cornu, P., Riva, M. and Carles C. (2007). "Selectivity and proofreading both contribute significantly to the fidelity of RNA polymerase III transcription." Proc.Natl.Acad.Sci.U.S.A. **104**(25):10400-5.

Anderson, K.S., Sikorski, J.A. and Johnson, K.A. (1988). "A tetrahedral intermediate in the EPSP synthase reaction observed by rapid quench kinetics." Biochemistry **27**(19): 7395-406.

Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A. and Nudler, E. (2005). "A ratchet mechanism of transcription elongation and its control." Cell **120**(2): 183-193.

Burgess, R.R. and Jendrisak, J.J. (1975). "A procedure for the rapid, large-scale purification of Escherichia coli DNA-dependent RNA polymerase involving Polymin P precipitation and DNA- cellulose chromatography." Biochemistry **14**(21): 4634-4638.

Coulombe, B. and Burton, Z.F. (1999). "DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms." Microbiol.Mol.Biol.Rev. **63**(2): 457-478.

Davenport, R.J., Wuite, G.J., Landick, R. and Bustamante, C. (2000). "Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase." Science **287**(5462): 2497-2500.

Echols, H. and Goodman, M.F. (1991) **"**Fidelity mechanisms in DNA replication." Annu.Rev.Biochem. **60**: 477-511.

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): 224-39.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C. and von Hippel, P.H. (1993). "Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription." Science **262**(5135): 867-873.

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

Foster, J.E., Holmes, S.F. and Erie, D.A. (2001). "Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation." Cell **106**(2): 243-252.

Gong, X.Q., Zhang, C., Feig, M. and Burton, Z.F. (2005). "Dynamic error correction and regulation of downstream bubble opening by human RNA polymerase II." Molecular Cell **18**(4): 461-470.

Guthold, M. and Erie, D.A. (2001). "Single-molecule study reveals a complex *E. coli* RNA polymerase." Chembiochem **2**(3): 167-70.

Holmes, S.F. and Erie, D.A. (2003). "Downstream DNA sequence effects on transcription elongation: Allosteric binding of nucleoside triphosphates facilitates translocation via a ratchet motion." J.Biol.Chem. **278**(37): 35597-35608.

Kennedy, S.K. and Erie, D.A. (2007). "In preparation."

Kubori, T. and Shimamoto, N. (1996). "A branched pathway in the early stage of transcription by *Escherichia coli* RNA polymerase." J.Mol.Biol. **256**(3): 449-57.

Kull, F.J., Vale, R.D. and Fletterick, R.J. (1998). "The case for a common ancestor: kinesin and myosin motor proteins and G proteins." J.Muscle Res.Cell Motil. **19**(8): 877-86.

Kunkel, T.A. and Bebenek, K. (2000) "DNA replication fidelity." Annu.Rev.Biochem. **69**:497-529.

Leipe, D.D., Wolf, Y. I., Koonin, E. V. and Aravind, L. (2002). "Classification and evolution of P-loop GTPases and related ATPases." J.Mol.Biol. **317**(1): 41-72.

Libby, R.T. and Gallant, J.A. (1991). "The role of RNA polymerase in transcriptional fidelity." Mol.Microbiol. **5**: 999-1004.

Matsuzaki, H., Kassavetis, G.A. and Geiduschek, E.P. (1994). "Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III." J.Mol.Biol. **235**(4): 1173-1192.

Mooney, R.A., Artsimovitch, I. and Landick, R. (1998). "Information processing by RNA polymerase:Recognition of regulatory signals during RNA chain elongation." J.Bacteriol. **180**(13):3265-75.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H., Yan, H. and Burton, Z.F. (2003). "NTP-driven translocation by human RNA polymerase II." J.Biol.Chem. **278**(20): 18303-18312.

Santangelo, T.J., Mooney, R.A., Landick, R. and Roberts, J.W. (2003). "RNA polymerase mutations that impair conversion to a termination-resistant complex by Q antiterminator proteins." Genes and Development **17**(10): 1281-1292.

Schulz, A.R. (1994). Enzyme Kinetics: From Diastase to Multi-enzyme Systems. NewYork, NY, Cambridge: Cambridge University Press.

Segel, Irwin H. (1993). Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems. New York, John Wiley and Sons, Inc.

Shaw, R.J., Bonawitz, N.D. and Reines, D. (2002) "Use of an *in vivo* reporter assay to test for transcriptional and translational fidelity in yeast." J.Biol.Chem. **277**(27):24420-6.

Taddei, F., Hayakawa, H., Bouton, M.-F., Cirinesi, A.-M., Matic, I. Sekiguchi, M. and Radman, M. (1997). "Counteraction by MutT protein transcriptional errors caused by oxidative damage." Science **278**(5335):128-30

Thomas, M.J., Platas, A.A. and Hawley, D.K. (1998). "Transcriptional fidelity and proofreading by RNA polymerase II." Cell **93**(4):627-37.

Tolić-Nørrelykke, S.F., Engh, A.M., Landick, R. and Gelles, J. (2004). "Diversity in the rates of transcript elongation by single RNA polymerase molecules." J.Biol.Chem. **279**(5): 3292-3299.

Toulokhonov, I., Zhang, J., Palangat, M. and Landick, R. (2007). "A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing." Cell **27**(3):406-19.

Uptain, S.M. and Chamberlin, M.J. (1997). "*Escherichia coli* RNA polymerase terminates transcription efficiently at rho-independent terminators on single-stranded DNA templates." Proc.Natl.Acad.Sci.U.S.A **94**(25): 13548-13553.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I. and Landick, R. (2007). "Structural basis for substrate loading in bacterial RNA polymerase." Nature **448**(7150): 163-168.

Via, A., Ferre, F., Brannetti, B., Valencia, A. and Helmer-Citterich, M. (2000). "Three-dimensional view of the surface motif associated with the P-loop structure: cis and trans cases of convergent evolution." J.Mol.Biol. **303**(4): 455-65.

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." Embo.J. **1**(8): 945-51.

Wong, I., Patel, S.S. and Johnson, K.A. (1991). "An induced-fit kinetic mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics." Biochemistry **30**(2):526-37.

Yin, H., Artsimovitch, I., Landick, R. and Gelles, J. (1999). "Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules." Proc.Natl.Acad.Sci.U.S.A **96**(23): 13124-13129.

Zhang, C., Zobeck, K.L. and Burton, Z.F. (2005). "Human RNA polymerase II elongation in slow motion: Role of the TFIIF RAP74 α1 helix in nucleoside triphosphate-driven translocation." Mol.Cell.Biol. **25**(9): 3583-3595.

CHAPTER 3: KINETIC INVESTIGATION OF MISINCORPORATION UTILIZING *ESCHERICHIA COLI* RNA POLYMERASE WITH MUTATIONS IN THE SECONDARY CHANNEL

Introduction

Crystal structures of both prokaryotic and eukaryotic RNA polymerases have revealed a funnel shaped pore that leads directly into the active site of the enzyme. This channel is 10-12Å in diameter and 45Å in length, which makes the channel large enough to accommodate one diffusing NTP at a time (Zhang *et al.* 1999; Korzheva *et al.* 2000). This pore is known as the secondary channel and has been considered by some investigators as the primary means of NTP entry into the catalytic site for nucleotide binding and incorporation during transcription (Zhang *et al.* 1999; Korzheva *et al. 2*000; Cramer *et al.* 2000; Cramer *et al.* 2001; Gnatt *et al.* 2001; Vassylyev *et al.* 2002; Batada *et al.* 2004; Kettenberger *et al.* 2004; Westover *et al.* 2004; Armache *et al.* 2005; Temiakov *et al*. 2005). Other researchers have proposed that the primary pathway for NTP entry into the catalytic site is through the main channel (Nedialkov *et al.* 2003; Burton *et al.* 2005; Gong *et al.* 2005; Zhang *et al.* 2005). Recent work by Kennedy and Erie has demonstrated that, while NTPs can come in through the main channel, the secondary channel does in fact play an important role in nucleotide incorporation and transcription elongation (manuscript in preparation).

This secondary channel is also believed to play a role in regulation of transcription elongation. Specifically, the secondary channel is believed to function as an extrusion point

of RNA during backtracking (Zhang *et al.* 1999; Artsimovitch & Landick 2000; Toulme *et al.*2000). Backtracking is the process in which RNAP translocates backwards along the DNA template displacing the 3' end of the RNA transcript from the catalytic site (Reeder & Hawley 1996; Komissarova & Kashlev 1997; Nudler *et al.* 1997). The extrusion of the RNA through the secondary channel provides the substrate for GreA and GreB induced cleavage and thereby plays a role in the regulation of RNAP during transcription elongation (Komissarova & Kashlev 1997; Artsimovitch & Landick 2000; Toulme *et al.* 2000).

In an attempt to understand the role of the secondary channel in NTP binding and transcription elongation, several mutant RNAPs were created with single or double amino acid substitutions of residues that are surface exposed at the junction of the secondary channel and the active site (Santangelo *et al.* 2003). These mutations were originally shown to disrupt Q-mediated antitermination both *in vivo* and *in vitro* without impairing the basic enzymatic activity of RNAP (Santangelo *et al.* 2003). More recently, these mutations were used in transient-state kinetic studies by Holmes *et al. (2*006). Correct and incorrect incorporation kinetic assays were performed and while none of the mutations significantly affected correct incorporation (Figure 3.1A and 3.1B), one amino acid substitution, βD675Y (*E. coli*), was determined to be a lower fidelity variant, significantly increasing the amount of misincorporation observed by the enzyme initiated from the promoter (Figure 3.2). Specifically, wild type misincorporated up to 20% while βD675Y misincorporated up to 80% (Holmes *et al.* 2006).

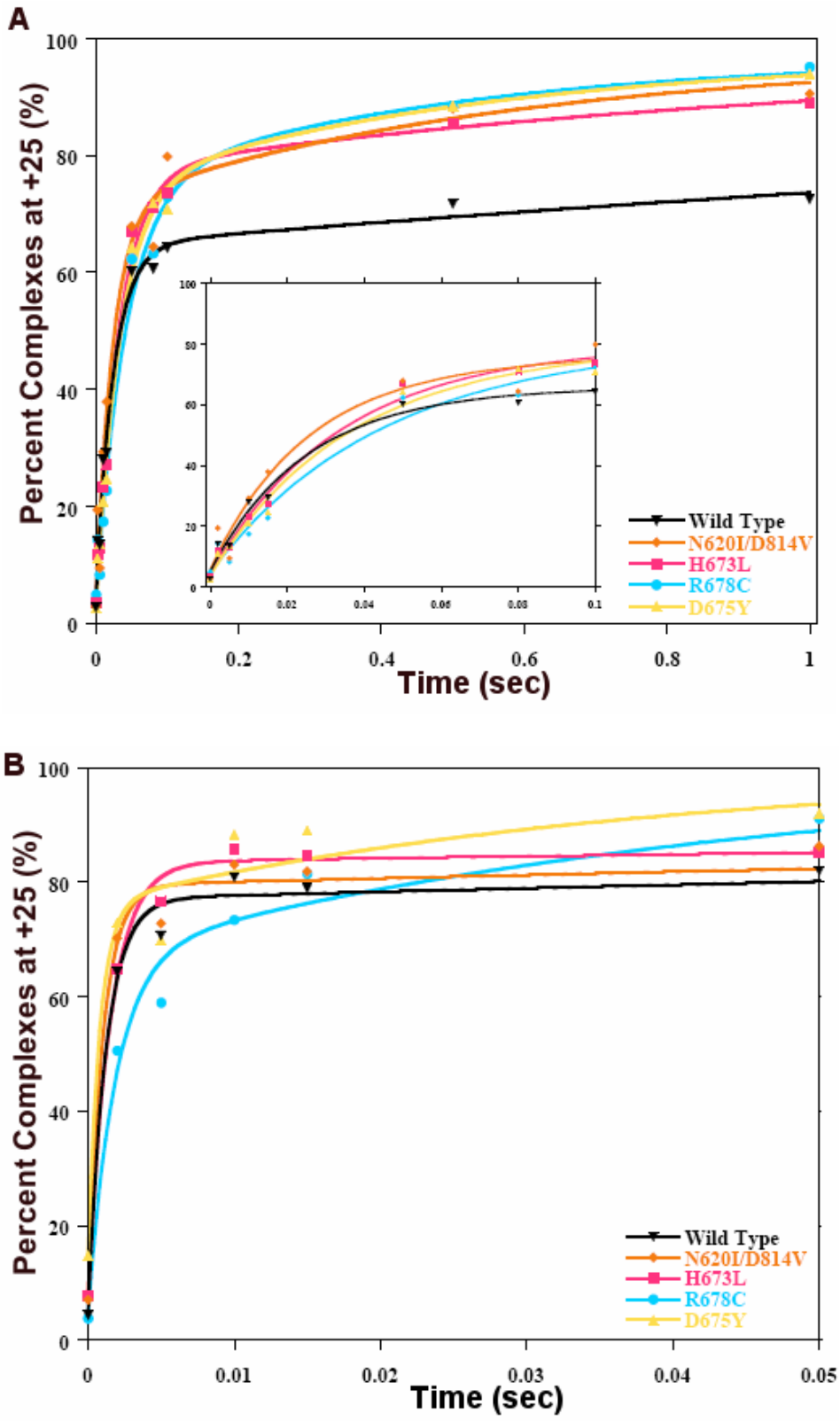

A
Percent Complexes at +25 (%)
Time (sec)
Wild Type
N620I/D814V
H673L
R678C
D675Y
B
Percent Complexes at +25 (%)
Time (sec)
Wild Type
N620I/D814V
H673L
R678C
D675Y

Figure 3.1: ***Plots of correct incorporation of CMP at position +25 versus time by wild type and RNAP variants with amino acid substitutions in the secondary channel.*** The experiment was performed at (A) low concentrations of CTP (5μM) and (B) high concentrations of CTP (100μM). All data are fit to double exponentials. The rate of correct incorporation of CTP at position +25 for wild type and all variant RNAPs was determined to be similar (Holmes *et al.* 2006).

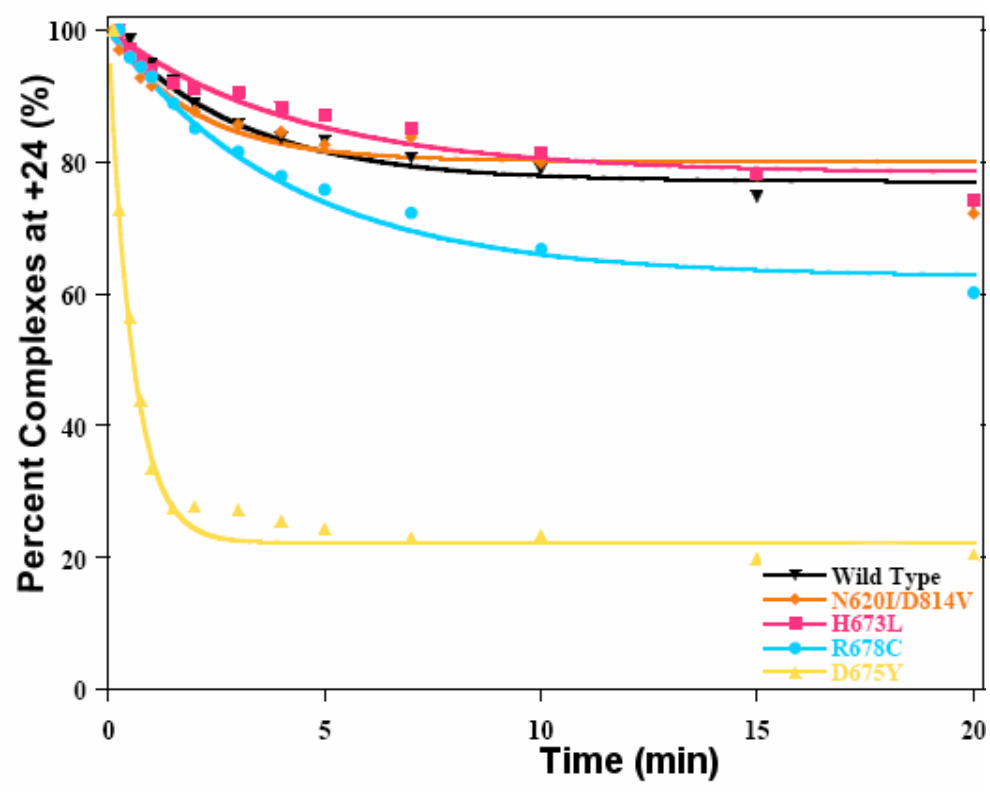


Figure 3.2: ***Plot of the disappearance of complexes at position +24 after incorrect nucleotide incorporation of UMP for CMP by wild type and RNAP variants with amino acid substitutions in the secondary channel.*** βD675Y RNAP shows a significant increase in the amount of misincorporation compared to wild type while R678C shows only a slight increase (Holmes *et al.* 2006).

Misincorporation experiments similar to those quantitatively represented in Figure 3.2 were performed using a variant RNAP where a valine was substituted for the aspartic acid at residue 675 (βD675V). Similar to the tyrosine substitution, the neutral valine side chain eliminates the charge that would be present on the wild type aspartic acid. However, unlike tyrosine, valine is similar in size to the aspartic acid. Experiments with βD675V demonstrate that the mutant behaves similarly to βD675Y (Figure 3.3) and further suggests that the charge on the amino acid side chain at residue βD675 is potentially playing a critical role in the fidelity of *E. coli* RNAP (Holmes *et al.* 2006).

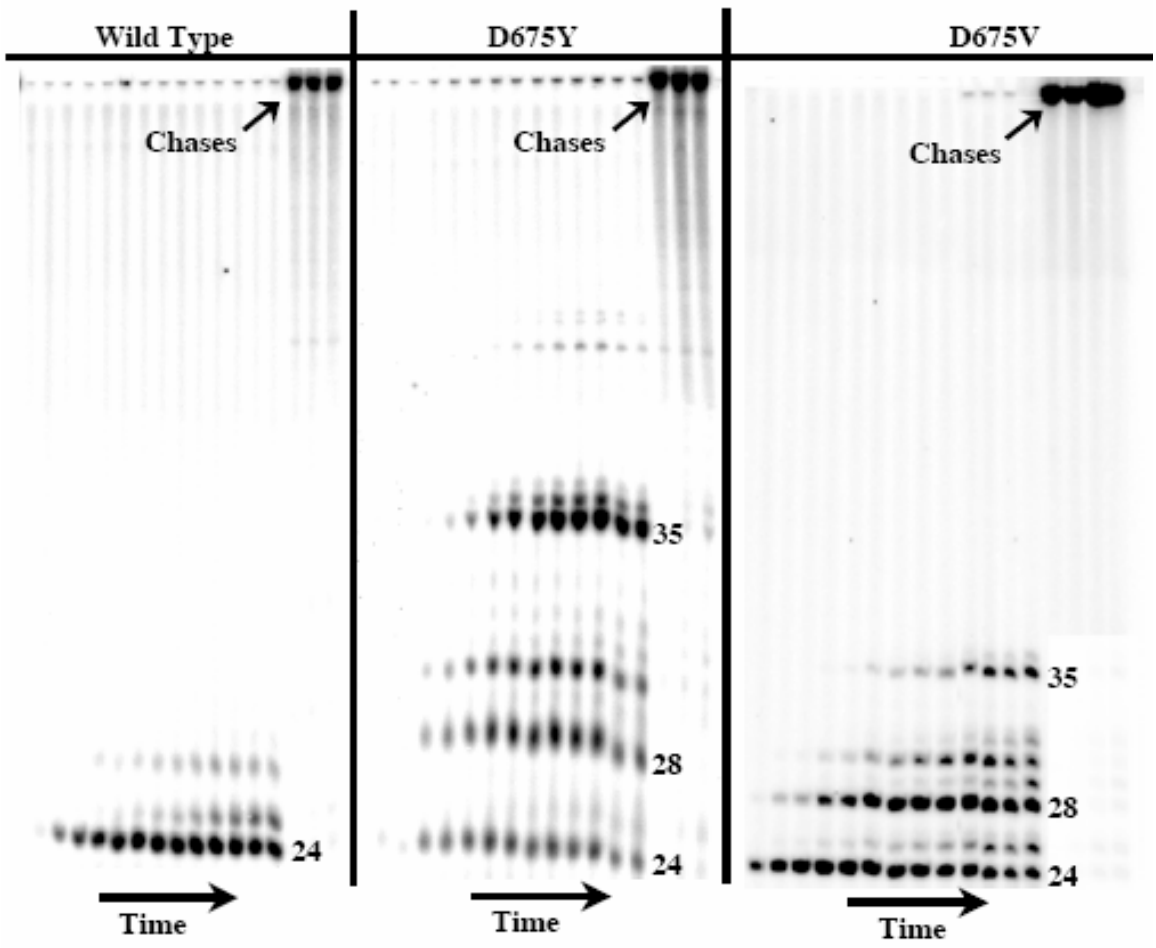


Figure 3.3: ***Misincorporation kinetics of wild type, βD675Y, and βD675V RNAPs.***

Denaturing polyacrylamide gels showing the misincorporation of UMP for CMP at position +25 in the nascent RNA chain over time. D657Y (middle panel) misincorporates at a rate and extent greater than that of wild type (left panel) with similar results shown in D675V (right panel) (Holmes *et al.* 2006).

Crystal structures of the yeast RNAP II elongation complex with an incorrect nucleotide bound and yeast RNAP II elongation complex with an NTP analog bound led to the proposal that there are potentially three sites in the RNAP that are involved in nucleotide binding and incorporation. These sites include an E site (entry site) adjacent to the catalytic site, a PS site (pre-insertion site) where the incoming NTP can pair with the DNA template, and the A site (active catalytic site) (Batada *et al.* 2004; Kettenberger *et al.* 2004; Westover *et al.* 2004). Combining the observation of the three sites with a two-step model for nucleotide incorporation proposed by Westover and co-workers, Temiakov *et al.* suggested a three-step model in which the NTP first binds to the E site and then rotates into the PS site where hydrogen bonding between the NTP and the DNA template base is checked before RNAP closes to bring the DNA template base and NTP pair into the A site (Batada *et al.* 2004; Westover *et al.* 2004; Temiakov *et al.* 2005). From this three-step model taken together with biochemical data, Holmes *et al.* (2006) suggested that mutations in the secondary channel affect either or both the conformational changes associated with moving the NTP from the E site to the PS site and from the PS to the A site. The βD675Y mutant RNAP substitution of tyrosine for an aspartic acid changes the surrounding structure and potentially creates a looser configuration in the tunnel around the small pore that separates the E and PS sites and could explain the observed increase in misincorporation (Holmes *et al.* 2006).

To further characterize the βD675Y RNAP mutant, we have performed UTP concentration-dependent kinetics using purified stalled elongation complexes. Surprisingly, we find that by purifying the complexes, βD675Y no longer misincorporates at rates and extents greater than that of wild type RNAP. In addition, we uncover a zero-order

dependence on the rate of misincorporation for concentrations of UTP less than 75μM. Also, we examined recent crystal structures of RNAP where the trigger loop is proposed to transport the NTP from the PS to the A site, subsequently closing over the catalytic site for synthesis (Toulokhonov *et al.* 2007; Vassylyev *et al.* 2007). We posit that the βD675Y mutation is affecting the closing of the trigger loop over the active site, thereby changing the misincorporation kinetics of the βD675Y mutant RNAP.

Results and Discussion

All experiments used the pDE13 DNA template where the first CMP to be incorporated in the RNA chain is at position +25 (Erie *et al.* 1993). This template is biotinylated at the 5' end and attached to streptavidin-coated magnetic beads. Elongation complexes were formed by adding *E. coli* RNAP, DNA, UTP, ATP, and [α-^{32}P] GTP and stalled at position +24 by omitting CTP. The complexes were then placed next to a magnet and purified by washing with buffer (See Methods). We then monitored the misincorporation of UMP for CMP at position +25 as a function of time. Because misincorporation happens at a rate much slower than that of correct incorporation, reactions were carried out by hand as opposed to using rapid quench techniques (Erie *et al*. 1993).

Concentration-Dependent Kinetics of UMP Incorporation Utilizing βD675Y RNAP

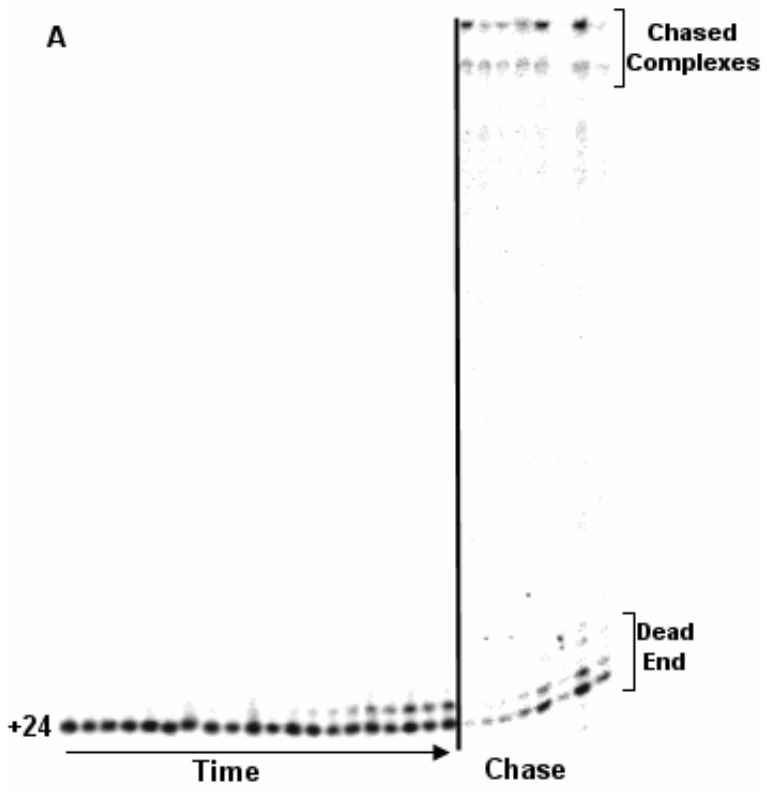
In this work, we have described the misincorporation kinetics for wild type RNAP using purified elongation complexes (Chapter 2). We report that misincorporation can be described via a non-essential activation mechanism where synthesis can only occur in the activated state while a subset of complexes are "trapped" in the unactivated state. These non-

productively bound complexes can be rescued in the presence of the correct NTP or at high concentrations of the incorrect nucleotide. The wild type kinetic data were fit to single exponentials and varying extents of maximum incorporation were reported. We have performed similar concentration-dependent kinetics for the βD675Y RNAP mutant. These experiments differ from the original misincorporation experiments performed with this enzyme in that the reaction is restarted after a stall from purified complexes, as opposed to the running start afforded by the promoter initiated reactions (Holmes *et al*. 2006).

Reactions of βD675Y RNAP elongation complexes with UTP were separated on 20% acrylamide, 8M urea denaturing gels and products of misincorporation of UMP for CMP at position +25 in the nascent RNA chain appearing as a function of time at a low (10μM) and high concentration (600μM) of UTP are shown in Figure 3.4A and 3.4B, respectively. These gels are representative of the range of concentrations of UTP used in these experiments. Similar to wild type RNAP, βD675Y RNAP exhibits a significant increase in the rate and extent of misincorporation at position +25 with increasing concentration of UTP. As seen with the wild type enzyme, the misincorporation reaction for βD675Y RNAP does not go to completion (100%) at lower concentrations of UTP.

The percent of complexes that misincorporate at position +25 were quantified and plotted as a function of time. Figure 3.5 shows kinetic data for misincorporation of UMP for CMP for 9 different UTP concentrations. Inspection of the data in Figure 3.5 reveals that the extents of misincorporation increase with increasing [UTP] as seen in wild type RNAP. However, when plotted individually, we see that the overall rate and extent of misincorporation is less in the βD675Y mutant than in wild type (represented by the plots of elongation complexes with low (10μM) and high concentrations (600μM) of UTP in Figure

3.6A and 3.6B, respectively). This decrease in the rate and extent of misincorporation in the βD675Y RNAP is surprising given the results presented by Holmes *et al.* where at 15μM UTP in the presence of 20μM ATP and 20μM GTP, βD675Y misincorporated up to 80% at a rate that was 20 times faster than that of wild type RNAP which misincorporated up to 20% (Figure 3.2, Figure 3.7, Table 3.1) (Holmes *et al.* 2006). The difference in rate and extent of misincorporation between wild type and βD675Y RNAP changes as concentration of UTP increases, becoming less significant at higher concentrations of UTP (Figure 3.6B).


A
Chased
Complexes
Dead
End
+24
Time
Chase

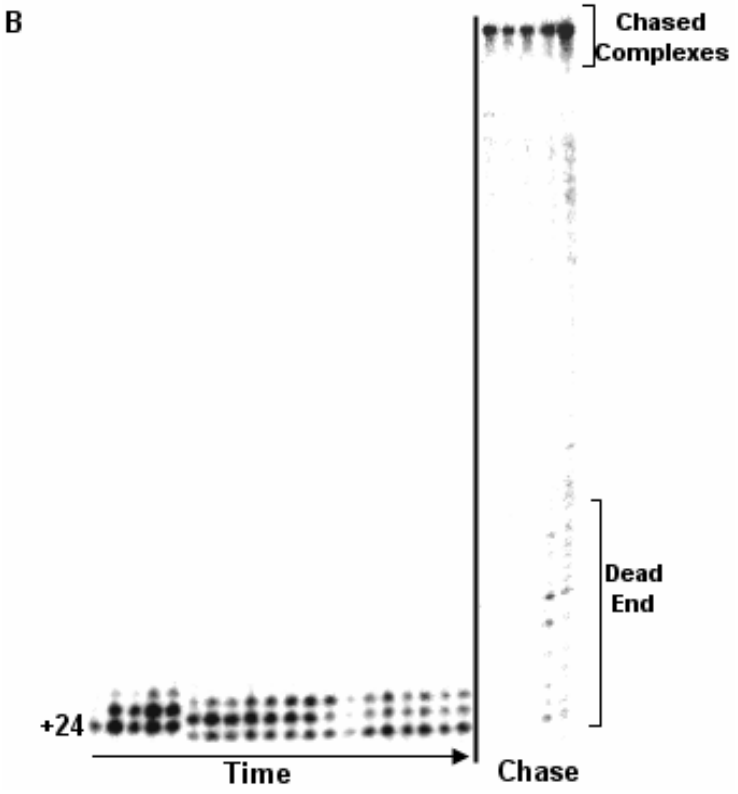

B
Chased
Complexes
Dead
End
+24
Time
Chase

Figure 3.4: ***Representative denaturing gels showing UMP misincorporation by βD675Y RNAP at position +25 at (A) 10μM UTP and (B) 600μM UTP added to purified complexes stalled at position +24 in the nascent RNA chain.*** The rate of misincorporation at position +25 increases with an increase in UTP concentration. Also, the percent of complexes misincorporated at position +25 increases with increasing UTP concentration. Time = 0 (prior to NTP addition), 0.12, 0.24, 0.35, 0.47, 0.59, 0.7, 0.82, 1, 1.17, 1.5, 2, 5, 10, 15, 20, 40, 60, 90, and 120 minutes. Chase reactions = 0, 4, 14, 24, 39, 59, 89, and 119 minutes.

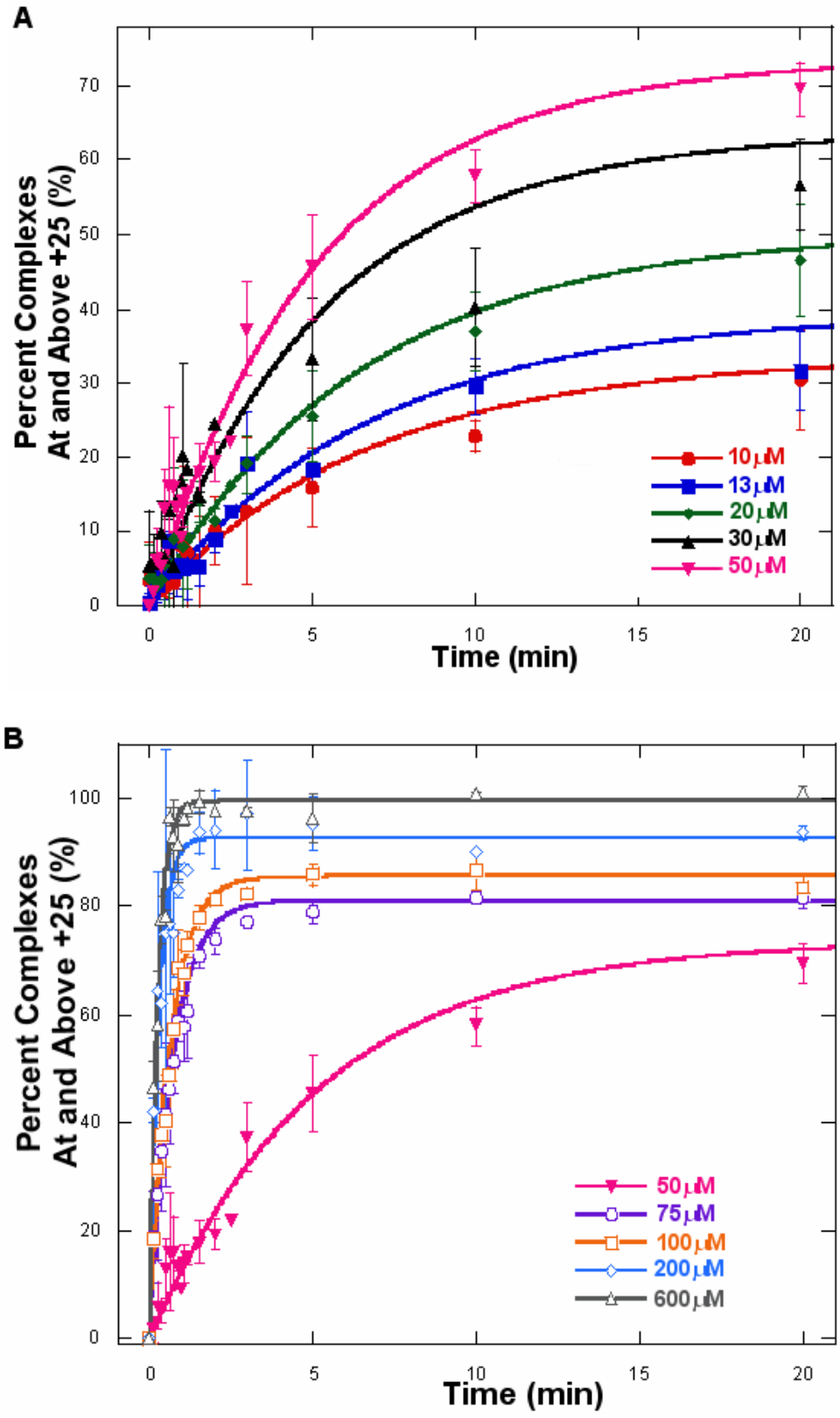
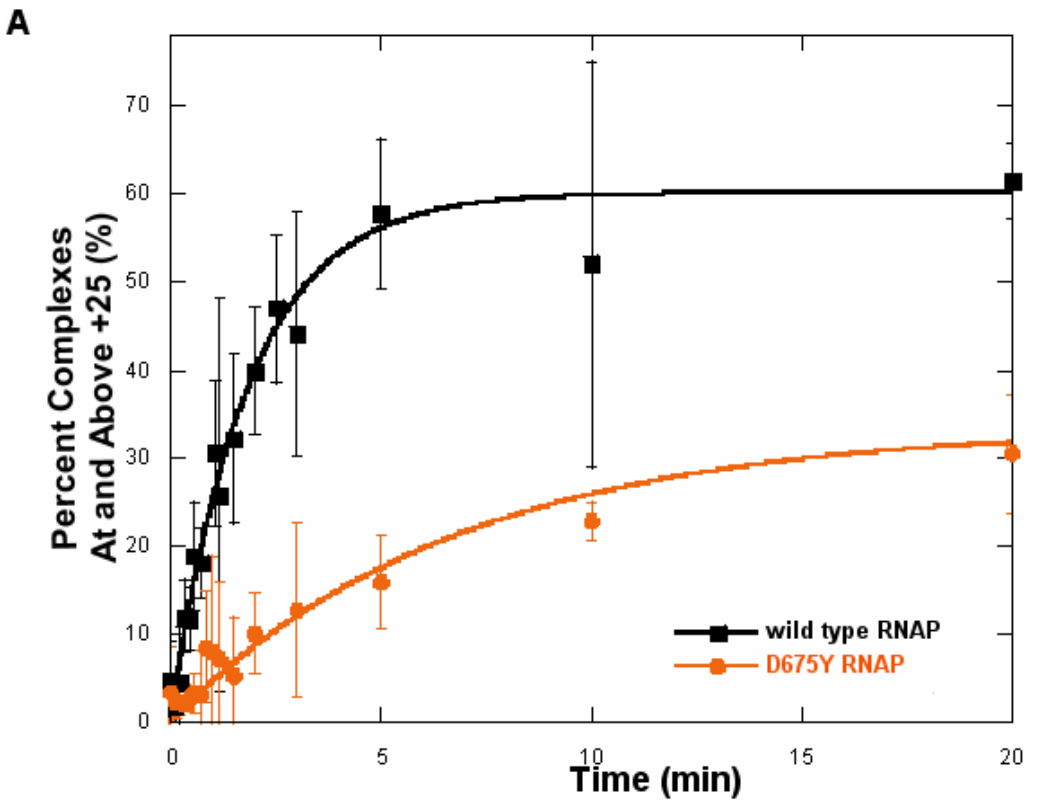

A
Percent Complexes
At and Above +25 (%)
Time (min)
10μM
13μM
20μM
30μM
50μM
B
Percent Complexes
At and Above +25 (%)
Time (min)
50μM
75μM
100μM
200μM
600μM

Figure 3.5: ***Plots of percent misincorporated complexes at position +25 from βD675Y RNAP versus time at (A) 10 – 50μM UTP and (B) 50-600μM UTP.*** These data are fit to single exponentials to obtain the apparent rate constant (k_{app}) and maximum extent of misincorporation (plateau value) for each concentration of UTP. Error bars represent the standard deviation for three to five sets of data for each concentration of UTP.


A
Percent Complexes
At and Above +25 (%)
70
60
50
40
30
20
10
0
0
5
10
15
20
Time (min)
wild type RNAP
D675Y RNAP

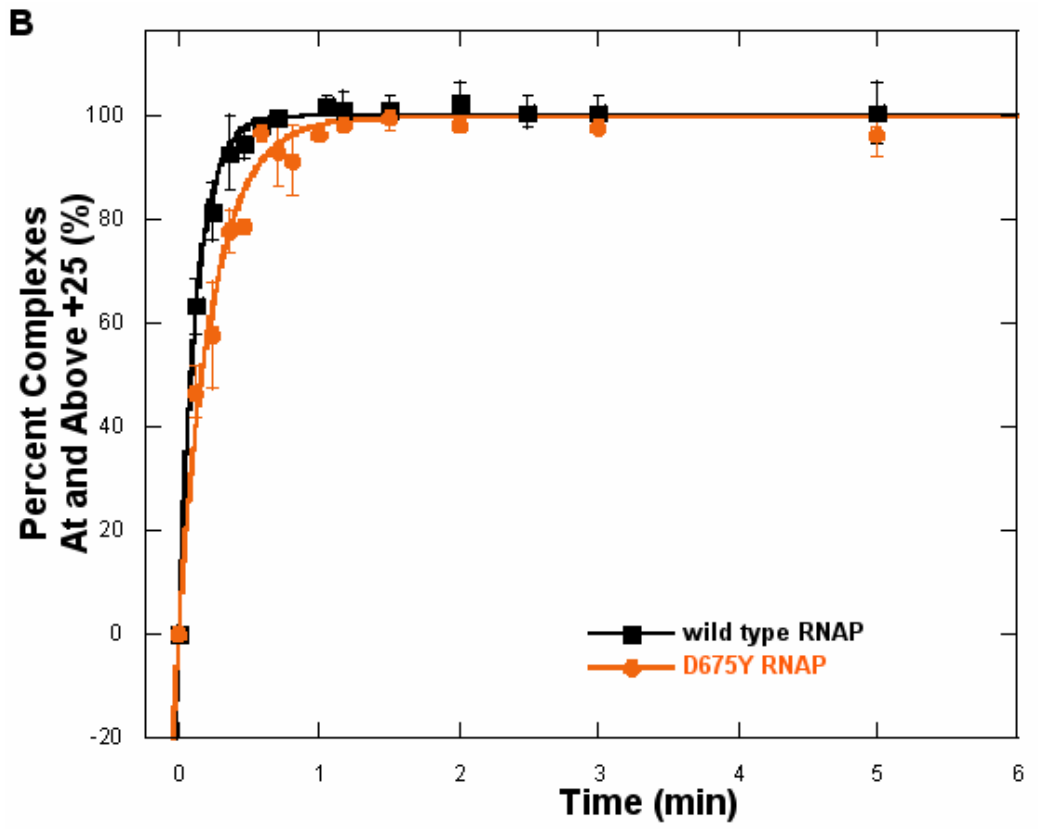

B
Percent Complexes
At and Above +25 (%)
100
80
60
40
20
0
-20
0
1
2
3
4
5
6
Time (min)
wild type RNAP
D675Y RNAP

Figure 3.6: ***Plots of percent misincorporated complexes at position +25 versus time at (A) 10μM UTP and (B) 600μM for wild type RNAP (black squares) and D675Y RNAP (orange circles).*** These data are fit to single exponentials and are representative of the trend in βD675Y to misincorporate slower and to a lesser extent than wild type with the difference becoming less significant with increasing UTP concentration.

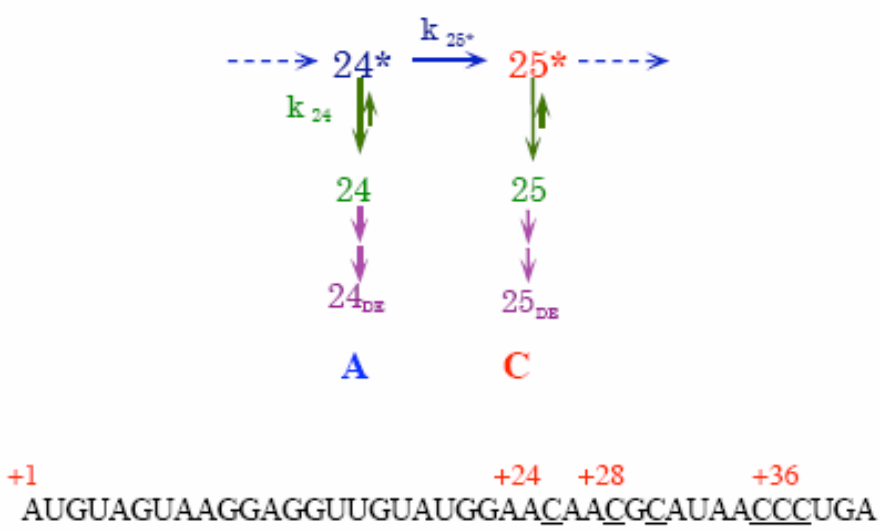


Figure 3.7: ***Basic branched kinetic pathway used to determine the simulated rates for misincorporation by wild type and variant RNAPs shown in Table 3.1*** (Erie *et al.* 1993; Holmes *et al.* 2006). The enzyme in the unactivated state of synthesis at a given template position is represented as n. The enzyme in the activated state of synthesis at a given template position is denoted by n* while n_{DE} represents the enzyme in a dead end state of synthesis.

Polymerase	k_{25*} (min^{-1})	k_{24} (min^{-1})
Wild-type	0.075	0.25
N620I/d814V	0.1	0.4
R678C	0.09	0.15
H673L	0.05	0.18
D675Y	1.4	0.4

Table 3.1: ***Rates of misincorporation simulated using the basic mechanism shown in Figure 3.7.*** Overall variant RNAPs behave similarly to wild type, with the exception of βD675Y. The βD675Y RNAP misincorporates at a rate (k_{25*}) approximately 20 times faster than that of wild type (Holmes *et al.* 2006).

Observing that βD675Y is a lower fidelity variant for reactions initiated from the promoter and a higher fidelity variant for purified elongation complex reactions, we posited that perhaps the presence of the downstream template base in the promoter initiated reactions is responsible for the lower fidelity in the βD675Y RNAP. The presence of ATP, encoded at positions +26 and +27 in the pDE13 template, may somehow enhance the rate of misincorporation in the βD675Y RNAP. To test this hypothesis for the purified elongation complexes, we performed the misincorporation reactions with simultaneous addition of 20μM UTP + 100μM ATP and 75μM UTP + 100μM ATP. These experiments resulted in misincorporation by βD675Y that showed no difference between addition of UTP + ATP and of UTP alone. In the presence of the downstream templated base, the reactions still proceeded at slower rates and to lesser extents than wild type enzyme for the purified elongation complexes. This result suggests that the shift in fidelity of the βD675Y RNAP is not caused by the simultaneous presence of the downstream NTP. The difference in experimental procedure (promoter initiation versus purified elongation complexes) did not significantly change the kinetics of UTP incorporation for CTP at position +25 with wild type enzyme; yet, the difference in experimental procedure does seem to be affecting the behavior of the βD675Y RNAP.

We further examined the misincorporation kinetics of βD675Y by plotting the maximum extents (plateau values obtained from the single exponential fit data in Figure 3.5) as a function of UTP concentration (Figure 3.8). Similar to wild type RNAP, only a subset of the RNA complexes from βD675Y RNAP misincorporates UMP at lower concentrations of UTP. Complete misincorporation (100% extent) is not observed until 600μM UTP, while completion is achieved at approximately 100μM in wild type RNAP. The difference in

extents between βD675Y and wild type RNAP is represented well by Figure 3.8. In the βD675Y RNAP, there is a continual increase in extent of misincorporation with increasing UTP concentration; however, the extent of misincorporation is clearly less than that of wild type enzyme until 600μM UTP. The plot of the extent of misincorporation as a function of UTP was fit to Michaelis-Menten kinetics and a K_m value of 21μM was obtained for βD675Y RNAP (Figure 3.8). This value is 3.5 times higher than the reported K_m of 6μM binding constant for wild type RNAP. The slight decrease in binding affinity does not demonstrate a significant change in the binding of NTPs to the RNAP due to the single amino acid substitution in the secondary channel.

To determine the concentration dependence of the rate of UMP misincorporation in βD675Y, the apparent rate constants (k_{app}) obtained from the single exponential fits were plotted versus UTP concentration (Figure 3.9). The rate of misincorporation presents an interesting phenomenon of the βD675Y RNAP variant. Unlike wild type enzyme, misincorporation by βD675Y RNAP appears to reach saturation at higher concentrations of UTP (Figure 3.9). Upon closer inspection of the rate versus UTP concentration plot, we observe that for concentrations less than 75μM there appears to be a zero-order dependence on the rate of misincorporation for the concentration of UTP (Figure 3.9B). This rate difference is also apparent in the single exponential plots of misincorporation where we see a significant increase in the rate of incorporation after 50μM (Figure 3.4B). Perhaps the zero-order rate dependence would not be surprising if we did not see a continual increase in the extent of misincorporation with increasing UTP concentration. However, as shown in Figure 3.8, the extent of misincorporation by βD675Y RNAP increases with increasing UTP concentration. For all concentrations lower than 75μM, this increase in extent of

misincorporation occurs unexpectedly at the same rate. Synthesis occurring at the same rate for varying UTP concentrations would be expected to extend to the same percentage of misincorporation, given the same time scale of reaction.

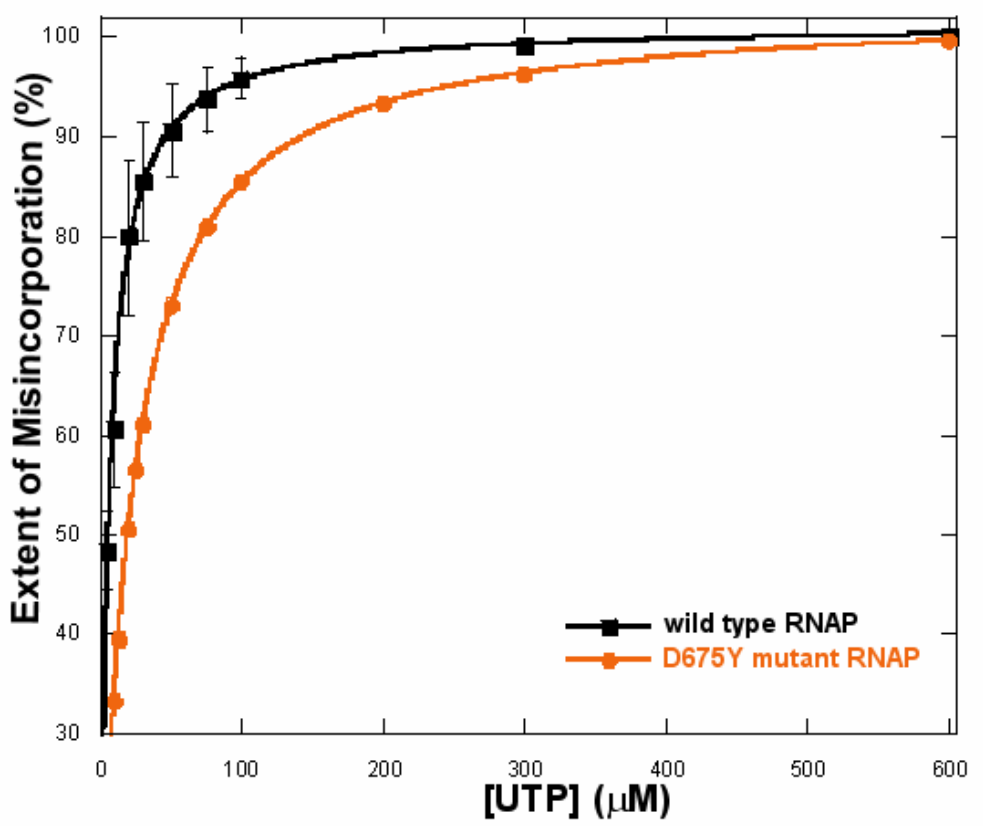


Figure 3.8: ***Plot of maximum extent of misincorporation (%) versus [UTP] (μM) for wild type enzyme (black squares) and βD675Y RNAP (orange circles).*** Data were fit to Michaelis-Menten kinetics (extent misincorporation $= \frac{V_{max} * [UTP]}{K_m + [UTP]}$) with wild type K_m equal to 6μM and βD675Y RNAP K_m equal to 21μM. Error bars represent standard deviation for three to five sets of data for each concentration of UTP.

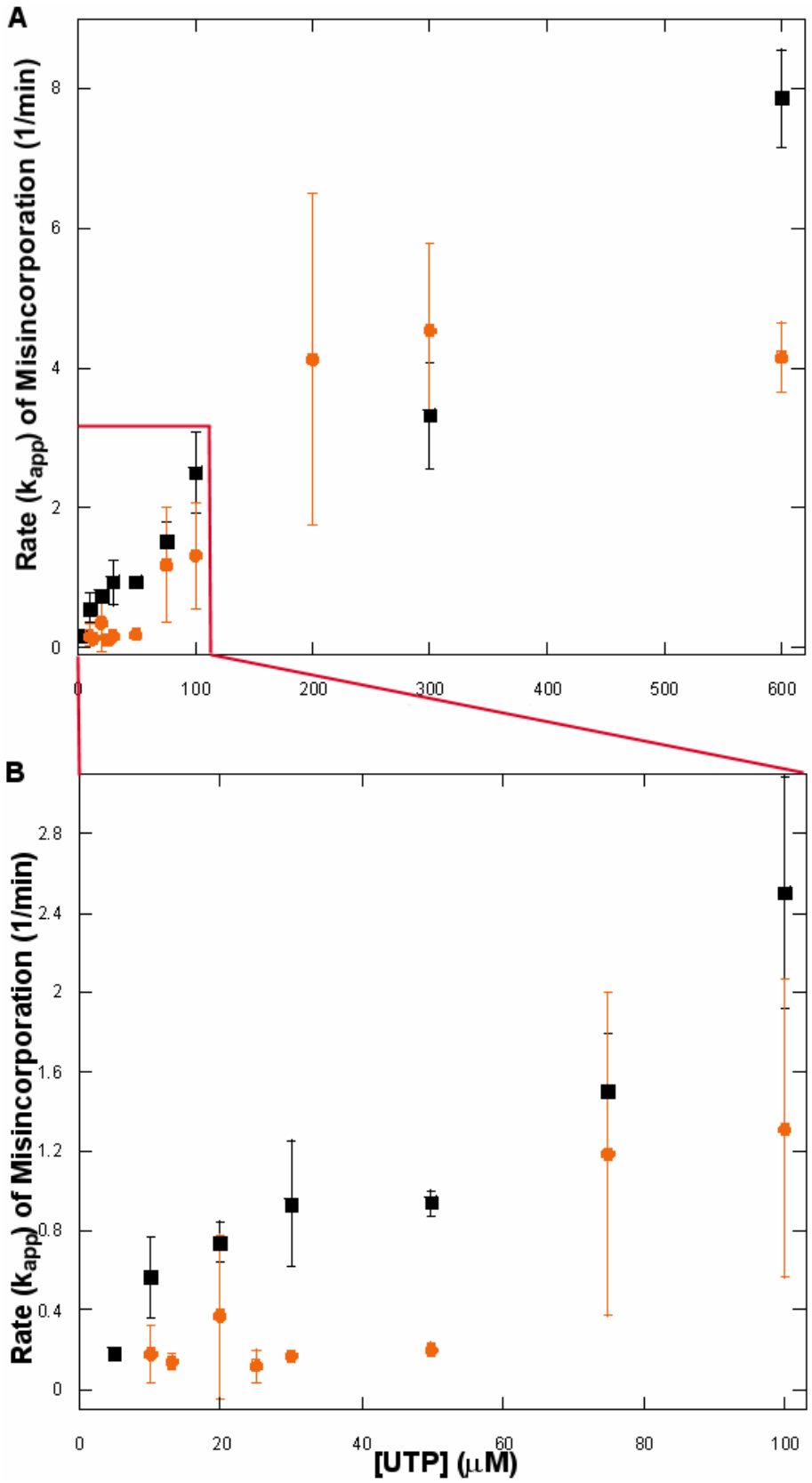

A
B
Rate (k_{app}) of Misincorporation (1/min)
Rate (k_{app}) of Misincorporation (1/min)
[UTP] (μM)
0
2
4
6
8
0
100
200
300
400
500
600
0
0.4
0.8
1.2
1.6
2
2.4
2.8
0
20
40
60
80
100

Figure 3.9: ***Plot of rate*** **(**k_{app}***, min***$^{-1}$***) versus [UTP] (***μ***M) for the wild type enzyme (black squares) and the*** β***D675Y RNAP (orange circles).** The area marked by the red box is expanded (B) to show the zero-order dependence of rate on UTP concentration at [UTP] < 75μM. Error bars represent standard deviation for three to five sets of data for each concentration of UTP.

We have attempted to fit the βD675Y data to the non-essential activation mechanism described previously for wild type kinetics (Chapter 2). However, the increasing extent of misincorporation with a zero-order rate dependence on UTP concentration has presented problems in determining the mechanism by which misincorporation occurs in the βD675Y mutant. We successfully fit concentrations greater than 50μM to a single set of rate constants in the non-essential activation mechanism previously described, with the rate of unactivated synthesis (k_{slow}) equal to zero and the rate out of the unactivated state (k_{-unact}) also equal to zero (Chapter 2, Figure 2.21). Yet, due to lack of rate dependence with continued extent dependence at concentrations of UTP 50μM and lower, we could not fit the lower concentrations to the same mechanism with the same rate constants. We have attempted a mechanism with a separate "trapped" state and found that this mechanism neither fit the high nor the low concentrations of UTP. Several attempts have been made to determine a mechanism, yet none have adequately represented the zero-order rate dependence at lower concentrations of UTP in the βD675Y RNAP.

A possible explanation for the zero-order dependence could be that a third NTP binding site exists in the RNAP and this third binding site is not reaching saturation until concentrations of UTP greater than 50μM. This suggested third NTP binding site is not completely irrational given the crystal structures of yeast RNAP II, which suggest that there are three sites in the RNAP that are involved in nucleotide binding and incorporation (Batada *et al.* 2004; Kettenberger *et al.* 2004; Westover *et al.* 2004). In these crystal structures, NTPs are not shown to occupy more than one of the three E, PS, and A sites simultaneously. However, as proposed by Holmes *et al.* (2006), the mutations in the secondary channel affect either or both the conformational changes associated with moving the NTP from the E site to

the PS site and from the PS to the A site. The βD675Y mutant RNAP substitution of tyrosine for an aspartic acid changes the surrounding structure and potentially creates a looser configuration in the tunnel around the small pore that separates the E and PS sites (Holmes *et al.* 2006). Perhaps this looser configuration allows for an NTP to share occupancy between the E and PS sites, and this shared occupancy should be considered in determining the possible mechanism of misincorporation in the βD675Y RNAP.

Recent Crystal Structures Provide Further Structural Insight into βD675Y RNAP

Crystal structures of the yeast RNAP II elongation complex revealed that binding of an incorrect NTP for synthesis was at a site termed the E site adjacent to the catalytic NTP binding site (A site) with the position of the base in the E site inverted and pointing out into the secondary channel (Westover *et al.* 2004). Comparison of structures with the correct and incorrect NTPs led to the proposal of a two-step model of NTP binding in which an incoming NTP binds first to the E site and then rotates through a narrow negatively charged pore (Batada *et al.* 2004) into the A site where it pairs with the template base (Westover *et al.* 2004).

Cramer and co-workers observed an NTP analog in a third site in which the incoming NTP was base paired with the DNA template base but was not positioned for catalysis and suggested that this site was a pre-insertion site (PS) (Kettenberger *et al.* 2004). They proposed a mechanism similar to the proposed mechanism for T7 RNAP in which the incoming NTP binds to the PS site with the RNAP in the "open" conformation and then RNAP closes down on the correctly paired NTP to align the NTP in the active site for catalysis (Yin & Steitz 2004; Landick 2004; Temiakov *et al.* 2004; Kettenberger *et al.* 2004).

Temiakov *et al*. (2005) integrated these two models and proposed a three-step mechanism in which the NTP first binds to the E site and then rotates into the PS site where hydrogen bonding between the bases is checked before RNAP closes to bring the DNA template base and NTP pair into the A site.

Based on the results of the correct and incorrect nucleotide incorporation previously discussed, Holmes *et al.*(2006) proposed that the amino acid changes at βD675 could affect either or both of the conformational changes associated with moving the NTP from the E site to the PS site and from the PS to the A site. Residue 675 (βAsp^{554} in *T. aquaticus*) is at the beginning of a 3 residue β-turn and one of the side chain oxygens of the aspartic acid is within hydrogen bonding distance to the back bone nitrogen of the first two residues in the turn (Zhang *et al.* 1999). In addition, the other oxygen is positioned to form a hydrogen bond and a salt bridge with the β' residue Gln^{739} and Arg^{744} in *E. coli* RNAP (Gln^{1037} and Arg^{1042} in *T.aquaticus*). This interaction presumably stabilizes the β-turn and anchors it to the side of the secondary channel (Zhang *et al.* 1999). Substitution of aspartic acid with tyrosine or valine would remove the hydrogen bond and salt bridge to β' as well as the hydrogen bonding interactions that stabilize the β-turn thereby changing the surrounding structure and potentially creating a looser configuration in the tunnel around the small pore that separates the E and PS sites. Removal of the charged aspartic acid would also reduce the negative electrostatic potential of this pore (Batada *et al.* 2004; Holmes *et al.* 2006).

Holmes *et al.* (2006) suggests that NTP discrimination is based on NTP rotation through the poor and subsequent rearrangement or closure of RNAP to align the NTP in the A site. Recent crystal structures have shown the trigger loop (β'1221-1265, *T. thermophilus*), a key structural element in the RNAP, exists in a "closed" and "open"

conformation. The “closed” conformation of the trigger loop appears to be closing down over the catalytic NTP in order for synthesis to occur (Vassylyev *et al.* 2007; Kennedy & Erie in preparation). Examining these crystal structures, we find that βD675 (βD554 in *T. thermophilus)* is within 5-6Å of the trigger loop in the closed conformation. Altering residue 675 may affect how the trigger loop closes over the NTP, loosening the restrictions of NTP binding into the catalytic site and allowing misincorporation to occur more rapidly during the reactions initiated from the promoter.

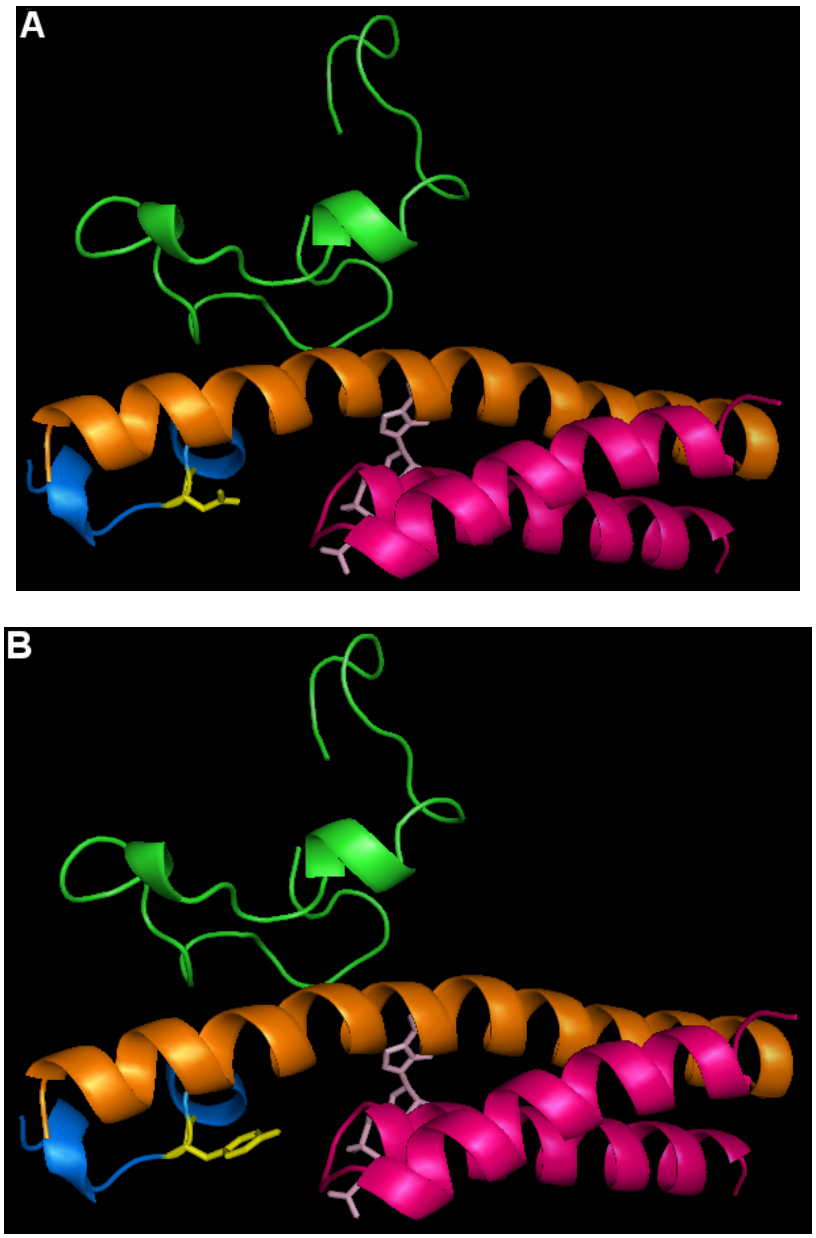


Figure 3.10: ***Potential changes in the trigger loop closing over the catalytic site with amino acid substitutions at βD675.*** (A) βD675 wild type enzyme (B) D675Y variant RNAP with aspartic acid replaced by a tyrosine. The structure is from PDB 2O5J. Fork loop 2 is represented in green. The bound GTP is light purple. The bridge helix is orange while the trigger loop in the "closed" conformation is shown in fuchsia. The β turn where D675 is located is shown in blue, with residue 675 shown in yellow.

Future directions

There are several questions that still need to be answered concerning the βD675Y mutant RNAP. First, what causes the change in misincorporation kinetics for the βD675Y RNAP compared to wild type when the experimental procedure changes from a promoter initiated running start experiment to a purified stalled elongation complex experiment? While simultaneous addition of the downstream NTP appears to have no affect on misincorporation in the purified elongation complexes, perhaps we will see a change in the misincorporation kinetics of βD675Y if we pre-incubate the purified elongation complexes with the downstream templated NTP. We can form purified stalled elongation complexes and then add ATP for a given amount of time prior to any UTP addition and monitor the reaction over time to see if the pre-incubation of ATP affects the rate and extent of misincorporation for the purified complexes made with βD675Y RNAP.

We also seek an explanation for the zero-order rate dependence on the kinetics of misincorporation at concentrations lower than 75μM UTP. We may be able to gain insight into this phenomenon by performing the same concentration-dependent kinetic misincorporation assay with the βD675V RNAP mutant. Will the similar size of the valine without the charge of the aspartic acid yield the same zero-order rate dependence observed in βD675Y or will we see a rate dependence that is similar to that for wild type enzyme where the rate increases approximately linearly with increasing UTP concentration? If so, what does the rate say about the affect D675 has on NTP binding and incorporation? We can perform the same UTP concentration-dependent series as well as the simultaneous and pre-incubation experiments with ATP to test the effect the downstream template base has on the misincorporation of UMP for CMP in βD675V RNAP. Determining the origin of the zero-

order rate dependence will shed light on the possible mechanism for βD675Y misincorporation and should allow us to determine a mechanism that will reasonably simulate the experimental kinetic data.

Based on the observation that misincorporation does not go to completion in the βD675Y RNAP, performing the same chase reactions that were carried out for wild type RNAP to determine if the rescue mechanism is affected by the mutation of residue 675 would be of interest. We proposed previously that the rescue mechanism was facilitated by NTP binding to the allosteric site (fork loop 2) and therefore we would expect no change in the rescue mechanism by changing the residues in the secondary channel. However, any effect βD6757Y has on the rescue mechanism might give insight into the exact structural model for the rescue.

Based on the previous structural model proposed for transcription elongation (Chapter 2), an experiment of interest would be to create a double mutant RNAP where both βD675 is substituted with Y675 and R542-F545 is deleted (Δ-loop mutant). With running start experiments, Δ-loop was shown to be a high fidelity mutant and βD675Y appears to be a low fidelity RNAP. Based on misincorporation reactions with purified complexes, misincorporation occurs at slower rates and to a lesser extent in both Δ-loop and βD675Y RNAPs compared to wild type enzyme. The concentration-dependent misincorporation kinetics utilizing an RNAP with both mutations present may offer greater insight into the role of the secondary channel and main channel elements during transcription elongation.

Acknowledgements

I would like to thank Dr. Tom Santangelo and Dr. Jeffrey Roberts for graciously providing the variant enzymes used in this work.

Experimental Procedures

Sources of protein and DNA

His-tagged wild type RNAP was purified from log phase cells of strain RL916 (gift of R. Landick) as described previously (Burgess & Jendrisak 1975; Uptain & Chamberlain 1997). Expression of D675Y and D675V mutant RNAP was carried out in the *E. coli* strain TOM100 and purified as previously described (Santangelo *et al.* 2003). The DNA template was prepared from pDE13 and amplified by PCR. The biotinylated 540 nucleotide fragment contains the λPR promoter and codes for a transcript in which the first cytosine to be incorporated is at +25 as indicated below:

pppAUGUAGUAAGGAGGUUGUAUGGAA$\overset{+25}{\underline{\mathbf{C}}}$AACGCAUAACCCUGA…

In vitro transcription reactions – misincorporation from promoter initiation

RNAP (60nM) and 5'-biotinylated DNA template (60nM) were incubated for 10 minutes at 37°C in 1X transcription buffer (30mM HEPES (pH 8.0), 10mM Mg^{2+}-glutamate, 200mM K^{+}-glutamate, 25μg/mL BSA, and 1mM DTT) to form open promoter complexes. Transcription was initiated by adding 15μM UTP, 20μM ATP, and 20μM [α-^{32}P] GTP (160Ci/mmol). The reaction was monitored over time at room temperature (~23°C). Reactions were quenched using 100% formamide and products were separated on 20% acrylamide (19:1), 8M urea denaturing gels.

In vitro transcription reactions – purified stalled elongation complexes

RNAP (60nM) and 5'-biotinylated DNA template (60nM) bound to streptavidin-coated magnetic beads were incubated for 10 minutes at 37°C in 1X transcription buffer (30mM HEPES (pH 8.0), 10mM Mg^{2+}-glutamate, 200mM K^{+}-glutamate, 25μg/mL BSA, and 1mM DTT) to form open promoter complexes. Complexes stalled at position +24 were formed by adding 20μM UTP, 20μM ATP, and 20μM [α-^{32}P] GTP (160Ci/mmol) and incubating at room temperature for 25 seconds. The complexes were washed ten to fifteen times using ice-cold 1X transcription buffer by holding the reaction tube next to a strong magnet to retain the complexes. Complexes were resuspended in 1X transcription buffer, aliquoted for the different reactions, and stored on ice. To ensure that the results were not dependent on the time complexes remained on ice, reactions were carried out in a different order with each experiment. Kinetic experiments were carried out by hand at room temperature. Misincorporation reactions were initiated by the addition of the indicated concentration of UTP to the purified SECs. Reactions were quenched using 100% formamide and products were separated on 20% acrylamide (19:1), 8M urea denaturing gels. UTP concentrations reported are the final concentration in solution.

In order to test the effect on misincorporation of pre-incubation and simultaneous addition with the downstream DNA base, purified stalled elongation complexes were formed as described previously. A fraction of these complexes were used for the *in vitro* transcription reaction adding 20μM or 75μM UTP. The remaining fractions were given 100μM ATP + 20μM UTP or 100μM ATP + 75μM UTP simultaneously. Reactions were monitored over time and quenched using 100% formamide before products were separated

on 20% acrylamide (19:1), 8M urea denaturing gels. Concentrations reported are final concentrations in solution.

At designated times during all *in vitro* transcription reactions, a portion of the reaction was added to the presence of all four NTPs (1mM) to extend the transcript to full length and ensure that the complexes were still active.

Data Analysis

Quantification and normalization of rate data

The amount of radioactivity in each lane of the gels was measured on an Amersham Biosciences PhosphorImager and analyzed with ImageQuant software. The percentage of complexes at each position on the template was calculated by dividing the amount of radioactivity in the indicated band by the total amount of radioactivity in all the bands +24 nucleotides in length and longer. To compare data from different experiments, data was normalized such that at time 0, there was 0% incorporation. Due to the incomplete misincorporation reaction, the maximum extent of incorporation could not be normalized to 100%. To normalize these data, the maximum extent of incorporation determined by the single exponential fit to the data was used as the maximum for each concentration. The experiments were conducted three to five times for each concentration.

Fits of the kinetic data to the mechanism

For the wild type and βD675Y RNAP enzyme, each data set was fit to the single-exponential equation using Kaleidagraph v4.01. The data from the single-exponential fits of the individual rate curves were used as a starting point to obtain initial values for binding constants to the catalytic and allosteric sites and the rate constants for the unactivated and activated states as previously described (Foster *et al.* 2001). For the non-essential activation

mechanism and all other mechanisms attempted, KinSim (Anderson *et al.* 1988) was used to fit the data "manually" – meaning the data were simulated using many combinations of rate and binding constants in an attempt to find the best fit possible.

Bibliography

Anderson, K.S., Sikorski, J.A. and Johnson, K.A. (1988). "A tetrahedral intermediate in the EPSP synthase reaction observed by rapid quench kinetics." Biochemistry **27**(19): 7395-406.

Armache, K.J., Mitterweger, S., Meinhart, A. and Cramer, P. (2005). "Structures of complete RNA polymerase II and its subcomplex, Rpb4/7." J.Biol.Chem. **280**(8): 7131-7134.

Artsimovitch, I. and Landick, R. (2000). "Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals." Proc.Natl.Acad.Sci.U.S.A. **97**(13): 7090-7095.

Batada, N.N., Westover, K.D., Bushnell, D.A., Levitt, M. and Kornberg, R.D. (2004). "Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center." Proc.Natl.Acad.Sci.U.S.A. **101**(50): 17361-17364.

Burgess, R.R. and Jendrisak, J.J. (1975). "A procedure for the rapid, large-scale purification of Escherichia coli DNA-dependent RNA polymerase involving Polymin P precipitation and DNA- cellulose chromatography." Biochemistry **14**(21): 4634-4638.

Burton, Z.F., Feig, M., Gong, X.Q., Zhang, C., Nedialkov, Y.A. and Xiong, Y. (2005). "NTP-driven translocation and regulation of downstream template opening by multi-subunit RNA polymerases." Biochem.Cell.Biol. **83**(4): 486-496.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R. and Kornberg, R.D. (2000). "Architecture of RNA polymerase II and implications for the transcription mechanism." Science **288**(5466): 640-649.

Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001). "Structural basis of transcription: RNA polymerase II at 2.8Å resolution." Science **292**(5523): 1863-1876.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C. and von Hippel, P.H. (1993). "Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription." Science **262**(5135): 867-873.

Foster, J.E., Holmes, S.F. and Erie, D.A. (2001). "Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation." Cell **106**(2): 243-252.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A. and Kornberg, R.D. (2001). "Structural basis of transcription: an RNA polymerase II elongation complex at 3.3Å resolution." Science **292**(5523): 1876-1882.

Gong, X.Q., Zhang, C., Feig, M. and Burton, Z.F. (2005). “Dynamic error correction and regulation of downstream bubble opening by human RNA polymerase II.” Mol.Cell. **18**(4): 461-470.

Holmes, S.F., Santangelo, T.J., Cunningham, C.K., Roberts, J.W. and Erie, D.A. (2006). “Kinetic investigation of *Escherichia coli* RNA polymerase mutants that influence nucleotide discrimination and transcription fidelity.” J.Biol.Chem. **281**(27): 18677-18683.

Kennedy, S.K. and Erie, D.A. (2007). “In preparation.”

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). “Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS.” Mol.Cell. **16**(6): 955-965.

Komissarova, N. and Kashlev, M. (1997). "RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA." J.Biol.Chem. **272**(24): 15329-15338.

Komissarova, N. and Kashlev, M. (1997). "Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded." Proc.Natl.Acad.Sci.U.S.A **94**(5): 1755-1760.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A. and Darst, S.A. (2000). "A structural model of transcription elongation." Science **289**(5479): 619-625.

Landick, R. (2004). “Active-site dynamics in RNA polymerases.” Cell **116**(3): 351-353.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H., Yan, H. and Burton, Z.F. (2003). "NTP-driven translocation by human RNA polymerase II." J.Biol.Chem. **278**(20): 18303-18312.

Nudler, E., Mustaev, A., Lukhtanov, E. and Goldfarb, A. (1997). "The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase." Cell **89**(1): 33-41.

Reeder, T.C. and Hawley, D.K. (1996). "Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism." Cell **87**(4): 767-777.

Santangelo, T.J., Mooney, R.A., Landick, R. and Roberts, J.W. (2003). "RNA polymerase mutations that impair conversion to a termination-resistant complex by Q antiterminator proteins." Genes and Development **17**(10): 1281-1292.

Temiakov, D., Patlan, V., Anikin, M., McAllister, W.T., Yokoyama, S. and Vassylyev, D.G. (2004). “Structural basis for substrate selection by T7 RNA polymerase.” Cell **116**(3): 381-391.

Temiakov, D., Zenkin, N., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., Kashkina, E., Savkina, M., Zorov, S., Nikiforov, V., Igarashi, N., Matsugaki, N., Wakatsuki, S., Severinov, K. and Vassylyev, D.G. (2005). “Structural basis of transcription inhibition by antibiotic streptolydigin.” Mol.Cell. **19**(5): 655-666.

Toulme, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M. and Rahmouni, A.R. (2000). "GreA and GreB proteins revive backtracked RNA polymerase *in vivo* by promoting transcript trimming." Embo.J. **19**(24): 6853-9.

Toulokhonov, I., Zhang, J., Palangat, M. and Landick, R. (2007). "A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing." Cell **27**(3):406-19.

Uptain, S.M. and Chamberlin, M.J. (1997). "*Escherichia coli* RNA polymerase terminates transcription efficiently at rho-independent terminators on single-stranded DNA templates." Proc.Natl.Acad.Sci.U.S.A **94**(25): 13548-13553.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I. and Landick, R. (2007). "Structural basis for substrate loading in bacterial RNA polymerase." Nature **448**(7150): 163-168.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S. and Yokoyama, S. (2002). “Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6Å resolution.” Nature **417**(6890): 712-719.

Westover, K.D., Bushnell, D.A. and Kornberg, R.D. (2004). “Structural basis of transcription: separation of RNA from DNA by RNA polymerase II.” Science **303**(5660): 1014-1016.

Westover, K.D., Bushnell, D.A. and Kornberg, R.D. (2004). “Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center.” Cell **119**(4): 481-489.

Yin, Y.W. and Steitz, T.A. (2004). “The structural mechanism of translocation and helicase activity in T7 RNA polymerase.” Cell **116**(3): 393-404.

Zhang, C., Zobeck, K.L. and Burton, Z.F. (2005). “Human RNA polymerase II elongation in slow motion: role of the TFIIF RAP74 alpha1 helix in nucleoside triphosphate-driven translocation.” Mol.Cell.Biol. **25**(9): 3583-3595.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K. and Darst, S.A. (1999). "Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3Å resolution." Cell **98**(6): 811-824.

CHAPTER 4: SUMMARY OF MISINCORPORATION BY *ESCHERICHIA COLI* RNA POLYMERASE DURING TRANSCRIPTION ELONGATION

A detailed kinetic mechanism is crucial to understanding the process of transcription elongation. Correct incorporation studies from our lab have led to a proposed mechanism for nucleotide addition during transcription elongation. This mechanism suggests RNAP can exist in an unactivated (slow synthesis) state or an activated (rapid synthesis) state. Transition between the two states is brought about by conformational changes in the RNAP following templated NTP binding to an allosteric site. Further investigation led to a structural model for translocation, where the movement of the allosteric site upon NTP binding facilitates translocation. This structural model has been expanded to include recently revealed structural information where the trigger loop of RNA polymerase plays a significant role during catalysis.

As described in Chapter 2, we have expanded our knowledge of the mechanism of transcription and the proposed allosteric site by returning the focus of study to misincorporation kinetics. We have proposed a non-essential activation mechanism similar to the proposed mechanism for correct incorporation with several key differences. During misincorporation, synthesis can only occur in the activated state while a subset of complexes are “trapped” in the unactivated state. We propose an incorrect NTP binding first to the catalytic site is interacting with the trigger loop locking the NTP into the catalytic site and blocking escape by occluding the secondary channel. We also demonstrate that “trapped”

complexes can be chased to complete reactions in the presence of the correct NTP.

Performing concentration dependent kinetics with Δ-loop RNAP, where four residues of the proposed allosteric site have been deleted, we determine that the activated state of synthesis is dependent on the fork loop 2 and as such fork loop 2 plays a key role in misincorporation. This information, taken together with the proposed structural model for correct incorporation, we propose an active displacement of NTPs during transcription elongation where a non-productively bound NTP in the catalytic site in the unactivated state of the RNAP can be displaced by the correct NTP. We propose a structural model for this displacement that uses the allosteric site as the site of binding for the correct NTP. NTP binding to the allosteric site shifts the conformation of the protein to allow the incorrect NTP to be released from the catalytic site, escaping through the secondary tunnel. By determining the mechanism and rates of misincorporation, we have expanded our understanding of the process of incorrect nucleotide incorporation during transcription elongation and gained insight into the fidelity of *E. coli* RNAP.

Further characterization of *E. coli* RNAP variant βD675Y, previously described as a lower fidelity mutant for running start reactions, suggests that the βD675Y RNAP is a higher fidelity mutant from purified elongation complexes. This result suggests that in the different experiments, βD675Y exists in two different states which affect the kinetics of misincorporation. In addition to the shift in fidelity, we demonstrate that from purified complexes the mutation of an aspartic acid to a tyrosine at residue 675 affects the kinetics of misincorporation in such a way that there is a zero-order concentration-dependence on the rate of misincorporation for concentrations of UTP less than 75μM. This result suggests that the single amino acid substitution is affecting the dynamics of the RNAP in a way that may

affect the NTP binding entry (E), pre-insertion (PS), and active (A) sites of the enzyme. We posit that the βD675Y mutation is affecting the closing of the trigger loop over the active site, thereby changing the misincorporation kinetics of the βD675Y RNAP. Future work with βD675Y RNAP along with studies of the βD675V and Δ-loop/βD675Y variant RNAP should shed light on the details affecting misincorporation and allow us to get a better picture of what is happening in this enzyme.

In conclusion, investigating the NTP concentration-dependent kinetics of misincorporation with various *E. coli* RNA polymerases has expanded our knowledge of the mechanism of transcription. We have gained insight into several structural elements that affect the fidelity of RNAP while gaining a better picture of the overall structural model of transcription elongation. We have successfully answered several questions regarding misincorporation during transcription elongation while simultaneously leaving more questions to be answered. Future experiments will only serve to answer these questions and leave us with a detailed description of the regulation and overall process of transcription elongation.