This is to certify that the
dissertation entitled

DOWNSTREAM NTP EFFECTS ON HUMAN RNA POLYMERASE II
TRANSCRIPTION ELONGATION
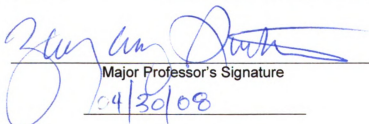
presented by

YALIN XIONG

has been accepted towards fulfillment
of the requirements for the

Ph.D.   degree in   Biochemistry and Molecular
Biology

_____
Major Professor's Signature

04/30/08
_____
Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

DOWNSTREAM NTP EFFECTS ON HUMAN RNA POLYMERASE II TRANSCRIPTION ELONGATION

By

YALIN XIONG

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry and Molecular Biology

2008

ABSTRACT

DOWNSTREAM NTP EFFECTS ON HUMAN RNA POLYMERASE II TRANSCRIPTION ELONGATION

By

YALIN XIONG

RNA polymerase II is the 12-subunit enzyme that synthesizes pre-messenger RNA in human cells. RNA polymerase II is among the largest and most dynamic enzymes in the human biosphere, and it is the target of many regulators. Transcription Factor IIF (TFIIF), TFIIS, and the mushroom toxin α-amanitin are used in this work as dynamic probes of the human RNA polymerase II mechanism, using millisecond phase kinetic analyses and related experimental strategies. The overall aim of this work is to determine the core mechanism of elongation catalyzed by human RNA polymerase II and to understand regulation by transcription factors.

The mushroom toxin α-amanitin is utilized as a transient state inhibitor of human RNA polymerase II. These studies indicate that NTP substrates bind to downstream DNA sites and apparently generate translocation force against a translocation block induced by α-amanitin. In response to the block, RNA polymerase II stalls in the core of the bond synthesis mechanism. It appears that the set point of the stall can be adjusted by changing the concentration of downstream NTPs and the number of downstream sites that can be accurately occupied by NTPs. These experiments show the value of transient state inhibition

experiments to analyze the internal mechanism of elongation catalyzed by human

RNA polymerase II and give insight into the core mechanism of RNA synthesis.

To my family

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my advisor, Dr. Zachary F. Burton, who has the attitude and the substance of a great mentor: he continually and convincingly conveyed a spirit of adventure in regard to research and scholarship. Without his guidance and persistent help this dissertation would not have been possible.

I would like to thank my guidance committee members, Dr. Honggao Yan, Dr. Michael Feig, Dr. Gregg Zeikus and Dr. Jim Geiger for their contributions during these years. I would also like to thank Dr. William Wedemeyer for his kind help on bioinformatics.

I thank all my colleagues in the Burton laboratory for sharing their expertise, reagents, discussions, and laughs with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS AND ABBREVIATIONS

TF	Transcription factor
RNAP	RNA polymerase
TBP	TATA-binding protein
mRNA	Messenger RNA
CTD	Carboxy-terminal domain
NTP	Nucleoside triphosphate
NMP	Nucleoside monophosphate
dNMP	Deoxy nucleoside monophosphate
dNTP	Deoxy nucleoside triphosphate
NDP	Nucleoside diphosphate
EC	Elongation complex
IR	Isomerization reversal
PPi	Pyrophosphate

CHAPTER ONE

THE MECHANISM OF ELONGATION CATALYZED BY

MULTI-SUBUNIT RNA POLYMERASES

1. Transcription catalyzed by RNA polymerases

The regulation of gene expression is fundamental to normal growth, development and survival of an organism. Gene regulation is executed mostly at the level of transcription, the synthesis of RNA from a DNA template. This RNA transcript is often processed during transcription to produce a mature RNA that can itself be functional as a regulatory element, but for most genes is a messenger (mRNA) for translation into protein. Transcription is carried out by the enzyme RNA polymerase (RNAP). Numerous factors regulate transcription by influencing the ability of the RNAP to access, bind and transcribe specific genes in response to appropriate signals. The general transcription factors are involved in various stages including recognition of promoter sequences, the response to regulatory factors and conformational changes essential to the activity of RNAP during the transcription cycle (Conaway and Conaway, 1997; Kuhlman et al., 1999; Ohkuma, 1997; Orphanides et al., 1996; Reinberg et al., 1998; Roeder, 1996; Sawadogo and Sentenac, 1990). Advances made over the past decades have revealed several of the key general transcription factors (Borukhov and Nudler, 2003; Murakami and Darst, 2003; Woychik and Hampsey, 2002), and

most recently, structures and models of RNAP II interacting with general transcription factors (Bushnell et al., 2004; Chen and Hahn, 2003; Chung et al., 2003). Combined with biochemical and genetic studies, these structures provide emerging views on the mechanism of the transcription machinery, the dynamic nature of protein-protein and protein-DNA interactions involved, and the mechanism of transcriptional regulation.

Although the transcription machinery of eukaryotes is much more complex than that of prokaryotes or archaea, the general principles of transcription and its regulation are conserved. Bacteria and archaea have only one RNAP, whereas eukaryotes use three nuclear enzymes, RNAP I-III, to synthesize different classes of RNA. The nuclear RNAPs share five common subunits, with the remainder showing strong similarity among the eukaryotic and archaeal enzymes (Bell and Jackson, 1998; Lee and Young, 2000). Although these enzymes have many more subunits than bacterial RNAP, subunits that make up most of RNAP II are homologous to subunits from all cellular RNAPs, suggesting that all these enzymes have the same or similar basic structure and mechanism (Ebright, 2000). In bacteria, the σ subunit is the sole general transcription factor-like polypeptide. σ recognizes promoter sequences, promotes conformational changes in the RNAP-DNA complex upon initiation and interacts directly with some transcription activators. In eukaryotes, σ factor function has been replaced by a much larger set of transcription factors, with each of the three forms of RNAP having their own set of associated general

transcription factors (Grummt, 2003; Lee and Young, 2000; Schramm and Hernandez, 2002). The RNAP II transcription machinery is the most complex, with a total of nearly 60 proteins, only a few of which are required for transcription by the other nuclear RNAPs. In contrast, archaea use a simplified version of a RNAP II - RNAP III-like system, relying on only two essential general factors, TBP (TATA-binding protein) and TFB (related to the RNAP II and RNAP III general factors TFIIB and Brf1)(Bell and Jackson, 1998).

The three main phases of the transcription cycle are known as initiation, elongation and termination, and each stage is subject to regulation. During transcription initiation, a transcription-competent RNA polymerase complex forms at the promoter and the DNA template is aligned in the active site of the polymerase. The active site is where nucleotides are paired with the template and are incorporated processively during elongation to produce the RNA transcript. Termination of transcription involves release of the RNA transcript and the dissociation of the transcription complex from the DNA template.

Recent studies have challenged the once commonly held view that transcription is predominantly regulated at the level of RNA polymerase recruitment to the promoter and it is now clear that regulation of transcription at the level of elongation is also important. In this study, I focus on RNA polymerase II (RNAP II) transcription elongation, which can be divided into three distinct stages: promoter escape, promoter-proximal pausing, and productive elongation. Each of these stages is defined by a marked difference in the stability and

behavior of the transcription complex, as well as a distinct group of factors that associate with it.

RNAP II is the 12-subunit enzyme that synthesizes all mRNA-encoding genes and some other genes that encode structural or regulatory RNAs. A striking feature that distinguishes RNAP II from the other two eukaryotic RNAPs is the extended carboxy-terminal domain (CTD) of Rpb1, the largest RNAP II subunit. The RNAP II CTD consists of tandem heptapeptide repeats of the consensus sequence, YSPTSPS (Prelich, 2002). The CTD is thought to extend from the core of the polymerase, and is subject to modification during the entire transcription cycle. Modification of the CTD markedly affects its conformation and ability to associate with factors that are involved in transcription elongation, RNA processing and termination. Therefore, modification of the CTD is important for the coordination of transcription events, and different modification states of the CTD are characteristic of different transcriptional stages.

A hallmark of RNAP II during the transition from transcription initiation to transcription elongation is a change from a hypophosphorylated (RNAP IIA) to a hyperphosphorylated (RNAP IIO) form (Sims et al., 2004). Phosphorylation of the CTD occurs predominantly at Ser2 and Ser5 by CTD-modifying enzymes (Meinhart et al., 2005). The level of Ser5 phosphorylation peaks early in the transcription cycle and remains constant or decreases towards the 3' end of the gene (Komarnitsky et al., 2000). By contrast, Ser2 phosphorylation predominates in the body and towards the 3' end of the gene and occurs concomitantly with productive elongation (Komarnitsky et al., 2000).

In eukaryotes, RNAP II transcribes a DNA template that is associated with a complex mixture of proteins, which is known as chromatin. The main constituent of chromatin is the nucleosome, which consists of 147 base pairs of DNA wrapped around an octamer of histone proteins. The mechanics of RNAP II transcription requires marked perturbation of chromatin. Chromatin in transcribed regions of the genome is more decondensed. Recent genomic analysis in *Saccharomyces cerevisiae* have revealed that genes, especially highly transcribed genes, have lower nucleosome occupancy than intergenic regions, with promoter regions showing evident nucleosome depletion (Pokholok et al., 2005). Nevertheless, nucleosomes are usually present in coding regions and pose a barrier to transcription elongation that must be overcome. Several protein factors regulate chromatin structure during transcription by covalently modifying histones or by temporarily moving or disassembling and reassembling nucleosomes (Henikoff and Ahmad, 2005). Several recent studies have revealed important insights into the regulation of chromatin structure during transcription elongation (Henikoff and Ahmad, 2005; Khorasanizadeh, 2004).

2. α-amanitin: the specific and potent inhibitor of RNAP II

α-amanitin selectively and specifically inhibits transcription by RNAPII (Kedinger et al., 1970; Lindell et al., 1970) and RNAP III (Weinmann and Roeder, 1974), RNAP II being the most sensitive. α-Amanitin is a cyclic peptide which binds to the Rpb1 subunit with high affinity (Kd $\sim 10^{-9}$ M) (Cochet-Meilhac and Chambon, 1974; Lutter, 1982). After three decades of study, however, the mode

of this toxin action has not been resolved. Recently, an X-ray structure was published of α-amanitin soaked into crystals of yeast RNAP II, indicating many of the key atomic contacts that contribute to RNAP II inactivation (Bushnell et al., 2002) . Multiple strong interactions between α-amanitin and RNAP II are found in the "funnel," the "cleft," and the key "bridge α-helix" regions of the Rpb1 subunit, which constitute the RNAP II "secondary pore" or "pore 1," a solvent accessible channel into the RNAP II active site. Contacts to the bridge helix are on the opposite face from the RNAP II active site, indicating that α-amanitin is unlikely to block catalysis directly. Bound α-amanitin partially occludes the pore, which has been presumed to be a channel for NTP substrate loading (Bushnell et al., 2002; Cramer et al., 2001; Gnatt et al., 2001) but may be primarily the route of pyrophosphate release (Nedialkov et al., 2003). Contacts of α-amanitin to the bridge helix are intriguing because bending of the bridge α-helix is thought to drive translocation of the RNA-DNA hybrid (Gnatt et al., 2001; Holmes and Erie, 2003; Nedialkov et al., 2003; Vassylyev et al., 2002). Translocation is essential during each bond addition cycle to move the RNA-DNA hybrid away from the space that must be cleared to load the next substrate NTP into the active site. Because α-amanitin binds to RNAP II with a relatively straight conformation of the bridge helix and because of key contacts between the bridge helix and α-amanitin, Kornberg and co-workers suggested that α-amanitin may block translocation by blocking bridge helix bending (Bushnell et al., 2002). Several other reports, however, indicate that RNAP II may be capable of forming multiple phosphodiester bonds without dissociation of α-amanitin (Chafin et al., 1995;

6

Rudd and Luse, 1996), so α-amanitin may allow occasional translocation without dissociation from RNAP II.

3. Translocation models

Mechanisms that describe translocation by multi-subunit RNA polymerases include the power stroke model, the Brownian/thermal ratchet model, the allosteric model, the pre-insertion model, the NTP-driven translocation model, and the "trigger loop-trigger helices" model (Abbondanzieri et al., 2005; Artsimovitch and Vassylyev, 2007; Bar-Nahum et al., 2005; Burton et al., 2005; Epshtein et al., 2002; Foster et al., 2001; Gong et al., 2005; Holmes et al., 2003; Landick, 2004; Oster, 2002; Sousa, 2003; Vassylyev and Artsimovitch, 2005; Xiong and Burton, 2007).

Powerstroke and Brownian ratchet models have been invoked to describe translocation by molecular motors. In a powerstroke mechanism, translocation is tightly coupled to an energy generating reaction, such as NTP hydrolysis. In many cases, however, molecular motors appear to keep separate the power generating step from the translocation step. For RNA and DNA polymerases, separating translocation from NTP hydrolysis may be essential to ensure fidelity of NMP or dNMP incorporation (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007). In a Brownian ratchet mechanism, translocation is rapid and spontaneous, driven by thermal motions. After translocation, thermal ratchet mechanisms require locking of the post-translocation state, i.e. by NTP binding. In general, molecular motors appear, however, to utilize binding of NTP

substrates as allosteric effectors to drive forward translocation. Hydrolysis of the NTP is dissociated from the translocation step, because linking NTP hydrolysis directly to energy generation for translocation might constitute a low fidelity mechanism. In multi-subunit RNA polymerase mechanisms, translocation is directly linked to fidelity through allosteric coupling (Gong et al., 2005; Xiong and Burton, 2007). Translocation may be a more deliberate and accurate process than could be described either by a pure power stroke or thermal ratchet model, lacking an allosteric component. The secondary pore NTP loading model requires a rapidly reversible Brownian ratchet (Figure I-1, **Images in this dissertation are presented in color.**), but dynamic studies of elongation may not be consistent with rapid reversibility of translocation or spontaneous pyrophosphate release (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2005).

The NTP-driven translocation model has features of a power stroke, thermal ratchet, and allosteric model (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007). The energy generating reaction is NTP hydrolysis during phosphodiester bond formation, and NTP hydrolysis is separated in the mechanism from the primary translocation step, to ensure fidelity. In the NTP-driven translocation mechanism, accurate NTP loading is coupled to translocation, limiting the possibility of NTP loading errors. At a stall position, in the absence of the next NTP substrate, the thermal ratchet is revealed, although, during rapid elongation, translocation appears to be NTP-driven. At a stall, once

Figure I-1. The secondary pore NTP loading hypothesis. The bridge α-helix is

blue (viewed from the C-terminal end). The DNA template strand is green.

Phosphates are gold. The RNA strand is red. Consequences of the secondary

pore NTP loading model are indicated.

Figure I-1.

Figure I-1 (continued).

Brownian ratchet

Rapid, reversible

main enzyme channel

bridge helix

secondary pore

"E" site

G
C
i+2
i+1
i
T
A
U

DNA

RNA

Post-translocated

Figure I-1 (contn'd).

main enzyme channel

bridge helix

secondary pore

G

C

i+2

i+1

i

T

A

A

U

R

P

P

DNA

RNA

"A" site
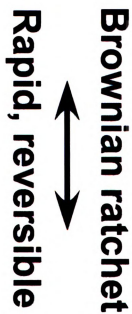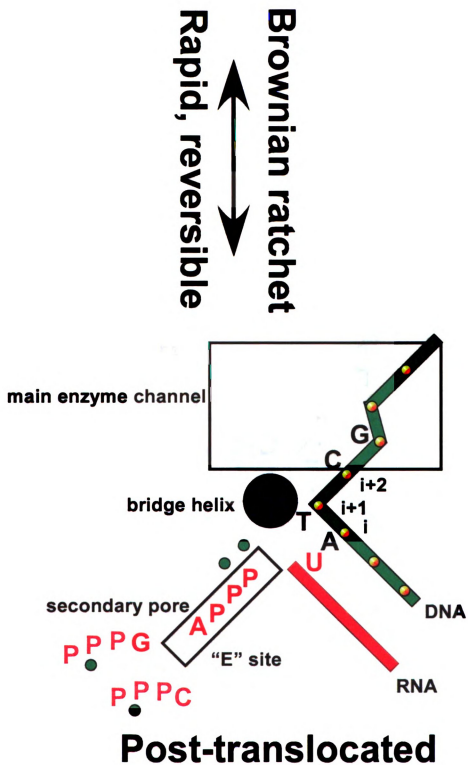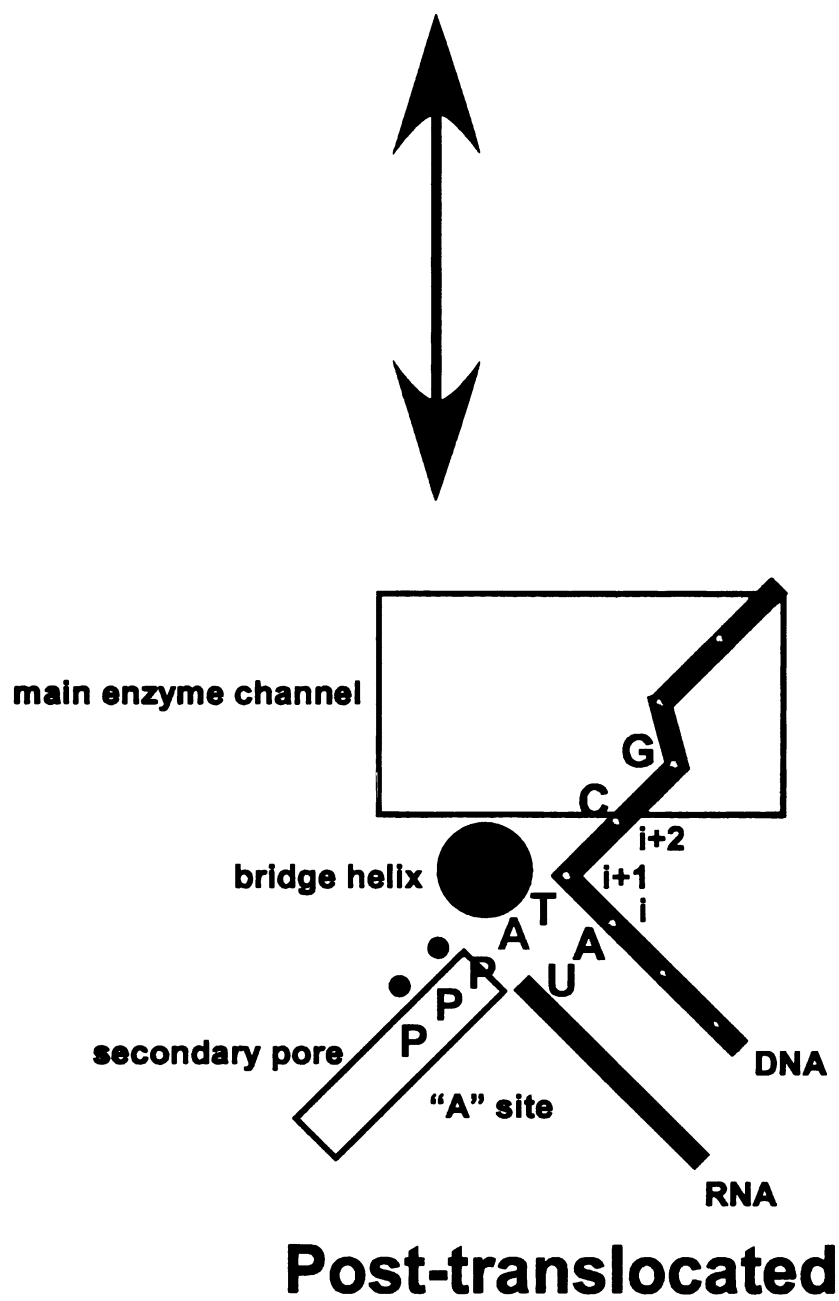
Post-translocated

Figure I-1 (continued).

Secondary pore NTP loading hypothesis
NTPs load at random to i+1 (active site)
NTPs usually mis-load (4 NTPs)
NTPs must load through the secondary pore
NTPs load only to the post-translocated EC
NTPs never load to the pre-translocated EC:
      pore is blocked by RNA
"Entry" site: backwards NTP near active site
NTPs flip from the "E" site to the "A" site
Incorrectly loaded NTPs must be expelled and replaced
      through the pore
Translocation is spontaneous, rapid, reversible
      (not observed)
Competition by non-templated NTPs (not observed)
NTP loading must be rate-limiting (not observed)
Low fidelity mechanism
Fidelity through trigger loop closing?

pyrophosphate has been slowly but spontaneously released, the elongation complex is observed to ratchet between pre- and post-translocated states. Rates of spontaneous translocation, however, appear to be slow compared to NTP-driven translocation (Burton et al., 2005; Gong et al., 2005; Nedialkov et al., 2003; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2003; Zhang et al., 2005). During ongoing synthesis, accurate binding of NTP substrates at the i+2 and i+3 downstream sites drives forward translocation and drives transfer of the i+2 NTP to the i+1 (substrate) position. In this manner, an accurately paired NTP at i+2 becomes an allosteric effector for its transfer into the active site.

Erie has proposed an allosteric mechanism, which may be similar to the NTP-driven translocation model, to describe elongation by *E. coli* RNA polymerase (Foster et al., 2001; Holmes et al., 2003; Holmes et al., 2006). Erie describes an "activated" state and an "unactivated" state for elongation. By considering the activated state to be the post-translocated state and the unactivated state to be the pre-translocated state, Erie's model becomes a version of the NTP-driven translocation model. Otherwise, the Erie model appears to require two templated NTP substrates (allosteric and substrate NTPs) interacting simultaneously at one template position, a model that does not appear to make physical sense. So far, Erie and co-workers have reported kinetic analyses of a single phosphodiester bond using a single quenching technique (EDTA). Part of the difficulty in interpreting Erie kinetic experiments derives from the lack of HCI quench data and data through formation of at least two

14

phosphodiester bonds. Availability of these data will significantly simplify the comparison of elongation catalyzed by eukaryotic and prokaryotic multi-subunit RNA polymerases. For human and yeast RNA polymerase II, the processive transition (translocation linked to pyrophosphate release) is rate-limiting (Zhang and Burton, 2004; Zhang et al., 2005). Monitoring only a single bond, how can one argue that a translocation event has been observed for the E. coli enzyme? What if the E. coli enzyme initiates elongation from a stall position only in the post-translocated state? At high NTP concentrations, E. coli appears to launch the elongation reaction from a single catalytic state (i.e. the post-translocated state) (Foster et al., 2001), in contrast to human RNA polymerase II, which may commence the reaction from both the pre- and post-translocated states (Zhang and Burton, 2004; Zhang et al., 2005). At low to moderate NTP concentrations, E. coli RNA polymerase can demonstrate bi-phasic kinetics, indicating that two distinct elongation modes can be resolved (Foster et al., 2001). Bi-phasic rate curves at lower NTP concentrations indicate the "activated" and "unactivated" pathways for elongation. Whether these faster and slower kinetic modes reflect reversibility between post- and pre-translocation states or relate to a characteristic of the EDTA quenching technique (i.e. chelating an i+2 NTP-Mg^{2+} rather than the i+1 NTP-Mg^{2+}, which may be protected within the active site) is not yet clear.

Cramer, Vassylyev, and Artsimovich have proposed the pre-insertion model to describe elongation (Artsimovitch and Vassylyev, 2006; Artsimovitch et al., 2005; Kettenberger et al., 2004; Vassylyev and Artsimovitch, 2005). This

15

model was initially based on an X-ray crystal structure from the Cramer laboratory in which an NTP analogue assumes a position ("P site" for pre-insertion site) that may be inconsistent with NMP incorporation at i+1 (Kettenberger et al., 2004). Cramer therefore suggests that the NTP is held in a pre-insertion site from which transfer occurs to the insertion A site. From crystallographic studies of the non-homologous bacteriophage T7 RNA polymerase, strong evidence for a pre-insertion model was obtained (Landick, 2004; Tahirov et al., 2002; Yin and Steitz, 2002). For T7 RNA polymerase, the Y helix can shift position to lock the pre-insertion NTP-dNMP base pair into the active site (insertion site) for chemistry.

Newer elongation complex structures from the Vassylyev (Vassylyev et al., 2007; Vassylyev et al., 2007) and Kornberg laboratories (Wang et al., 2006) indicate that conformational dynamics of the β' and Rpb1 trigger loop may account for pre-insertion (also called pre-selection) and insertion site intermediates. The pre-selection intermediate appears to have a relaxed trigger loop structure. The insertion site structure is characterized by a closed conformation in which the trigger loop forms extended helices ("trigger helices") that enclose the NTP substrate within the active site.

The trigger loop-trigger helices model is consistent with secondary pore NTP loading, but does not necessarily require NTP loading by this route. NTPs might load through the secondary pore, first to the E site. The NTP could then flip into the P or PS (pre-insertion or pre-selection site). Tightening of the trigger loop-trigger helices assembly would close and lock the active center for

chemistry. Mis-loading of an NTP would be expected to block closing of the trigger helices, resulting in release and ultimately accurate replacement of the NTP. Loading of the accurately templated substrate NTP would permit formation of the closed A site structure. After NMP addition to the nascent RNA chain, the trigger loop might be expected to relax. Relaxation of the active site would allow pyrophosphate release and release of the thermal ratchet for translocation. It appears that an open trigger loop conformation would be required for NTP loading through the secondary pore. In recent *T. thermophilus* RNA polymerase structures, there does not appear to be sufficient space to load NTP substrates through the main enzyme channel. Also, the downstream transcription bubble is closed, preventing NTP base-pairing to the DNA template at the i+2 and i+3 positions. The trigger loop-trigger helices model has aspects of secondary pore NTP loading, thermal ratchet translocation, and induced fit to ensure transcriptional fidelity. This model indicates that most NTP selection occurs in and in close proximity to the active site.

NTPs may load through the secondary pore to the "entry" site or "E" site before flipping into the pre-insertion site and/or the insertion site (Figure I-1) (Bar-Nahum et al., 2005; Batada et al., 2004; Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004; Zhang et al., 1999). The secondary pore presents a negative electrostatic barrier to NTP entry (Batada et al., 2004). NTPs can load through the secondary pore when the trigger loop-trigger helices assembly is in a relaxed conformation. When the trigger helices are in the closed conformation, there is not very much space to load an NTP through the secondary pore. In

recent *T. thermophilus* RNA polymerase structures (Vassylyev et al., 2007; Vassylyev et al., 2007), there does not appear to be much space to load NTPs to the active site through the main enzyme channel, indicating that, at least for bacteria, NTP loading most likely proceeds through the secondary pore.

Nudler and Ruckenstein have proposed a dual ratchet mechanism to describe translocation (Bar-Nahum et al., 2005). Because the dual ratchet model relies on a forward driving ratchet, which is suggested to be the Rpb1 bridge α-helix driven by a conformational change of the Rpb1 G-loop (trigger loop), this model has elements of power stroke, ratchet, and allosteric changes. Experimental support for the dual ratchet model was primarily derived from steady state experiments, so dynamic effects have been inferred from steady-state approaches, including cross-linking experiments. Kinetic modeling was based on data from the Erie laboratory for wild type RNA polymerase, and comparative data were not reported for β' I1134V and G1136S, the Nudler slow and fast elongation mutants, which locate to the trigger loop (G-loop) region. The dual ratchet mechanism relies on secondary pore NTP loading, a thermal ratchet to support translocation, and dynamics of the trigger loop and bridge a-helix in translocation and active site tightening.

4. The NTP-driven translocation model

4.1. The secondary pore NTP loading hypothesis

The secondary pore NTP loading hypothesis and its predictions are summarized in Figure I-1. According to this model, NTPs load to the RNA

polymerase II active site only through the secondary pore, never through the main enzyme channel. Based on available structures, secondary pore NTP entry appears to require loading to the post-translocated elongation complex. In the pre-translocated elongation complex, the template DNA base faces the main enzyme channel, not the secondary pore, so NTP loading to the pre-translocated elongation complex appears impossible when NTPs load through the pore. Single molecule elongation studies indicate that NTPs may load to a pre-translocated elongation complex (Abbondanzieri et al., 2005), potentially contradicting the secondary pore NTP entry model.

Because NTPs are posited to load directly to the i+1 active site, and NTPs are selected at random from solution, mis-loading must be more frequent than accurate NTP loading. Significant competition by non-templated NTPs is expected, unless each selection of NTPs involves many NTP encounters. Another consideration is that the pore is a 15 angstrom deep and 7 angstrom wide channel to the active site from the "funnel" on the outside of the RNA polymerase II structure (Batada et al., 2004). This channel is only wide enough to admit a single NTP at one time (NTPs have a 6 angstrom minimal diameter). When an incorrect NTP is loaded (the usual case according to the secondary pore NTP entry model), the NTP must be rejected at the active site and then released through the pore, prior to loading another incorrect or correct NTP. Bacteriophage T7 RNA polymerase, which is not a homologue of multi-subunit RNA polymerases, appears to load NTPs through a deep and narrow secondary pore.

Using X-ray crystallography, NTPs have been visualized in different positions near the RNA polymerase II active site: 1) the "E" site (for "entry" site); and 2) the "A" site for "acceptor" site or insertion site (Kettenberger et al., 2004; Wang et al., 2006; Westover et al., 2004). In this review, we consider a "P" site (for "pre-insertion" site) structure to be a variant of the "A" site structure (Kettenberger et al., 2004). Closed trigger loop-trigger helices conformations are probably the best approximations to "A" site structures. Relaxed trigger loop conformations characterize "P" site structures.

In E site structures, base-pairing functions of the NTP base face away from the DNA template strand, so base-pairing to template is impossible in the E site (Wang et al., 2006; Westover et al., 2004) (Figure I-1). An NTP in the E site, therefore, cannot recognize whether or not it is accurately templated. The E site structure has been obtained: 1) using NTPs that are not accurately templated at i+1; or 2) using the templated NTP in the presence of elevated Mg^{2+} (80 mM). The E site may represent an "exit" site in addition to an "entry" site (Wang et al., 2006; Westover et al., 2004). Closed trigger loop-trigger helices structures have an accurately paired NTP (or analogue) near the active site.

The secondary pore NTP loading hypothesis indicates that non-templated NTPs compete for binding to the E and P sites. Such competition is not observed (Gong et al., 2005; Xiong and Burton, 2007), indicating that multiple NTPs can be tested at the active site in each NTP loading through the pore. NTPs that are accurately templated at downstream sites, but not at the i+1 site, stimulate transcription upon i+1 NTP addition, rather than compete for NTP entry

20

(Gong et al., 2005; Xiong and Burton, 2007). This result does not appear consistent with the secondary pore NTP loading hypothesis. According to the secondary pore NTP loading hypothesis, translocation must be rapidly reversible and pyrophosphate release must be rapid and spontaneous. Kinetic studies, however, indicate that translocation coupled to pyrophosphate release may be a rate-limiting step during RNA synthesis (Zhang and Burton, 2004; Zhang et al., 2005). Because of secondary pore dimensions and electrostatics, the secondary pore NTP loading hypothesis appears to require that the rate of NTP loading be rate-limiting (Batada et al., 2004). The rate for NTP loading to the i+1 active site of human RNA polymerase II, however, is at least 1450 +/- 330 s^{-1}, so NTP loading is not rate-limiting (Zhang and Burton, 2004). The rate limiting steps in RNA synthesis are approximately 20-30 s^{-1}. Dynamic studies, therefore, appear inconsistent with the secondary pore being the sole route for NTP loading.

4.2. The NTP-driven translocation hypothesis

A simpler model to describe NTP loading and translocation has recently been hypothesized (Figure I-2). The NTP-driven translocation model posits that NTPs interact at downstream positions i+2, i+3, and possibly i+4. NTP substrates are sorted at a downstream site, i.e. i+3 or possibly i+4, so little competition for NTP loading is observed at the i+1 active site (Burton et al., 2005; Gong et al., 2005; Nedialkov et al., 2003; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2005). NTPs enter the i+1 active site by transfer during translocation paired with their cognate dNMP base. Pyrophosphate

Figure I-2. The NTP-driven translocation model. Colors are as in **Figure I-1**.

Consequences of the NTP-driven translocation model are indicated.
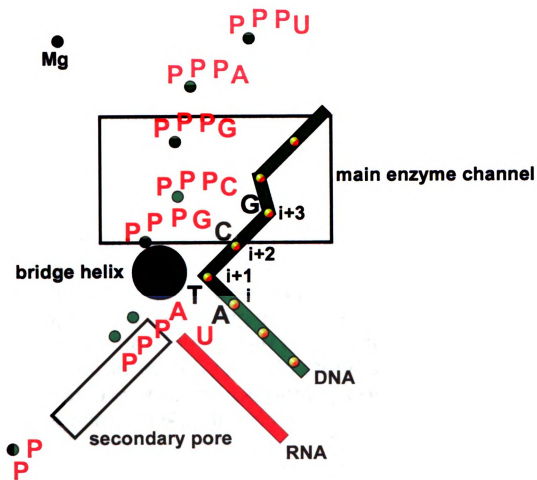
Post-translocated

Figure I-2.

Figure I-2 (continued).

NTP-driven translocation model
Little competition by non-templated NTPs
NTPs at downstream sites (confirmed)
Templated pre-loading and pre-sorting of NTPs
Translocation is coupled to PPi release
      (supported by kinetic studies)
PPi release through secondary pore
Main channel NTP loading
Translocation coupled to fidelity (confirmed)
Transcription errors sensed as translocation blocks
      (confirmed)
High fidelity mechanism

release is coupled to NTP-driven translocation (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007; Zhang et al., 2005). Pyrophosphate is eliminated through the secondary pore, consistent with the constrictive dimensions and repulsive electrostatics of the pore (Batada et al., 2004). Flow of NTPs into the enzyme and excretion of pyrophosphate through the pore are maintained in a consistent direction.

If the i+2 NTP can be shown to affect incorporation of the NTP at i+1, this supports the NTP-driven translocation model and appears less consistent with the secondary pore NTP loading model. Gong et al. (2005) have indicated that, in the presence of the α-amanitin translocation block, i+2 and i+3 NTPs appear to force rejection of a large fraction of stably loaded i+1 NTP (Gong et al., 2005; Xiong and Burton, 2007). NTPs that are not templated at the i+2 and i+3 positions do not induce i+1 NTP rejection, demonstrating base-pairing specificity at i+2 and i+3 sites. Furthermore, dNTPs will not substitute for NTPs at i+2 and i+3 downstream positions, demonstrating chemical selectivity at downstream sites (Gong et al., 2005; Xiong and Burton, 2007). These experiments appear to support the NTP-driven translocation model.

NTP-driven translocation was initially hypothesized based on dynamic studies of RNA polymerase II elongation (Nedialkov et al., 2003; Zhang and Burton, 2004; Zhang et al., 2005). Using the method of rapid chemical quench flow with millisecond reaction timing, the Burton laboratory has observed rapid elongation by multi-subunit RNA polymerases through formation of multiple bonds. Using a combination of reaction quenching (stopping) techniques, stable

NTP loading (EDTA quench) can be tracked independently from phosphodiester bond formation (HCl quench) (Arnold and Cameron, 2004; Arnold et al., 2004; Gong et al., 2004; Gong et al., 2005; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2005). NTP-Mg^{2+} loading to the active site can be tracked by EDTA quenching because free NTPs are disabled by Mg^{2+} chelation. Phosphodiester bond formation is tracked by HCl quenching, which stops the reaction instantly. Interestingly, EDTA quench and HCl quench curves are well separated, showing that NTP-Mg^{2+} can be stably sequestered in the active site prior to formation of a new bond (Zhang and Burton, 2004; Zhang et al., 2005). After formation of the bond, another delay is observed before the next NTP-Mg^{2+} substrate is stably loaded into the active site and becomes resistant to EDTA quenching. These experiments clearly reveal two rate-determining steps in each bond addition cycle during ongoing RNA synthesis. There is a rate-limiting step between stable NTP-Mg^{2+} loading in the active site and phosphodiester bond formation. It is likely that this delay represents an "induced fit" step in which the NTP-Mg^{2+} is tightened prior to incorporation of NMP into the growing RNA chain. Substrate tightening is likely to represent a fidelity check to ensure accurate NMP incorporation, because a mis-paired NTP would not support tightening. The second rate-limiting step during elongation occurs after phosphodiester bond formation but before the next NTP-Mg^{2+} is stably loaded (Zhang and Burton, 2004; Zhang et al., 2005). Based on the RNA polymerase mechanism, this step must be translocation, pyrophosphate release, a reverse isomerization (i.e. relaxation of the trigger loop) or NTP loading. This step is unlikely to be NTP

loading, because this step can be very rapid (>1450 s-1). The secondary pore NTP loading model demands that translocation be rapidly reversible and that pyrophosphate release be rapid and spontaneous. Neither of these requirements of the secondary pore NTP loading model is supported by RNA polymerase II elongation kinetics (Zhang and Burton, 2004; Zhang et al., 2005). The secondary pore NTP-loading model appears to require that NTP loading be rate-limiting (Batada et al., 2004). Stable NTP-Mg^{2+} loading occurs with a rate constant of at least 1450 +/- 330 s^{-1} (Zhang and Burton, 2004), at least 50 times faster than can easily be accommodated by the secondary pore NTP loading hypothesis (Batada et al., 2004).

NTP-driven translocation is expected to be a high fidelity mechanism. The secondary pore NTP loading mechanism, by contrast, may be a low fidelity mechanism. According to NTP-driven translocation, incorrect NTPs are rarely loaded into the active site because of NTP pre-screening at downstream sites. dNTPs are excluded at downstream sites (Gong et al., 2005; Xiong and Burton, 2007), ensuring that multi-subunit RNA polymerases synthesize RNA and not DNA. Furthermore, NTP-driven translocation appears to be coupled to error prevention and error correction mechanisms. Experiments with α-amanitin indicate that transcription errors are sensed as translocation blocks and that NTP-driven translocation against a translocation block can lead to dynamic error prevention and error correction (Gong et al., 2005; Xiong and Burton, 2007).

In "isomerization reversal" experiments (Gong et al., 2005; Xiong and Burton, 2007), using an α-amanitin block and an EDTA quench procedure,

translocation driven by downstream NTPs induces reversal of stable loading of the i+1 NTP-Mg^{2+}. If downstream i+2 and i+3 NTP-Mg^{2+} are disabled at an early time (i.e. 0.002 s) by EDTA chelation, little i+1 NTP-Mg^{2+} is released from the active site, resulting in robust bond synthesis. If downstream NTP-Mg^{2+} is allowed to strain against the α-amanitin translocation block for a longer period of time (i.e. 0.1 s), a significant percentage of i+1 NTP-Mg^{2+} is released (Gong et al., 2005; Xiong and Burton, 2007). Isomerization reversal experiments show that incoming NTPs are used to sense a translocation block, which normally would signal a transcription error. Incoming NTPs help to prevent or correct the error by dislodging the i+1 NTP from the active site. The secondary pore NTP loading hypothesis appears to be a low fidelity mechanism because it requires frequent mis-loading of NTPs to the i+1 active site (the usual case according to the secondary pore NTP loading model). The secondary pore NTP loading hypothesis cannot allow for coupling of translocation to transcriptional fidelity, as in the NTP-driven translocation model (Figures I-1 and 2).

5. NTP loading into the RNA polymerase II active site through main enzyme channel

The active site of a multi-subunit RNAP is deeply buried within the structure. Whether NTPs load through the main enzyme channel or the secondary pore, some explanation of how NTPs navigate the constrained space to the catalytic center will ultimately be necessary.

The NTP-driven translocation hypothesis posits that RNAP II may load at least three NTP substrates simultaneously to templated sites (Burton et al., 2005; Gong et al., 2005; Nedialkov et al., 2003; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2005). Substrates loaded to the active site as NTP-dNMP base pairs, and translocation is driven by protein conformational changes induced by accurate NTP loading. In support of the model, Burton and colleagues have demonstrated sequence-specific NTP interactions with template at i+2 and i+3 downstream sites while the i+1 active site is occupied (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007).

If the NTP-driven translocation hypothesis is correct, then NTP substrates have to be first loaded into the RNAP II main enzyme channel and not the secondary pore (pore 1). NTPs have been suggested to be loaded to the i+1 active site and the i+2, i+3, and, possibly, the i+4 downstream sites (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007). Binding strengths of NTPs are likely to be tuned so that the i+1 site has the highest affinity for an accurately paired NTP, followed by the i+2 site and then the i+3 site. During normal elongation, in the presence of all four NTP substrates, NTPs are freely exchanged at a downstream site, i.e. i+3 or beyond. Downstream template access may be regulated by the non-template DNA strand, because closing of the downstream bubble would disrupt pairing of downstream NTPs. At i+1, i+2, and i+3 sites, improperly paired NTPs are exchanged more rapidly than accurately paired NTPs. Altering the order of affinities of these three sites (i.e. through mutation or by mis-pairing the downstream DNA template) might reduce

transcriptional fidelity, by forcing misloading of NTP substrates to the active site. For instance, enhancing NTP affinity at the i+3 site through mutation would be expected to lower discrimination between accurately and inaccurately paired NTPs at the i+2 and i+1 sites, increasing misincorporation at i+1. This is a possible interpretation of the yeast Rpb1 E1103G mutant RNAP II, which is a fast mutant with defects in transcriptional fidelity. According to these views, NTP-driven translocation is directly coupled to transcriptional fidelity mechanisms.

The NTP-driven translocation mechanism provides an explanation for evolution of the vast, enclosed main enzyme channel, which is a singular characteristic of multi-subunit RNAPs. According to this view, a major consequence of the RNAP "crab claw" structure is to enhance transcriptional fidelity by supporting and tuning NTP-driven translocation. During processive synthesis, the enclosed main channel space is occupied by at least two templated NTP substrates, allowing verification and re-affirmation of NTP identity prior to active site loading. Thus, multi-subunit RNAPs have evolved for high fidelity and transcriptional efficiency. Multi-subunit RNAPs with different gene specificities (i.e. RNAPs I and III) or those derived from different organisms are evolved to adjust error and elongation rates. Extensive conservation of multi-subunit RNAPs through three kingdoms attests to the enduring adaptability of this enzyme. Another reason for the enclosed channel is to maintain high processivity during elongation, by maintaining tight DNA and RNA contacts.

In some yeast RNAP II elongation complex structures, the downstream transcription bubble appears to be open to position i+4 (Gnatt et al., 2001;

30

Westover et al., 2004) or i+6 (Wang et al., 2006), consistent with NTP pairing to the i+1, i+2 and i+3 DNA template sites (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007). After NMP incorporation at i+1, the NTP in the i+2 position is transferred to the i+1 site during translocation. Movement of the i+3 NTP to the i+2 position vacates the i+3 site, which is filled by random exchange of incoming NTPs. In support of this model, the Burton laboratory has demonstrated that NTPs templated at the i+2 and i+3 sites can determine the fate of the NTP loaded at the i+1 position, and incoming NTPs stimulate forward translocation and bond completion, which is pyrophosphate release (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007; Zhang et al., 2005).

The RNAP II active site is buried deep within the enzyme structure, limiting accessibility and free exchange of NTP substrates. We suggest that NTP substrates may enter the enzyme through the main enzyme channel and not necessarily through the secondary pore. In this manner, the negative electrostatic potential of the pore (Batada et al., 2004) can be bypassed in NTP loading. We posit that loading of NTP substrates to downstream DNA sites helps to induce conformational changes in the elongation complex that drive forward translocation (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007) and substrate NTP tightening.

6. The rock and roll model for RNA polymerase II translocation

The "rock and roll" model is proposed to describe the multi-subunit RNAP catalytic mechanism in molecular details. The model, which might apply to both eukaryotic RNAP II and bacterial RNAP, is shown schematically in Figure I-3. "Rock and roll" refers to rocking of the "rocking assembly" (*S. cerevisiae* rocking assembly I: Rpb1 526 to 764, Rpb8, Rpb9 48 to 120; rocking assembly II: Rpb9 1 to 46, Rpb1 869 to 1063, Rpb1 1113 to 1308), stimulated by i+1 NTP tightening and NTP-driven translocation. Motion of the rocking assemblies induces reversible rolling of the "rolling assembly" (Rpb1 810 to 868 (bridge α-helix and C-terminal cap), Rpb1 764 to 810, Rpb2 500 to 542 ("fork loop 2" region), Rpb2 750 to 778, Rpb2 706 to 738, Rpb1 1384 to 1402, Rpb1 1063 to 1112) (Figures I-4, 5, 6 and 7). Briefly stated, we propose that at least three NTP substrates can be on template simultaneously at i+3, i+2, and i+1 (the active site). We propose significant conformational coupling between the NTP substrate at i+1 and downstream NTPs at i+2 and i+3. Essentially, forward translocation, partly driven by i+2 and i+3 NTPs, is coupled to i+1 NTP tightening, which in turn is coupled to conformational changes that drive stepped forward translocation. The rock and roll model unifies NTP-driven translocation and events at the active site, explaining how NTP-driven translocation can be coupled to pyrophosphate release (Gong et al., 2005; Xiong and Burton, 2007; Zhang et al., 2005). Close resemblance of the schematic to the structure is indicated in Figure I-3F.

Figure I-3. Schematic diagram of rock and roll. A) Relaxed elongation complex. B) Tightened elongation complex. C) Phosphodiester bond formation. D) NTP-driven translocation coupled to switching DNA contacts and return of the translocation arm. E) Relaxed elongation complex. F) Similarity between the elongation complex structure and the schematic. Tightening of the i+1 NTP and NTP-driven translocation cause the rocking assembly to rock forward and the rolling assembly to roll forward. Tightening of the i+1 NTP also draws the translocation arm forward through a set of anti-parallel β-sheets linked to active

site loops. The translocation arm projects multiple basic groups to interact with the DNA template. Rpb1 K752 is posited to be a major sensor of the stage of the bond addition cycle.
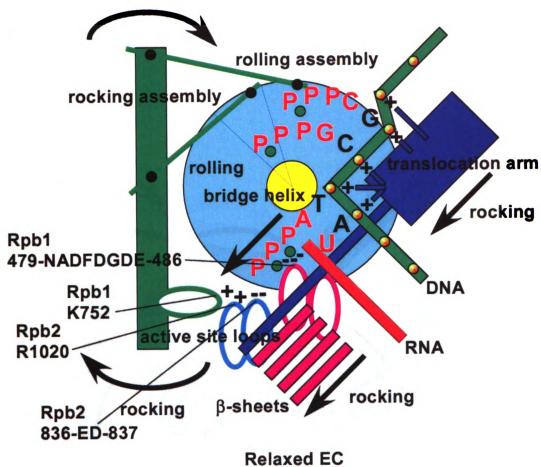
A

Figure I-3.

Figure I-3 (continued).

NTP Tightening
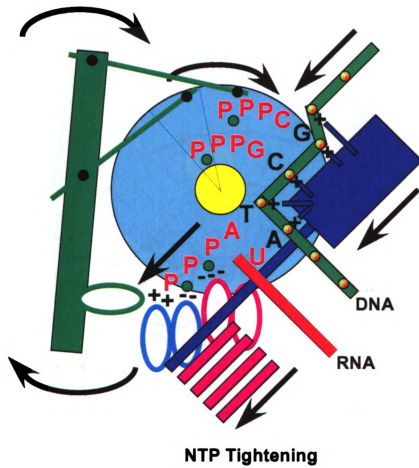
B

Figure I-3 (continued).
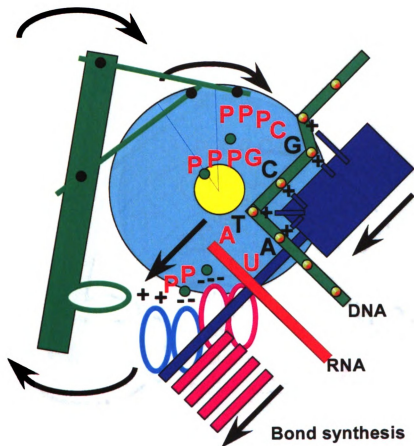
C

Figure I-3 (continued).
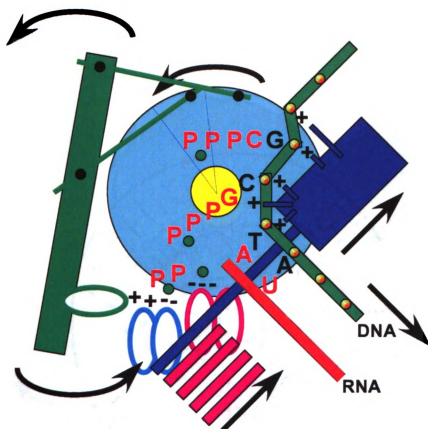
NTP-driven translocation,
switching of DNA contacts
return of translocation arm

D

Figure I-3 (continued).
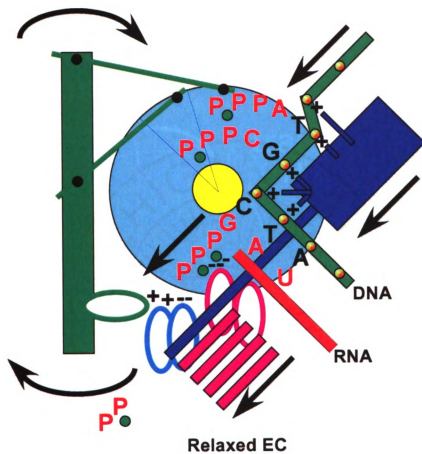
Relaxed EC

E

Figure I-3 (continued).

F

39
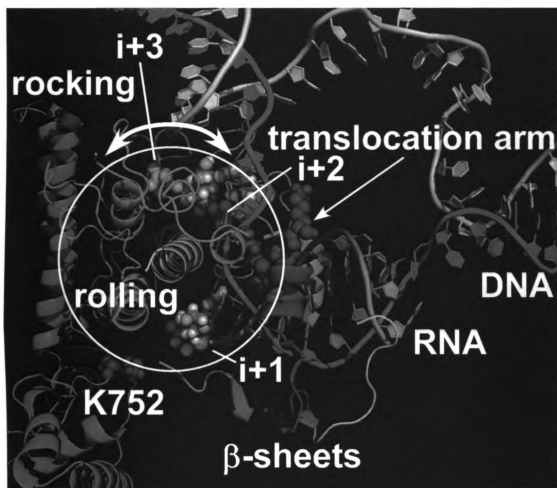
Figure I-3A shows a relaxed elongation complex structure with an ATP-Mg^{2+} loaded at i+1, a GTP-Mg^{2+} at i+2, and a CTP-Mg^{2+} at i+3. Current X-ray crystal structures with NTPs loaded at i+1 resemble the image shown in this figure. The ATP-Mg^{2+} is in a relaxed conformation in the active site. Mg-A is held by Rpb1 479-NA<u>DF</u>D<u>G</u>DE-486. Mg-B is out of range of Rpb2 830-YSGYNQ<u>ED</u>S-838 (see Figure I-8; 2E2H structure). Figure I-3B depicts tightening ("isomerization") of the i+1 ATP-Mg^{2+}. Mg-A is held by Rpb1 479-NA<u>DF</u>D<u>G</u>DE-486. Mg-B is held by Rpb2 830-YSGYNQ<u>ED</u>S-838, as observed in the 2NVT crystal structure (Figure I-8). Rpb2 R1020 and Rpb1 K752 are brought close to the γ-phosphate of the ATP. Contact of Rpb1 K752 with the γ-phosphate of ATP generates translocation force through movement of the helix 20-21-22-23 assembly (Rpb1 674 to 763), which includes K752 on the Rpb1 750-GSKG-753 active site loop. Figure I-9C indicates that further translocation force is generated as the phosphodiester bond is formed. Furthermore, after chemistry, pyrophosphate is retained in the active site. This feature of the model is indicated by kinetic studies, which show that bonds can be formed and then reversed, apparently through endogenous pyrophosphorolysis, which is detected as bond reversal using a combination of EDTA quench and HCl quench procedures. By coupling NTP-Mg^{2+} tightening, bond formation, and pyrophosphate extraction to ongoing translocation, induced fit within the i+1 active site is coupled to transcriptional fidelity. At any point in the catalytic cycle, failure to achieve optimal geometry within the active site results in backwards winding of the ratchet and substrate rejection. In Figures I-3A-3C, the translocation arm exerts forward

translocation pressure on the DNA template strand through lysine and arginine contacts to DNA phosphates. Maximal translocation strain in the system occurs after chemistry but prior to the next NTP loading, which in this case requires GTP-Mg^{2+} transfer from the i+2 downstream position to i+1. Figure I-3D shows a feature of the model, which is NTP-driven translocation coupled to pyrophosphate release. We have depicted the later stage of this transition as the return of the rocking assembly and bridge helix ratchet to their relaxed positions. Pyrophosphate is released, perhaps because of return of the rocking assembly and Rpb1 K752. Contacts between the translocation arm and the DNA template must first be broken and then be re-set at the next downstream position. Changing the bend point in the DNA template strand is likely to be an important feature of re-setting phosphate contacts during return of the translocation arm. In Figure I-3E, the relaxed elongation complex is shown with GTP-Mg^{2+} in the i+1 active site. Figure I-3F shows a crystal structure that resembles the relaxed complexes shown in Figures I-3A and 3E.

Mutations in yeast RNAP II residues that are defective in elongation may support the rock and roll model. Remarkably, α-amanitin, a potent translocation blocker (Gong et al., 2004), is poised to block interaction between the rocking assembly and the rolling assembly, inhibiting rocking and preventing rolling (Figure I-9). The structure of α-amanitin bound to RNA polymerase II and the distribution and precise locations of mutant RNA polymerase II proteins provide support for the rock and roll model. Trigger loop dynamics could be incorporated into the "rock and roll" model.

Figure I-4. The rocking assembly rocks, and the rolling assembly rolls. The left panel is the view from the N-terminal end of the bridge a-helix (blue; Rpb1 810 to 845). The N-terminal end of the bridge a-helix can be identified by the "propeller" (P; Rpb1 813-FFFH-816). The right panel is the view from the C-terminal end of the bridge α-helix. The i+1 NTP is orange. The i+2 NTP (placed by modeling) is pink. The i+3 NTP (placed by modeling) is lime. The DNA template strand is green. The RNA strand is red. The non-template DNA strand is white. Phosphates along the DNA template strand are gold. Mg is green. Basic amino acids are blue. Rpb1 K752 is proposed to be a sensor of the state of the catalytic cycle. Hydrophobic amino acids are white. DNA chains were extended and disordered parts of the non-template DNA strand were placed by modeling.

Figure I-4.

Rocking

Rolling

P

K752

43
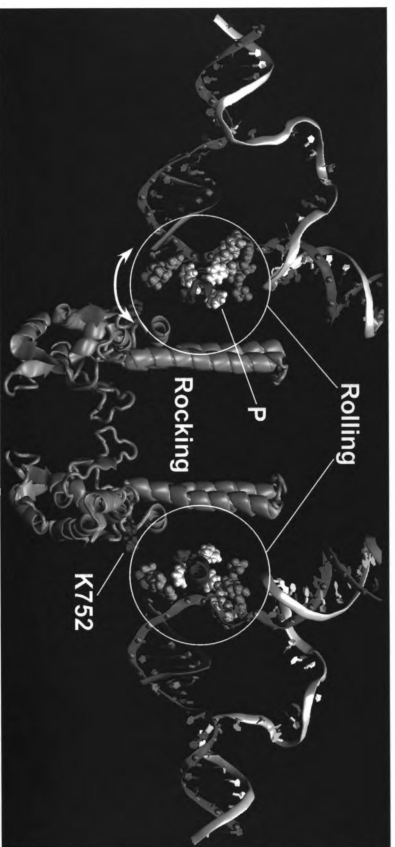
Figure I-5. The rolling assembly. The image is the same as shown in **Figure I-4** except the rolling assembly is shown in additional detail. The protein segment that connects the rocking assembly to the bridge α-helix is shown in purple (Rpb1 764 to 809). The fork loop 2 region is shown in cyan (Rpb2 500 to 542; proposed i+2 NTP interaction region). The helix 36-trigger loop-helix 37 region is yellow (Rpb1 1060 to 1112; proposed i+3 NTP interaction region).

Figure I-5.

Rocking

Rolling
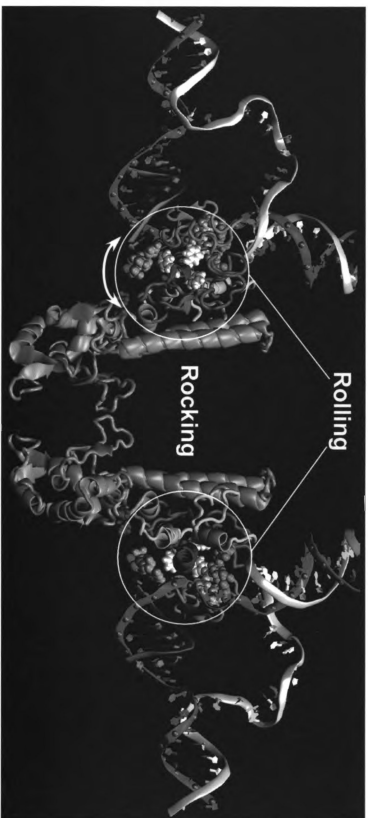
Figure I-6. Functional domains of the *S. cerevisiae* Rpb1 subunit. Active site loops are indicated. In some cases, important residues are highlighted in red. Relevant mutations are in red. Homologous residues and structures found in bacterial RNA polymerase from *T. thermophilus* are indicated in green.

R326 (-)
K330 (-)
K332 (K610)
R337 (R615)
R344 (R622)
R350 (R628)
Translocation Arm
326
(604)

Active Site

1
(1)

347   356   525
(625)   (634)   (778)

Rpb1
(β' homolog)

446-RQPSLHKMS-454
(684-RAPTLHRLG-692)
479-NADFDGDE-486
(737-NADFDGDN-744)
Mg-A
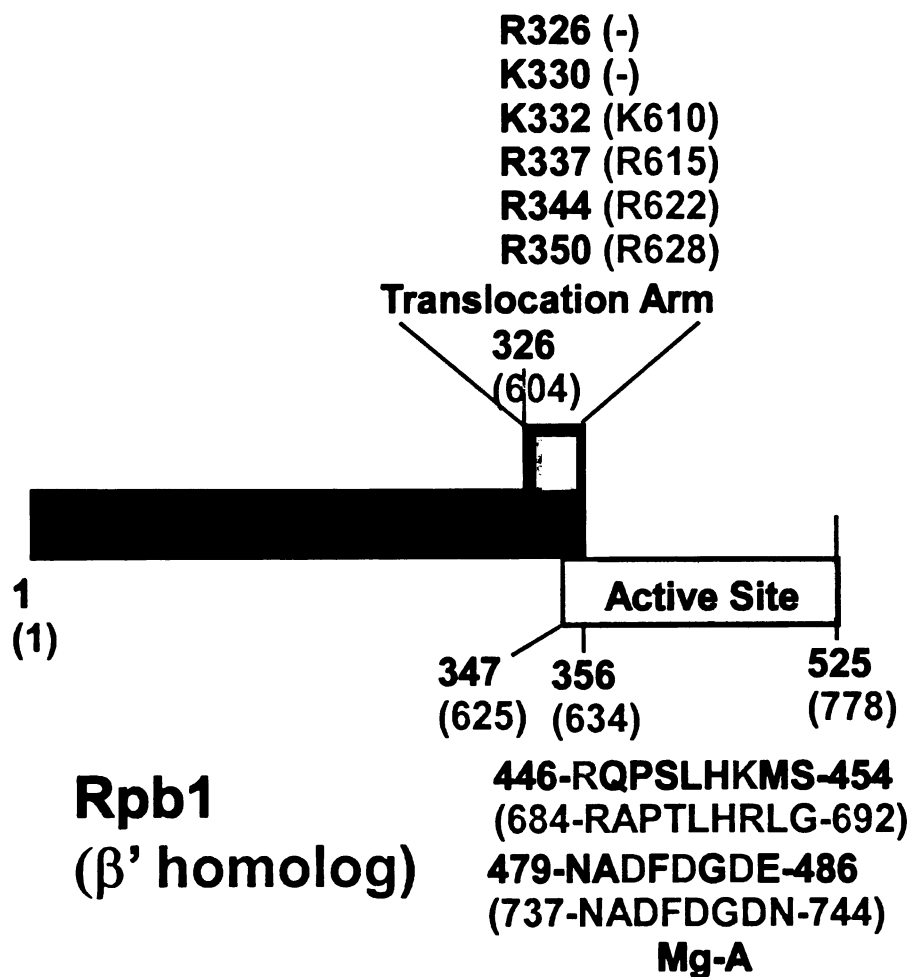
Figure I-6.

Figure I-6 (continued).

Rpb5 interaction
Rpb6 interaction
(ω interaction)
C-terminal Cap on Bridge
Piston/Rolling Assembly

Y836 (Y1093)
R839 (R1096)
R840 (K1097)
K843 (-)
E846 (-)
819-GGREG-823
(1076-GARKG-1080)
813-FFFH-816
(1070-YFI-1073)

Bridge Helix Ratchet
(Rolling Assembly)

Rpb8 interaction

673
(954)
763
(1040)
809
(1067)
845
(1105)

Rocking Assembly I

Connector to Bridge
(Rolling Assembly)

Helix 20-21-22-23 Assembly
750-GSKG-753
(1026-GARG-1029)
G730D
S751F
A759P
C764F

Rpb9 interaction

Figure I-6 (continued).

i+3 NTP
P1099 (P1257)
R1100 (R1258)
K1102 (-)
E1103 E1103G (E1261)
N1106 (-)
K1112 (K1269)

Helix 36-Trigger loop-Helix 37
(Rolling Assembly)

Rpb5 interaction
Rpb6 interaction
(ω interaction)
1446
(1505)

Piston | CTD

869
(1125)

1063
(1220) **Rpb9 interaction**

1308

I1327N
E1351K
H1367D

E1351 (E1351)
V1384 (V1424)
R1386 (K1426)
F1389 (L1429)
R1391 (-)
F1402 (F1440)

1113
(1271)

Rocking Assembly II
(Piston)

E1062A (E1219)
A1076V
S1096F
R1281L

Figure I-7. Functional domains of the *S. cerevisiae* **Rpb2 subunit.** Colors are

as in **Figure I-6.**
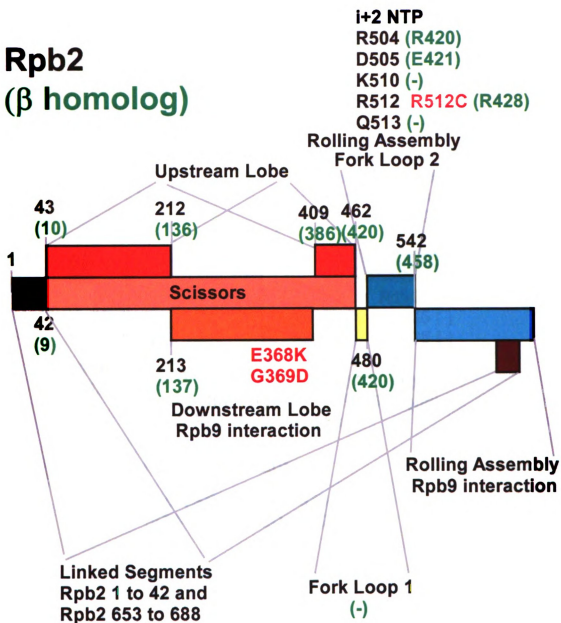
Figure I-7.

Figure I-7 (continued).

830-YSGYNQ**ED**S-838
(679-FDGYNFEDA-687)

Mg-B?

979-**K**FAS**R**HGQ**K**G-988
(838-KLANRHGNKG-847)

1013-NPHAIPS**R**MT-1022
(872-NPLGVPSRMN-881)

A753V
H761Y
A777T
R983G
A1016T
P1018S

R983-E1028
(R842-E887)
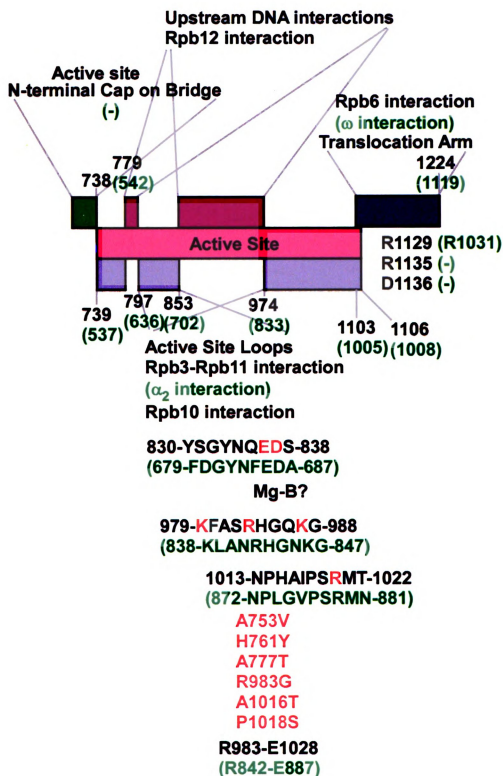
Figure I-8. The two-metal catalytic mechanism for RNA polymerase II and the trigger loop hypothesis. Three identical views from Protein Data Bank 2NVT (upper panels) and 2E2H (lower panels) crystal structures are shown. The 2NVT structure has a straighter conformation of the bridge α-helix (blue), a more ordered helix 36 (H36) (yellow), an "open" conformation of the trigger loop (yellow), and a wider spacing of Mg-A and Mg-B (green; A and B). The 2E2H structure has a bent conformation of the bridge α-helix, a disordered helix 36, a "closed" conformation of the trigger loop, and little separation of Mg-A and Mg-B. Acidic amino acids are red (Rpb1 D481, D483, D485 and Rpb2 E836, D837). The i+2 (pink) and i+3 (lime) NTPs were placed by modeling. Colors are the same as in **Figures I-4** and **5**.
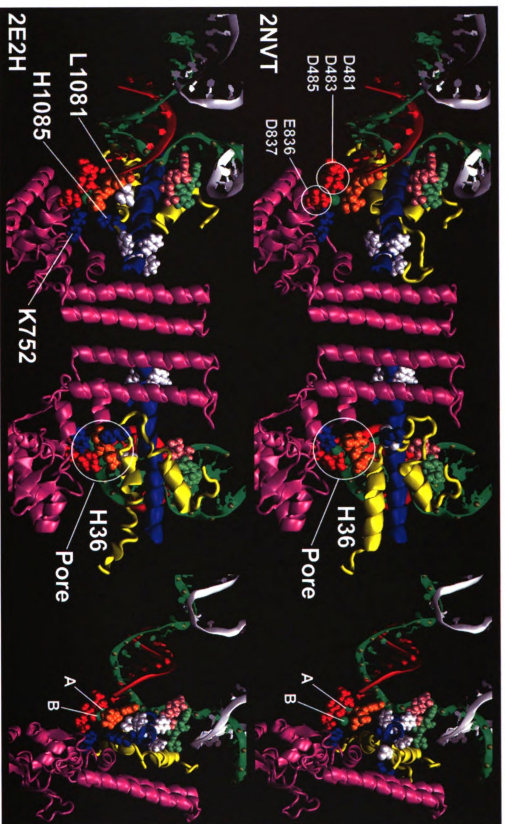
Figure I-8.

Figure I-9. α-amanitin (red) is a cyclic octapeptide that blocks rocking of the rocking assembly (mauve) and inhibits rolling of the rolling assembly. The helix 36-trigger loop-helix 37 region (Rpb1 1060 to 1112) is yellow. The connecting segment between the rocking assembly and the bridge α-helix (Rpb1 764 to 809) is cyan. The positions of three clustered SIT mutants in Rpb1 rocking assembly I (G730D, A759P, and C764F) border the α-amanitin binding site. i+2 and i+3 NTPs and portions of the non-template DNA strand and template DNA strand were placed by modeling. Other colors are as in **Figures I-4** and **5**.
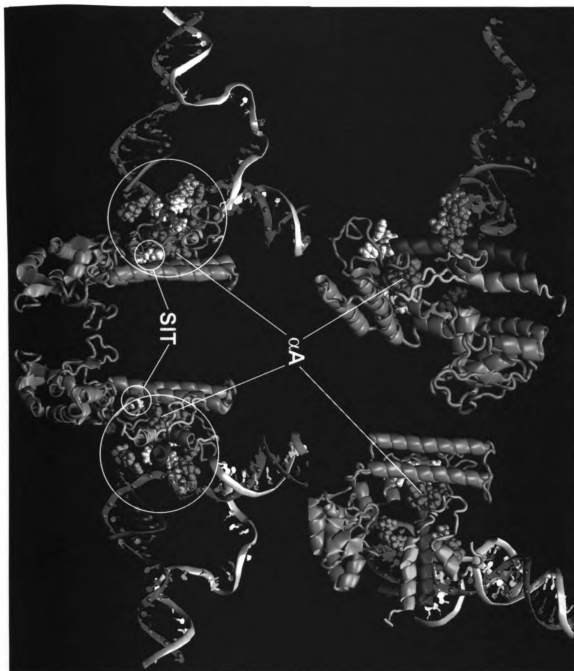
Figure I-9.

7. Fidelity and efficiency

NTP-driven translocation coupled to rock and roll provides a new view of the efficiency and fidelity of RNA synthesis (Burton et al., 2005; Gong et al., 2005; Xiong and Burton, 2007). Pre-loading and pre-aligning of NTP substrates results in highly efficient polymerization, causing RNA polymerase to resemble a miniature industrial assembly line for building an RNA chain. The RNA assembly line metaphor also suggests mechanisms for sophisticated quality control in RNA polymerization, some of which have been demonstrated to be novel error prevention and error correction mechanisms (Gong et al., 2005; Xiong and Burton, 2007). NTPs are tested for accurate base pairing at downstream i+3 and i+2 sites, before loading into the i+1 active site, which is the only position at which bond formation and hence transcription errors can occur. The triphosphate tails of incoming NTPs appear to be scanned at the i+3 site (yeast Rpb1 K1102, N1106, and K1112) and again at the i+2 site (Rpb2 R504, R512, and Q513) (Figures I-10 and 11). For RNA polymerase II, the 2'-OH of the ribose ring of the incoming NTP may be monitored at the i+2 position by an aspartic acid (Rpb2 D505) (Burton et al., 2005). An aspartate residue provides a similar function in the pre-insertion site of RNA-dependent RNAP from (Tellez et al., 2006; Thompson and Peersen, 2004). The space for NTP-dNMP passage from the i+2 to the i+1 site is limited, so translocation through a constrained space is proposed to provide a fidelity check, in which inappropriate base pairs are disrupted. In the active site, induced fit is utilized to confirm the accuracy of NTP

loading. Prior to pyrophosphate release, there is a mechanism for reversing phosphodiester bond synthesis in response to a block to translocation (Gong et al., 2005; Xiong and Burton, 2007). RNAP II, therefore, senses a transcription error as a translocation block. The purpose of the bond reversal mechanism is to remove a 3'-terminal, improperly loaded or incorporated base from the RNA chain prior to pyrophosphate release and essentially irreversible elongation.

Figure I-10. The proposed i+2 NTP interaction site. Rpb2 R504, D505, R512, and Q513 on fork loop 2 (cyan; Rpb2 504 to 513) are posited to make contacts to the i+2 NTP. The bridge α-helix is yellow. NTPs and amino acids are colored by chemistry. Mg-A and Mg-B are purple. The white arrow indicates the 90 degree bend in the DNA template over the bridge α-helix. The i+2 NTP was placed by modeling.
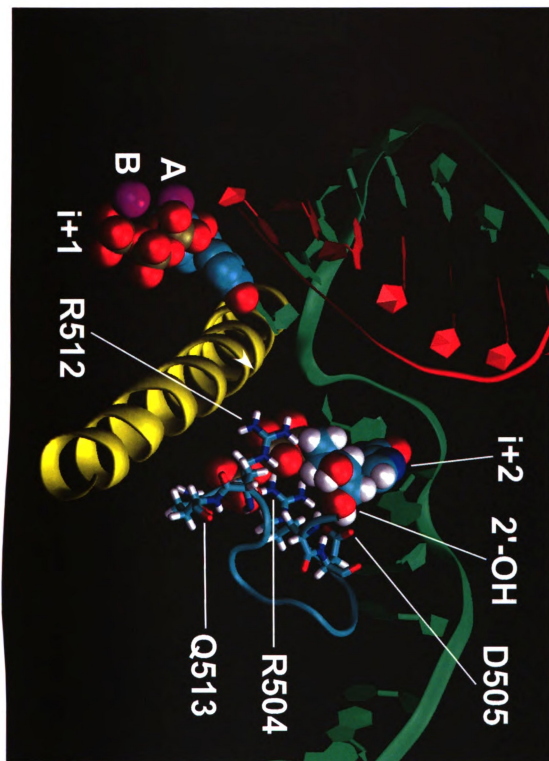
Figure I-10.

Figure I-11. The proposed i+3 NTP interaction site. Rpb1 K1102, N1106, and K1112 are posited to make important contacts to the i+3 NTP. Rpb1 R840 on the bridge α-helix (yellow) makes main chain contacts to Rpb1 N1106. The helix 36-trigger loop-helix 37 region (Rpb1 1063 to 1112) is mauve. The fork loop 2 region (Rpb2 500 to 542) is cyan. The i+2 and i+3 NTPs were placed by modeling.
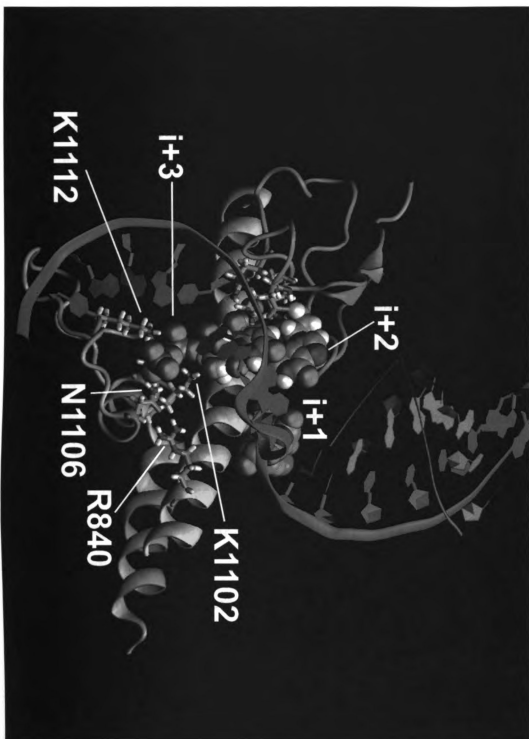
Figure I-11.

8. Conclusions

A model is proposed to unify kinetic, structural, evolutionary, and genetic data on RNAP II. In addition to aiding in the interpretation of a vast collection of available data, the model makes extensive and precise predictions for future experiments. Based on evolutionary comparisons (Figures I-12, 13, 14, 6 and 7), the rock and roll model may be applicable to all multi-subunit RNAPs. Multi-subunit RNAPs, found in eubacteria, archaea, and eukarya, appear to represent a significant evolutionary advance allowing development of complex life on earth. Eubacteria may have existed on earth for 3.2 to 3.5 billion years, indicating the age of multi-subunit RNAPs. The enclosed main enzyme channel may have evolved to support NTP-driven translocation coupled to rock and roll in order to: 1) enhance transcriptional fidelity and efficiency; 2) support processivity; and 3) enforce single base stepping during elongation to prohibit unrestrained sliding. Alternatively, the trigger loop-trigger helices model posits that NTP selection and induced fit occur only within the enzyme active site and do not involve downstream NTP interactions.

Figure I-12. Six active site loops act as sensors of the phase of the bond addition cycle and respond to i+1 NTP tightening. The active site is conformationally coupled to the rocking and rolling assemblies. Black rectangles indicate ion pairs. Ovals indicate adjacent residues. In this figure, dotted lines indicate coordination to Mg^{2+}. A = archaea, B = eubacteria, I, II, III = RNA polymerases I, II, and III, to indicate conservation of residues (similarity or identity). Relevant mutations are indicated in red.

Rpb1 446-RQPSLHKMS-454
Active site loop sensor

N445T,S

Rpb1
K452
A, B, I, II, III

Rpb1
E1062
A, B, II

Rolling Assembly
A1076V
S1096F
V1107(AR,AR)A

Rpb1
R350
A, B, I, II, III

Arm
Translocation

Rpb1
E486
A, I, II, III

Rpb1 479-NADFDGDE-486
Active site loop sensor

R983G

Rpb2
R983
A, B, I, II, III

N479Y

Mg-A

N488D
P492A

N487A,I,L

Rpb2
K979
A, B, I, II, III

Rpb2
E1028
A, B, I, II, III

Rpb2 830-YSGYNQEDS-838
Active site loop sensor

Rpb2
D837
A, B, I, II, III

Mg-B?

P1018S
A1016T

Rpb2
R1020
A, B, I, II, III

γ-phosphate
i+1 NTP

Rpb2
E836
A, B, I, II, III

Rpb1
K752
A, B, I, II, III

S751F

Rpb1 750-GSKG-753
Active site loop sensor
Rocking Assembly

G730D
A759P
C764F

i-phosphate RNA

Rpb2
K987
A, B, I, II, III

i+1 NTP

Rpb2 979-KFASRHGQK-987
Active site loop sensor

Rpb2 1013-NPHAIPSRMT-1022
Active site loop sensor

Figure I-12.

Figure I-13. Conformational coupling between the i+1 NTP, the rocking assembly, the piston assembly, and the proposed i+3 NTP site. Symbols are as in **Figure I-12**. The black bar indicates a proposed 35 angstrom charge relay. Parentheses indicate that one ion pair interaction (Rpb5 E148-Rpb1 K1350) is slightly out of range (6.5 angstroms) in available structures. Dotted lines indicate structural proximity and probable coupled movement.

Rpb1
D1323
A, I, II

Rpb1
R1289 ■ E1303 ■ R1326
II

Rpb1
A, I*, II

Rpb1
A, I*, II, II

Rpb5
E148 (■) I1327N
I, II, III

Rpb1
D1309 ■ R1281 ■ E712
A, I, II, III A, I, II, III A, II*, III*
R1281L G730D
A759P
C764F

Rocking Assembly

S751F

Rpb1
K350 ●
A*, I*, II, III

Rpb1
K752 ■ i+1 NTP
A, B, I, II, III

Active site loop
i+1 NTP sensor
750-GSKG-753

E1351K

Piston Assembly
Charge Relay
i+3 NTP interaction

Rpb1
E1351 ■
A, B, I, II, III

Rpb1
R100 ●
A, B, I, II, III

E1103G

Rpb1
D1342 ■
A, I, II, III

Rpb1
E1103 ●
I, II, III

V1107(ARAR)A
E1062A
A1076V
S1096F

Rolling
Assembly

Rpb1
K1102 ■
I, II, III

i+3 NTP
triphosphate

Rpb1
K1112
B, II

Rpb5
R200 ●
I, II, III

Rpb1
R1345 ■
A, I, II, III

Rpb1
D1373
A, I, II, III

H1367D

Piston Assembly

Figure I-13.

67

Figure I-14. Connections of the i+2 NTP interaction site on the rolling assembly with the rocking assembly and active site. Symbols are as in Figures I-12 and 13.

Rolling Assembly

i+2 NTP
triphosphate
2'-OH

A777T

Rpb2
K775
A, B, I, II, III ■

H761Y ● D760
Rpb2
A, B, I, II

A753V

Rpb2
R728
II*
E801
Rpb1

Bridge
R730 N802
Rpb2
II*
R806

Rpb2
D1049
II ■

Rpb2
E526
A, B, I, II, III

i+2 NTP
triphosphate

Rpb2
R504 ● D505
A, B, II, III A*, B, I*, II

Rpb2
K510 ■
I, II

Rpb1
E771 - - - **Rocking Assembly**

i+2 NTP
γ-phosphate

Rpb2
Q513 ● R512
II A, B, I, II, III
R512C

Rpb1
R821
819-GGREG-823
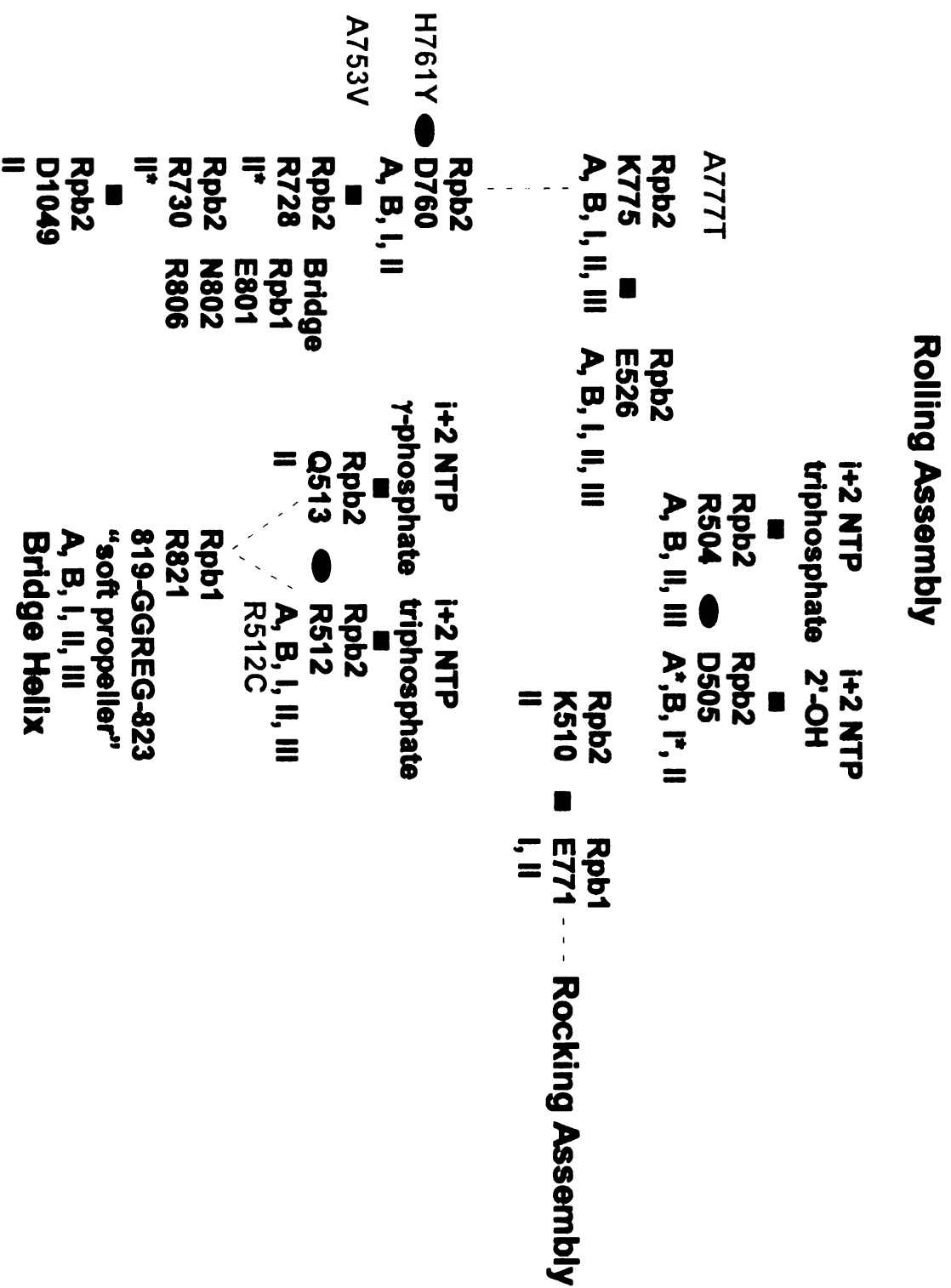"soft propeller"
A, B, I, II, III

Bridge Helix

Figure I-14.

69

References:

Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R., and Block, S. M. (2005). Direct observation of base-pair stepping by RNA polymerase. Nature *438*, 460-465.

Arnold, J. J., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg2+. Biochemistry *43*, 5126-5137.

Arnold, J. J., Gohara, D. W., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn2+. Biochemistry *43*, 5138-5148.

Artsimovitch, I., and Vassylyev, D. G. (2006). Is it easy to stop RNA polymerase? Cell Cycle *5*, 399-404.

Artsimovitch, I., and Vassylyev, D. G. (2007). Merging the RNA and DNA worlds. Nat Struct Mol Biol *14*, 1122-1123.

Artsimovitch, I., Vassylyeva, M. N., Svetlov, D., Svetlov, V., Perederina, A., Igarashi, N., Matsugaki, N., Wakatsuki, S., Tahirov, T. H., and Vassylyev, D. G. (2005). Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins. Cell *122*, 351-363.

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*, 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*, 17361-17364.

Bell, S. D., and Jackson, S. P. (1998). Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol *6*, 222-228.

Borukhov, S., and Nudler, E. (2003). RNA polymerase holoenzyme: structure, function and biological implications. Curr Opin Microbiol *6*, 93-100.

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*, 486-496.

Bushnell, D. A., Cramer, P., and Kornberg, R. D. (2002). Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. Proc Natl Acad Sci U S A *99*, 1218-1222.

Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science *303*, 983-988.

Chafin, D. R., Guo, H., and Price, D. H. (1995). Action of alpha-amanitin during pyrophosphorolysis and elongation by RNA polymerase II. J Biol Chem *270*, 19114-19119.

Chen, H. T., and Hahn, S. (2003). Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. Mol Cell *12*, 437-447.

Chung, W. H., Craighead, J. L., Chang, W. H., Ezeokonkwo, C., Bareket-Samish, A., Kornberg, R. D., and Asturias, F. J. (2003). RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. Mol Cell *12*, 1003-1013.

Cochet-Meilhac, M., and Chambon, P. (1974). Animal DNA-dependent RNA polymerases. 11. Mechanism of the inhibition of RNA polymerases B by amatoxins. Biochim Biophys Acta *353*, 160-184.

Conaway, R. C., and Conaway, J. W. (1997). General transcription factors for RNA polymerase II. Prog Nucleic Acid Res Mol Biol *56*, 327-346.

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

Ebright, R. H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. J Mol Biol *304*, 687-698.

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*, 623-634.

Foster, J. E., Holmes, S. F., and Erie, D. A. (2001). Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation. Cell *106*, 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 3.3 A resolution. Science *292*, 1876-1882.

Gong, X. Q., Nedialkov, Y. A., and Burton, Z. F. (2004). Alpha-amanitin blocks translocation by human RNA polymerase II. J Biol Chem *279*, 27422-27427.

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*, 461-470.

Grummt, I. (2003). Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. Genes Dev *17*, 1691-1702.

Henikoff, S., and Ahmad, K. (2005). Assembly of variant histones into chromatin. Annu Rev Cell Dev Biol *21*, 133-153.

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*, 35597-35608.

Holmes, S. F., Foster, J. E., and Erie, D. A. (2003). Kinetics of multisubunit RNA polymerases: experimental methods and data analysis. Methods Enzymol *371*, 71-81.

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*, 18677-18683.

Kedinger, C., Gniazdowski, M., Mandel, J. L., Jr., Gissinger, F., and Chambon, P. (1970). Alpha-amanitin: a specific inhibitor of one of two DNA-pendent RNA polymerase activities from calf thymus. Biochem Biophys Res Commun *38*, 165-171.

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*, 955-965.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. Cell *116*, 259-272.

Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev *14*, 2452-2460.

Kuhlman, T. C., Cho, H., Reinberg, D., and Hernandez, N. (1999). The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. Mol Cell Biol *19*, 2130-2141.

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

Lee, T. I., and Young, R. A. (2000). Transcription of eukaryotic protein-coding genes. Annu Rev Genet *34*, 77-137.

Lindell, T. J., Weinberg, F., Morris, P. W., Roeder, R. G., and Rutter, W. J. (1970). Specific inhibition of nuclear RNA polymerase II by alpha-amanitin. Science *170*, 447-449.

Lutter, L. C. (1982). Photoreactivation of amanitin-inhibited RNA polymerase II. J Biol Chem *257*, 1577-1578.

Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. Genes Dev *19*, 1401-1415.

Murakami, K. S., and Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol *13*, 31-39.

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*, 18303-18312.

Ohkuma, Y. (1997). Multiple functions of general transcription factors TFIIE and TFIIH in transcription: possible points of regulation by trans-acting factors. J Biochem *122*, 481-489.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev *10*, 2657-2683.

Oster, G. (2002). Brownian ratchets: Darwin's motors. Nature *417*, 25.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., *et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Prelich, G. (2002). RNA polymerase II carboxy-terminal domain kinases: emerging clues to their function. Eukaryot Cell *1*, 153-162.

Reinberg, D., Orphanides, G., Ebright, R., Akoulitchev, S., Carcamo, J., Cho, H., Cortes, P., Drapkin, R., Flores, O., Ha, I., *et al.* (1998). The RNA polymerase II general transcription factors: past, present, and future. Cold Spring Harb Symp Quant Biol *63*, 83-103.

Roeder, R. G. (1996). Nuclear RNA polymerases: role of general initiation factors and cofactors in eukaryotic transcription. Methods Enzymol *273*, 165-171.

Rudd, M. D., and Luse, D. S. (1996). Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. J Biol Chem *271*, 21549-21558.

Sawa
trans

Schr
its ta

Sims
RNA

Sou
Meth

Tah
D. O
com

Tell
Bru
RNA
Bio

Tho
dep
23.

Vas
ste

Va
S.,
hol

Va
Art
RN

Va
ang
pol

Wa
(20
spe

Wea
poly
Sci

Sawadogo, M., and Sentenac, A. (1990). RNA polymerase B (II) and general transcription factors. Annu Rev Biochem *59*, 711-754.

Schramm, L., and Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. Genes Dev *16*, 2593-2620.

Sims, R. J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev *18*, 2437-2468.

Sousa, R. (2003). On models and methods for studying polymerase translocation. Methods Enzymol *371*, 3-13.

Tahirov, T. H., Temiakov, D., Anikin, M., Patlan, V., McAllister, W. T., Vassylyev, D. G., and Yokoyama, S. (2002). Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. Nature *420*, 43-50.

Tellez, A. B., Crowder, S., Spagnolo, J. F., Thompson, A. A., Peersen, O. B., Brutlag, D. L., and Kirkegaard, K. (2006). Nucleotide channel of RNA-dependent RNA polymerase used for intermolecular uridylylation of protein primer. J Mol Biol *357*, 665-675.

Thompson, A. A., and Peersen, O. B. (2004). Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. Embo J *23*, 3462-3471.

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

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 2.6 A resolution. Nature *417*, 712-719.

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*, 157-162.

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*, 163-168.

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*, 941-954.

Weinmann, R., and Roeder, R. G. (1974). Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. Proc Natl Acad Sci U S A *71*, 1790-1794.

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*, 481-489.

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*, 1014-1016.

Woychik, N. A., and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. Cell *108*, 453-463.

Xiong, Y., and Burton, Z. F. (2007). A tunable ratchet driving human RNA polymerase II translocation adjusted by accurately templated nucleoside triphosphates loaded at downstream sites and by elongation factors. J Biol Chem *282*, 36582-36592.

Yin, Y. W., and Steitz, T. A. (2002). Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science *298*, 1387-1395.

Zhang, C., and Burton, Z. F. (2004). Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. J Mol Biol *342*, 1085-1099.

Zhang, C., Yan, H., and Burton, Z. F. (2003). Combinatorial control of human RNA polymerase II (RNAP II) pausing and transcript cleavage by transcription factor IIF, hepatitis delta antigen, and stimulatory factor II. J Biol Chem *278*, 50101-50111.

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*, 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 3.3 A resolution. Cell *98*, 811-824.

CHAPTER TWO

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, 486-496.

M

th

fo

er

m

m

co

co

tra

elo

enz

the

dow

CHAPTER TWO

NTP-DRIVEN TRANSLOCATION AND REGULATION OF DOWNSTREAM TEMPLATE OPENING BY MULTI-SUBUNIT RNA POLYMERASES

Abstract

Multi-subunit RNA polymerases bind nucleotide triphosphate (NTP) substrates in the pre-translocated state and carry the dNMP–NTP base pair into the active site for phosphoryl transfer. NTP-driven translocation requires that NTP substrates enter the main-enzyme channel before loading into the active site. Based on this model, a new view of fidelity and efficiency of RNA synthesis is proposed. The model predicts that, during processive elongation, NTP-driven translocation is coupled to a protein conformational change that allows pyrophosphate release: coupling the end of one bond-addition cycle to substrate loading and translocation for the next. We present a detailed model of the RNA polymerase II elongation complex based on 2 low-affinity NTP binding sites located in the main-enzyme channel. This model posits that NTP substrates, elongation factors, and the conserved Rpb2 subunit fork loop 2 cooperate to regulate opening of the downstream transcription bubble.

1

fu

th

1

m

ca

tr.

ZI

2(

al

Ea

N

2(

w

th.

da

be

re

1. Translocation models

The virtue of transient-state (pre-steady-state) kinetic analysis is that functional dynamic changes in the elongation complex can be tracked through the formation of specific bonds on a millisecond time scale (Johnson 1992, 1995). As such, the relationship between kinetic analyses and elongation mechanism becomes apparent. Based on transient-state studies of elongation catalyzed by human RNA polymerase II, we have proposed the nucleotide triphosphate (NTP) - driven translocation model (Nedialkov et al. 2003a, 2003b; Zhang et al. 2003, 2005; Gong et al. 2004; Zhang and Burton 2004; Gong et al. 2005). Competing models to describe elongation and translocation include the allosteric model, posited by Erie and colleagues (Foster et al. 2001; Holmes and Erie 2003), and the thermal or Brownian-ratchet model, championed by Sousa, Nudler, and others (Guajardo and Sousa 1997; Oster 2002; Wang and Oster 2002; Sousa 2003, 2005; Landick 2004; Bar-Nahum et al. 2005). We find fault with these models, but describe the merits of disparate views. The purpose of this review is to consider NTP-driven translocation in the context of structural data and alternative models. Because much of the evidence for our model has been published, we will concentrate on qualitative descriptions of our model and related models.

2. Rate-limiting steps during elongation

During processive RNA synthesis, human RNA polymerase II has 2 primary rate-limiting steps: NTP-driven translocation coupled to pyrophosphate release and phosphodiester bond synthesis (or a strongly coupled but unknown conformational change) (Figure II-1A) (Nedialkov et al. 2003a, 2003b; Zhang et al. 2003, 2005; Gong et al. 2004; Zhang and Burton 2004). These 2 rate-limiting steps are sufficient to explain overall rates of RNA polymerase II elongation (15 to 30 nucleotides per second). NTP loading is rapid but is initially low-affinity, as expected for an interaction mediated primarily through base-pairing, recognition of the triphosphate (i.e., through basic side chain contacts), and scanning of the ribose ring (to exclude dNTP loading). For human RNA polymerase II, isomerization is rapid, followed by relatively slow phosphodiester-bond synthesis. During a primary isomerization step, the incoming substrate NTP-Mg^{2+} is tightened in the RNA polymerase II active site, where the metal is sequestered from chelation by EDTA (Zhang and Burton 2004; Zhang et al. 2005). Despite tight binding of the incoming NTP-Mg^{2+}, subsequent phosphoryl transfer is delayed. This is a somewhat confusing result, because the phosphoryl transfer reaction is expected to proceed very rapidly (Patel et al. 1991). One way to explain this apparent contradiction is to speculate that, after the NTP-Mg^{2+} is initially tightened in the RNA polymerase II active site, the substrate must be further distorted (i.e., puckering the ribose sugar and (or) aligning the triphosphate) before bond formation can occur (Castro et al. 2005).

Figure II-1. Nucleotide trip-hosphate (NTP)-driven translocation for (A) multi-subunit RNA polymerases (RNAPs) and (d)NTP-driven translocation for (B) simple DNA (DNAPs) and RNA polymerases (proposed). Pre-translocation (pre) and post-translocation (post) elongation complexes (ECs) are indicated. Bold arrows indicate rate-limiting steps in the human RNA polymerase II mechanism. Rate-limiting steps may be different for some multi-subunit RNA polymerases. Bold type with an asterisk indicates an isomerized (closed) elongation complex. Elongation is shown for advancement from RNA or DNA length n to length n+2. Question marks indicate uncertainty in simple enzyme mechanisms (B). Continuous preloading of NTPs is indicated in the multi-subunit RNA polymerase mechanism (see subsequent figures). Elongation complexes that relate to published crystal structures are labeled: (*a*) Gnatt et al. 2001; (*b*) Westover et al. 2004*a*, 2004*b*; (*c*) Kettenberger et al. 2004; (*d*) Westover et al. 2004*b*; (*e*) Johnson et al. 2003; Johnson and Beese 2004; (*f*) Yin and Steitz 2002; Johnson et al. 2003; Johnson and Beese 2004; (*g*) Temiakov et al. 2004; (*h*) Yin and Steitz 2004; and (*i*) Yin and Steitz 2004. Active site (*i+1*) and downstream (*i+2* and *i+3*) positions are indicated.

Figure II-1.

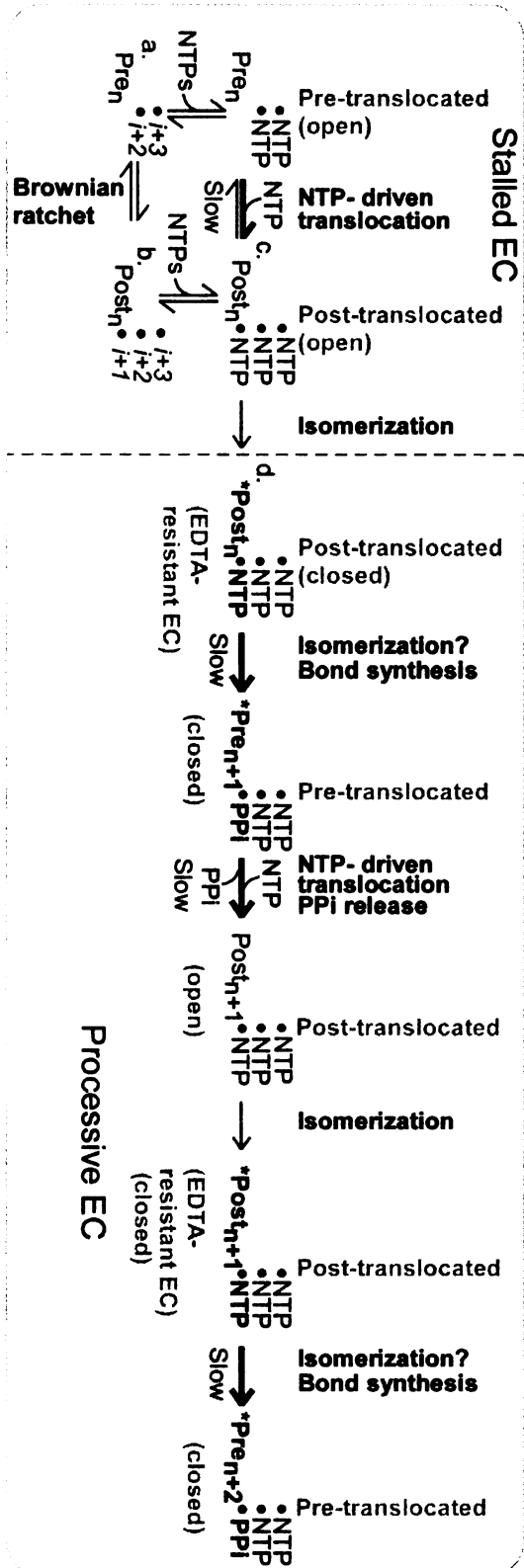**A. Multi-subunit RNAPs**

Stalled EC

Pre$_n$
$\bullet\bullet$NTP
$\bullet$NTP
Pre-translocated (open)

NTPs
a. Pre$_n$ $\begin{smallmatrix}\bullet\bullet & i+3 \\ \bullet & i+2\end{smallmatrix}$

Brownian ratchet

Slow — NTP-driven translocation

Pre$_n$ $\bullet\bullet$NTP $\bullet$NTP c. Post$_n\bullet$NTP

NTPs
b. Post$_n\bullet$ $\begin{smallmatrix}\bullet & i+3 \\ \bullet & i+2 \\ & i+1\end{smallmatrix}$

Post$_n\bullet$NTP
$\bullet\bullet$NTP
Post-translocated (open)

Isomerization

---

d. *Post$_n\bullet$NTP
$\bullet\bullet$NTP
(EDTA-resistant EC)
Post-translocated (closed)

Slow — Isomerization? Bond synthesis

*Pre$_{n+1}\bullet$PPi
$\bullet\bullet$NTP
$\bullet$NTP
(closed)
Pre-translocated

Slow — NTP-driven translocation PPi release
PPi NTP

Post$_{n+1}\bullet$NTP
$\bullet\bullet$NTP
(open)
Post-translocated

Isomerization

*Post$_{n+1}\bullet$NTP
$\bullet\bullet$NTP
(EDTA-resistant EC) (closed)
Post-translocated

Slow — Isomerization? Bond synthesis

*Pre$_{n+2}\bullet$PPi
$\bullet\bullet$NTP
(closed)
Pre-translocated

Processive EC

81

Figure II-1 (continued).

B. Simple RNAPs + DNAPs

Stalled EC

Pre-translocated, Pre-insertion EC (open)

•(d)NTP

Pre$_n$

Brownian ratchet

e. Pre$_n$ •i+2

(d)NTP

f. Post$_n$•i+1

(d)NTP

g. Post$_n$•(d)NTP

(d)NTP- driven translocation

Post-translocated, Pre-Insertion EC (open)

Post$_n$•(d)NTP

Isomerization

h. *Post$_n$•(d)NTP (EDTA-resistant EC)

Post-translocated, Insertion EC (closed)

Bond synthesis

i. *Pre$_{n+1}$•PPi (closed)

(d)NTP

translocation PPi release

*Pre$_{n+1}$•PPi (closed)

PPi

PPi

(d)NTP

Post$_{n+1}$• (open)

Processive EC

*Pre$_{n+1}$•PPi

•(d)NTP

Pre-translocated

(d)NTP

PPi

(d)NTP- driven translocation PPi release

Post$_{n+1}$•(d)NTP

Post-translocated

Isomerization

*Post$_{n+1}$•(d)NTP (EDTA-resistant EC) (closed)

Post-translocated

Bond synthesis

*Pre$_{n+2}$•PPi (closed)

Pre-translocated

82

Interesting similarities have been observed in the kinetics of RNA polymerase II (12 subunits) and simpler DNA and RNA polymerases, which are generally single subunit enzymes and are not homologous to multi-subunit enzymes. For the RNA-dependent RNA polymerase of poliovirus, analyzed in the presence of Mn^{2+}, isomerization is rapid, and is followed by rate-limiting bond formation (Arnold and Cameron 2004; Arnold et al. 2004; Castro et al. 2005). In this case, isomerization is defined by sequestration of NTP-Mn^{2+} in the active site, so that the metal is no longer available to react with EDTA as the reaction-quenching agent. On the other hand, with Mn^{2+} as the metal cofactor, for poliovirus RNA polymerase, isomerization (resistance to EDTA quench) is apparently coincident with bond formation (acid quench). This result indicates that, in the Mn^{2+}-supported reaction, after initial tight sequestration of the NTP-Mn^{2+} in the active site, another isomerization step might be necessary to complete bond synthesis. Of significant interest is the startling similarity of the poliovirus RNA polymerase elongation kinetics in the presence of Mn^{2+} to the human RNA polymerase II elongation reaction in the presence of the physiological metal Mg^{2+}.

3. NTP-driven translocation

Two relatively rapid kinetic phases are clearly resolved for human RNA polymerase II elongation from a stall position (Nedialkov et al. 2003a, 2003b; Zhang et al. 2003, 2005; Gong et al. 2004; Zhang and Burton 2004) (Figure II-

1A). These phases have been interpreted as elongation from the pre- and post-translocation states of the elongation complex. As expected, elongation from the post-translocation state is much faster than from the pre-translocation state. Occupancy of the 2 phases is changed in the presence of different sets of elongation factors, showing that accessory factors shift RNA polymerase II between functional modes (translocation states), as expected. The pre-translocation state is highly sensitive to α-amanitin inhibition, but the post-translocation state is resistant; the distinct rates of elongation clearly distinguish the 2 phases (Gong et al. 2004). The mushroom toxin α-amanitin has been shown to block translocation. Both pre- and post-translocation states are responsive to NTP substrates, showing that the pre-translocated elongation complex binds a templated NTP substrate.

The binding of an NTP to the pre-translocated elongation complex is consistent with the NTP-driven translocation model and the allosteric model, but not the Brownian-ratchet model. The Brownian-ratchet model has little meaning if the NTP substrate binds prior to translocation and drives translocation forward (Wang and Oster 2002; Landick 2004). The Brownian-ratchet mechanism requires that thermal motion drive translocation and that the NTP act as a pawl (as on a mechanical ratchet) to fix the elongation complex in the post-translocation state (Bar-Nahum et al. 2005; Sousa 2005). The observed binding of the substrate NTP to the pre-translocation state of the elongation complex converts the Brownian-ratchet model into the NTP-driven translocation model. The elongation mechanism for human RNA polymerase II, however, has

characteristics of both NTP-driven translocation (the favored pathway) and a Brownian-ratchet (the default pathway in the absence of NTP substrates) (Figure II-1A). When NTPs are not present, the elongation complex fractionates into both the pre- and post-translocation states, so RNA polymerases can function as Brownian ratchets, at least in the absence of NTPs or during transcriptional stalling.

The allosteric model is based on millisecond-phase kinetic studies of *Escherichia coli* RNA polymerase elongation (Foster et al. 2001; Holmes and Erie 2003). The model describes an allosteric transition between an unactivated state and an activated state of the RNA polymerase elongation complex, apparently requiring multiple interactions with the next incoming NTP substrate. The difficulty with the allosteric model is that, because the allosteric site is templated, allostery appears to require 2 NTP substrates simultaneously base-paired to the same DNA template base. As far as we can determine, there is no satisfactory explanation for this conundrum. We do, however, support the following predictions of the allosteric model. First, an NTP substrate bound in the pre-translocation state acts as an allosteric effector to drive translocation forward from the unactivated state (pre-translocation state) to the activated state (post-translocation state) (Zhang and Burton 2004; Zhang et al. 2005). As far as we can discern, the allosteric model reduces to the NTP-driven translocation mechanism, and no other view of the allosteric model makes physical sense. Second, our laboratory has demonstrated other allosteric effects of the incoming NTP substrate during transfer from the pre- to the post-translocation positions

(Gong et al. 2005). Specifically, the fate of a templated NTP tightened in the RNA polymerase II active site is linked to the presence of NTPs paired at the next 2 downstream template positions. Because these experiments demonstrate simultaneous occupancy of the pre- and post-translocation states of the elongation complex, they demonstrate the major tenet of the NTP-driven translocation model, which requires that both template positions be occupied by NTP substrates. So the allosteric model is verified to this extent: the incoming NTP substrate acts in a manner similar to an allosteric effector to drive translocation forward.

Another prediction of our model is that, during processive RNA synthesis, NTP-driven translocation is coupled to pyrophosphate release from the previous bond-addition cycle (Zhang and Burton 2004; Zhang et al. 2005). We posit that, during processive synthesis, pyrophosphate remains locked in a tightened (closed) conformation of the elongation complex. For pyrophosphate to be released, the active site must first be relaxed to the open conformation. NTP-driven translocation induces the closed→open conformational change in the elongation complex that results in pyrophosphate release (Figure II-1A).

This feature of our model is demonstrated when human RNA polymerase II elongation in the presence of TFIIF is compared with TFIIF-mutant proteins defective in stimulating elongation (Zhang et al. 2005). We identified TFIIF-mutant proteins that fail to fully support NTP-driven translocation. As a result, the processive transition between phosphodiester bonds, which requires release of pyrophosphate, becomes very slow. Because bond completion

(p

de

fa

pc

to

py

in

sta

or

be

inq

4.

po

20

tra

ass

por

cha

por

entr

(pyrophosphate release) relies on NTP-driven translocation, this step is highly dependent on the concentration of the incoming NTP. This NTP-driven step is facilitated by wild-type TFIIF. In the presence of the mutant TFIIF protein, RNA polymerase II escapes from a transcriptional stall by a mechanism that appears to be NTP-independent. Stalled RNA polymerase II lacks tightly bound pyrophosphate. Pyrophosphate release results in the acceleration of the NTP-independent translocation pathway, because relaxation of the active site during stalling frees the Brownian ratchet (Figure II-1A). During a processive transition, only the NTP-driven pathway is available until pyrophosphate can spontaneously be released, a slow process unless this step is facilitated by TFIIF and (or) the incoming NTP.

4. NTP loading into the RNA polymerase active site

Elongation complex structures have been published for yeast RNA polymerase II (Gnatt et al. 2001; Kettenberger et al. 2004; Westover et al. 2004a, 2004b) (Figures II- 2–4), and these structures are consistent with the NTP-driven translocation model. These structures, however, indicate that the previously assumed route of NTP entry into the RNA polymerase active site, the secondary pore, is not the main route of NTP entry. The secondary pore presents a narrow channel with a highly negative electrostatic potential (Batada et al. 2004), so the pore does not appear to be a favorable NTP-loading route. Rather, the main entry port for NTP substrates must pass through the main enzyme channel. For a

Figure II-2. Three NTP binding sites in RNA polymerase II. (A and B) Two views of the yeast RNA polymerase II elongation complex structure (Gnatt et al. 2001; Westover et al. 2004*b*). (C and D) Simplified views of main channel (*i+2* and *i+3*) and active site (*i+1*) NTP-loading sites. (C) Stick figures with labeled amino acids thought to be involved in NTP binding and NTP-driven translocation. Fork loop 2 (Rpb2 region Asn499 to His515) is a light blue ribbon. (D) Space-filling representation, indicating the open space between the *i+2* and *i+1* sites. Minor movements of Rpb2 fork loop 2 (light blue; amino acids Rpb2 500–540 are shown) is necessary for NTP passage from *i+2* to *i+1*. DNA-template strand is red (from crystal structures) or pink (modeled, upstream DNA). The non-template strand is white (from crystal structures) or silver (modeled, bubble and upstream DNA). The modeled open complex is 18 bases. RNA is green. The bridge α-helix is a blue cylinder. The *i+1* (active site) (purple), *i+2* (yellow), and *i+3* (orange) NTPs are indicated. Rpb2 Arg504, Rpb2 Arg512, Rpb1 Lys1102, and Rpb1 Lys1112 (blue) are indicated. Rpb2 Asp505 and Rpb2 Glu529 (red) and Rpb1 Ala828 (brown) are shown.
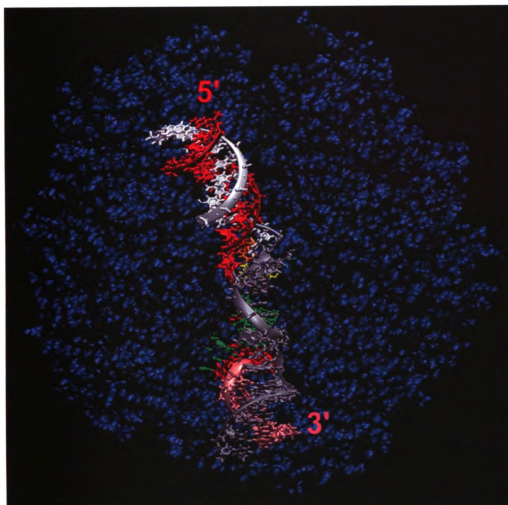
A

Figure II-2.

Figure II-2 (continued).
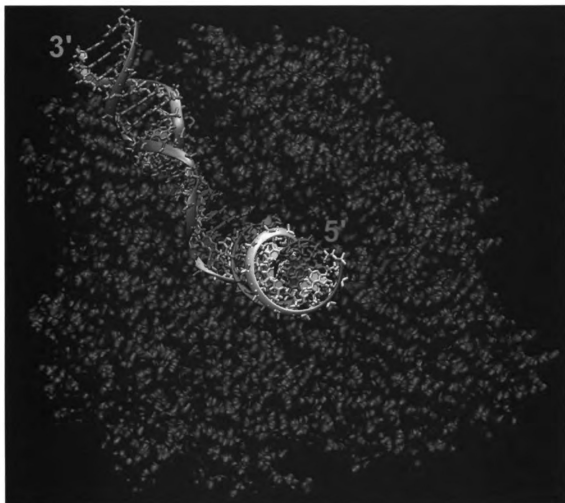
B

Figure II-2 (continued).
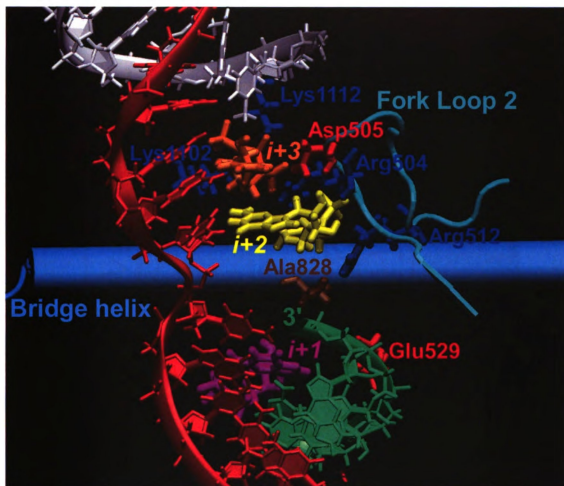
C

Figure II-2 (continued).
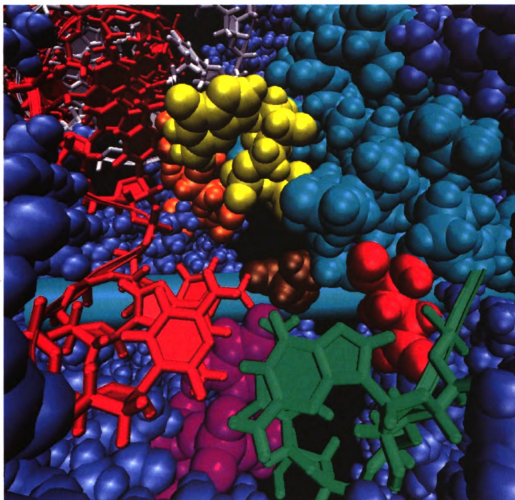
D

Figure II-3. NTP-driven translocation by human RNA polymerase II. (A) Pre-translocation elongation complex. Colors are the same as those indicated in **Figure II-2.** (B) Translocation intermediate. (C) Post-translocation elongation complex. Space is available to load an NTP (yellow) from the $i+2$ main-channel site to the $i+1$ active site. An active site, Mg^{2+} (white), is shown.
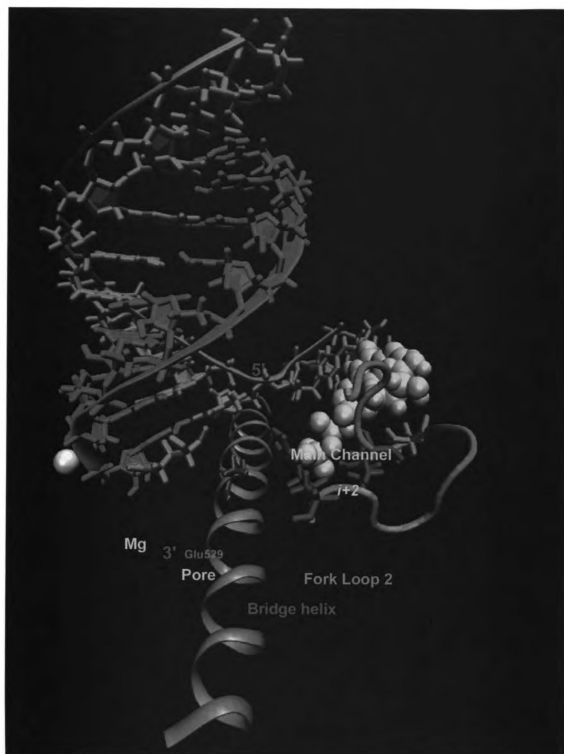
A

Figure II-3.

Figure II-3 (continued).

B

Figure II-3 (continued).

C
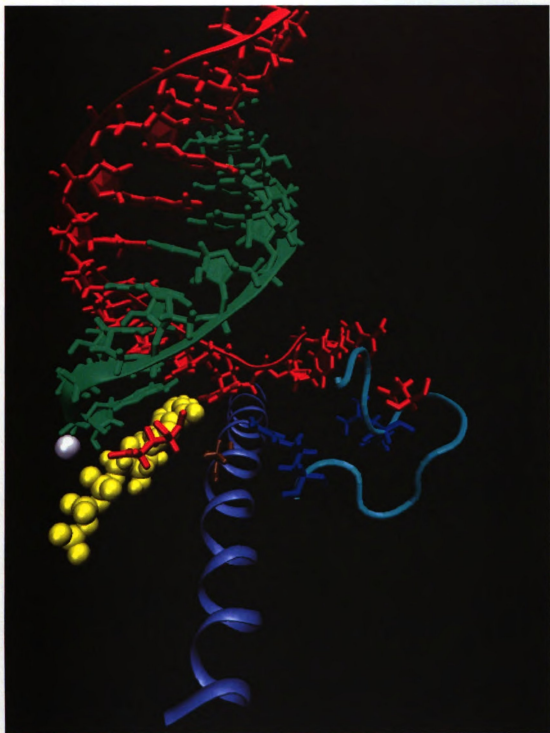
Figure II-4. α-amanitin blocks translocation. (A) Structural model of α-amanitin (pink) bound to the post-translocation elongation complex with 3 substrate NTPs bound. Colors are the same as those indicated in **Figure II-2**. (B) Structural comparison of α-amanitin (left panel) and Microcin J25 (right panel) (Bayro et al. 2003; Rosengren et al. 2003; Wilson et al. 2003; Gong et al. 2004). For (B) similar amino acids are yellow.
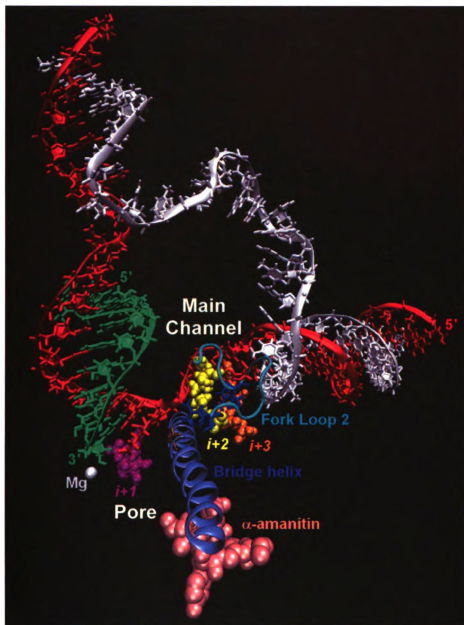
Figure II-4.

Figure II-4 (continued).

B

α-amanitin

hydroxy-Pro

hydroxy-Trp

Gly

Gly

Microcin J25

Pro

Gly

Gly

Tyr

99
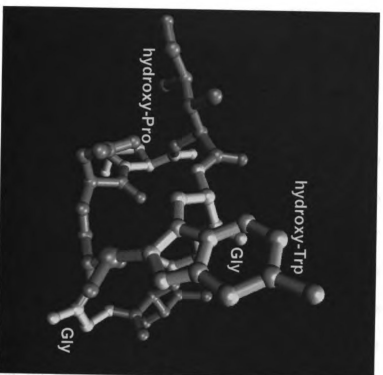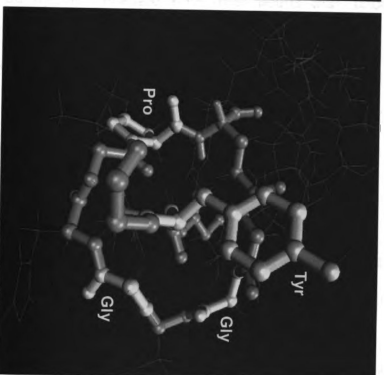
stalled elongation complex in the post-translocation state, the secondary pore is the likely route of NTP entry (Kettenberger et al. 2004; Westover et al. 2004a, 2004b). In the pre-translocation state, however, the DNA template base is single stranded and faces the main enzyme channel, not the secondary pore (Gnatt et al. 2001). To the pre-translocation state, therefore, the main enzyme channel is the likely route of NTP entry, and the secondary pore is excluded as a potential loading route. Kinetic studies demonstrate that the pre-translocation state of the elongation complex binds NTPs (Nedialkov et al. 2003a; Gong et al. 2004, 2005; Zhang and Burton 2004; Zhang et al. 2005). During processive synthesis, NTPs occupy the pre-translocation position, while the post-translocation position is occupied by the substrate NTP; this means that, under normal elongation conditions, NTPs enter the active site from the main enzyme channel. Only when the RNA polymerase stalls, releases pyrophosphate, and then translocates, can NTPs be loaded through the secondary pore, a situation that does not arise during processive synthesis (Figure II-1A). Sufficient space is available to load NTPs from the main channel into the RNA polymerase II active site (Figure II-2D).

What is the energetic driving force that propels the dNMP–NTP base pair to advance from the $i+2$ position to the $i+1$ active site (Figure II-3)? From observation of available elongation complex structures, there is no obvious answer to this question. We assume that the accurately paired NTP acts in a manner similar to an allosteric effector to induce a conformational change in RNA polymerase II that drives translocation of the RNA–DNA hybrid forward. Bending

of the bridge α-helix (Gnatt et al. 2001), rearrangement of the trigger loop (Vassylyev et al. 2002; Bar-Nahum et al. 2005), and movement of other mobile protein switches (Holmes and Erie 2003) may be involved in the NTP-driven translocation mechanism.

Available elongation complex structures of yeast RNA polymerase II present different views of the number of single-stranded DNA template bases available for NTP pairing in the main enzyme channel. A structure from the Kornberg laboratory indicates at least 3 main channel single-stranded template bases ($i+2$, $i+3$, and $i+4$; $i+1$ represents the position of an NTP in the active site poised for bond formation) (Gnatt et al. 2001). More recent structures from the same researchers have shown a similar position for bubble melting ($i+4$ is unpaired). The nontemplate DNA strand in these structures, however, lacked bases $i+2$ and $i+3$, inhibiting its capacity to anneal (Westover et al. 2004b). The structure from the Cramer laboratory shows that the downstream bubble can close, leaving (at most) a single base ($i+2$) unpaired in the main enzyme channel (Kettenberger et al. 2004). The Cramer structure is not informative about the unpaired $i+2$ base, because the $i+2$ base was selected to be unpaired. To support the NTP-driven translocation mechanism, a single unpaired DNA base ($i+2$) in the main-enzyme channel is minimally sufficient (Nedialkov et al. 2003a; Zhang and Burton 2004), and none of the available structures contradict this requirement.

Based on experiments with human RNA polymerase II, however, at least 3 DNA bases appear to be simultaneously available for pairing NTP

substrates (*i+1*, *i+2*, and *i+3*), (Gong et al. 2005; Figure II-2). This result is in apparent contradiction to the Cramer structure, but it appears to be consistent with the Kornberg structures. We therefore propose that the Cramer elongation complex suffers downstream bubble collapse. The Cramer elongation complex has a 7- rather than an 8-base-pair RNA–DNA hybrid (Kireeva et al. 2000; Gnatt et al. 2001), and the crystal is badly disordered upon addition of a substrate analogue (GMPcPP), which fails to load accurately into the catalytic site (Kettenberger et al. 2004). In the Cramer structure, the *i+2*, *i+3*, and *i+4* DNA base pairs are strained, as if they are likely to separate in a more favorable conformation of the elongation complex.

There is no clear demonstration that the Cramer structure is highly active, leading us to wonder whether the crystals would disorder upon addition of a natural GTP substrate, as they appear to do attempting to load the GMPcPP analogue. The Cramer elongation complex does not represent the only configuration of the downstream elongation bubble; the Kornberg structures, with 3 unpaired DNA bases in the main enzyme channel (*i+2*, *i+3*, and *i+4*) (Gnatt et al. 2001; Westover et al. 2004*b*), appear to be more accurate portrayals of a functional RNA polymerase II elongation complex. An NTP substrate loaded in the active site (*i+1*) senses the presence of at least 2 downstream templated NTP bases (*i+2* and *i+3*), using verified and functional human RNA polymerase II elongation complexes (Gong et al. 2005). Because ours is a kinetic experiment, we do not know the structural positions occupied by incoming NTPs, but we have modeled possible positions in Figures II-2C and 2D. Our experiments are done

with human enzyme, and available crystal structures are of the yeast enzyme. It is possible that there are differences in the regulation of the downstream transcription bubble in different multi-subunit RNA polymerases, although we do not believe that this is the case for RNA polymerase II from different species. We guess that all multi-subunit RNA polymerases use similar NTP-driven translocation mechanisms.

5. Evolutionary conservation of residues proposed to participate in NTP-driven translocation

Multi-subunit RNA polymerases are conserved in all three living kingdoms, eukarya, archaea, and eubacteria. Eukarya possess RNAPs I, II and III, encoded by the Rpa, Rpb and Rpc genes, respectively. Multiple sequence alignments show the regions that we propose to be important in the NTP-driven translocation mechanism are either invariant or highly conserved (data not shown). The sequence alignments were initially done using 3D-Coffee (O'Sullivan et al., 2004; Poirot et al., 2004) and then edited by hand.

The proposed functions of particular amino acids are listed in Table II-1. Rbp1 Ala828 is on the bridge α-helix. This residue is invariant in this alignment except for the highly divergent vaccinia virus RNAP for which this residue is valine. Rpb1 Lys1102 is conserved in most Rpbs, Rpcs and Rpas (as arginine). Rpb1 Lys1112 is conserved in all Rpbs and most Rpas. Rpb1 Lys1102 and Lys1112 are proposed to recognize the *i+3* NTP triphosphate, so some multi-

subunit RNAPs may use a different strategy to recognize the *i+3* NTP or may use a simpler version of the NTP-driven translocation mechanism (i.e. pre-loading only the *i+2* NTP). Rpb2 Arg504, Rpb2 Asp505, Rpb2 Arg512 and Rpb2 Glu529 are highly conserved among Rpbs. Arg504 is conserved except in Rpas. Asp505 is found in Rpbs and eubacteria. Asp505 is proposed to recognize the 2'- and 3'-hydroxyls of the ribose ring of the *i+2* NTP. RNA polymerases lacking a similar residue at this position must use a somewhat different strategy to discriminate ribose and deoxyribose sugars. Rpb2 Arg512 is invariant in this alignment. Because of the position of this buried arginine, it is difficult to account for its high conservation without considering the NTP-driven translocation mechanism. Arg512 is not positioned to interact with DNA, RNA or an NTP at the active site. Rpb2 Glu529 is also very highly conserved throughout evolution. Based on the high conservation of the fork loop 2 region and particular amino acids, we propose that most or all multi-subunit RNA polymerases utilize NTP-driven translocation and pre-loading of the *i+2* NTP.

Table II-1. Yeast RNAP II residues proposed to be important in NTP-driven translocation.

Subunit	Amino Acid	Proposed Function
Rpb1	Ala828 (Bridge Helix)	Loading of *i+2* NTP to the *i+1* active site
Rpb1	Lys1102 (Helix 37)	Recognition of the *i+3* NTP triphosphate
Rpb1	Lys1112	Recognition of the *i+3* NTP triphosphate
Rpb2	Arg504 (Fork Loop 2)	Recognition of the *i+2* NTP triphosphate
Rpb2	Asp505 (Fork Loop 2)	Recognition of the 2'- and 3'- OH of the *i+2* NTP ribose sugar
Rpb2	Arg512 (Fork Loop 2)	Recognition of the *i+2* NTP triphosphate
Rpb2	Glu529	Orientation of the triphosphate of the incoming NTP substrate as it enters the *i+1* active site

6. A structural model for NTP-driven translocation

Based on our work and available crystal-structure x-rays, we present a detailed model to describe the preloading of NTP substrates in the RNA polymerase II main channel prior to translocation (Figure II-2). The non-template strand is disordered in crystal structures, so its position was modeled in the images (Figures II-2A, 2B). Our model posits 2 NTP- Mg^{2+} substrates ($i+2$ and $i+3$) paired with cognate DNA bases in the main channel; the active site is occupied with another NTP substrate ($i+1$). We do not believe a third NTP can be accommodated in the main channel because, at the $i+4$ position, an NTP would clash with the non-template DNA strand. We posit that main channel NTP substrates help to keep the downstream transcription bubble open, explaining why the bubble can collapse in the absence of appropriate incoming substrates (Kettenberger et al. 2004). For the post-translocated elongation complex, in the presence of substrates, we suggest bubble opening at the $i+2$, $i+3$ and $i+4$ positions, as indicated in the Kornberg structure (Gnatt et al. 2001). To create our model, we placed some amino acids that were disordered in the Kornberg structures (yeast Rpb2 fork loop 2 region, residues Gly503-Leu508) (Figures II-2C, 2D). The $i+2$ and $i+3$ NTPs pair to their main channel template sites. A binding site forms around the NTPs, bordered by Rpb2 subunit fork loop 2 protein side chains. Available amino acids project the expected basic groups (Rpb2 Arg504, Rpb2 Arg512, Rpb1 Lys1102, and Rpb1 Lys1112) and 1 acidic group (Rpb2 Asp505). It is proposed, therefore, that the binding of NTPs in the

main channel sculpts and maintains the downstream edge of the transcription bubble.

This model provides a mechanism by which RNA polymerase II can anticipate upcoming transcriptional stalls. As RNA polymerase II approaches a stall, we have observed that the elongation complex enters the paused state more rapidly than can be accommodated by rate constants into and out of the paused state, indicating that RNA polymerase II senses an upcoming stall before the active site is deprived of substrate (Zhang et al. 2003). According to our current model, the downstream DNA template, Rpb2 fork loop 2, and the nontemplate DNA strand encompass the *i+2* and *i+3* NTP- Mg^{2+} substrates to regulate the opening of the downstream edge of the transcription bubble. When NTP substrates are not available, the downstream bubble tends to collapse, leading to enhanced transcriptional pausing. The binding of elongation factors is also expected to influence the extent of downstream bubble opening and, hence, the frequency of pausing.

As described above, the NTP-Mg^{2+} binding site in the main RNA polymerase II channel is expected to screen incoming NTPs for accurate base pairing, for triphosphate tails, and for 2'- and 3'-hydroxyls on the ribose ring. Our model allows for templating at main channel sites and explains why, during NTP deprivation, the transcription bubble can collapse, as observed in the Cramer structure (Kettenberger et al. 2004). Deprived of the *i+2* and *i+3* NTPs, RNA polymerase II might close the downstream bubble as it approaches the stall. We posit that yeast RNA polymerase II amino acids Rpb2 Arg504 and Arg512

(contacting the *i+2* NTP triphosphate) and Rpb1 Lys1102 and Lys1112

(contacting the *i+3* NTP triphosphate) are involved in screening triphosphate

tails. Rpb2 Asp505 is positioned to scan the 2'- 3'-cis diol of the *i+2* ribose ring.

We suggest that this interaction may be mediated by a mobile Mg^{2+} atom. For T7

RNA polymerase, a Mg^{2+} is used in a similar manner to prescreen the 2'-hydroxyl

of the incoming NTP substrate in the pre-insertion site (Temiakov et al. 2004).

For RNA polymerase II, neither the NTP- Mg^{2+} binding pockets at the *i+2* or *i+3*

positions of the transcription bubble are stable binding sites. NTPs are not

expected to loiter in the main channel, but rather to rapidly dock and then

advance toward the active site, paired with their cognate DNA base. As the DNA

template translocates, another NTP- Mg^{2+} occupies the *i+3* position. (See

supplementary Figures II-1A–1C and supplementary Table I for a summary of

predictions based on our model and for evolutionary conservation of relevant

protein regions.)

As the *i+2* NTP- Mg^{2+} is transferred to the *i+1* active site during

translocation, other RNA polymerase II residues become important (Figures II-

2C, 2D, and 3). Bridge helix Rpb1 Ala828 is encountered during NTP passage

past the bridge α-helix (Nedialkov et al. 2003*a*). Rpb2 Glu529 is expected to help

orient the tri-phosphate, through charge repulsion, during its passage into the

active site. Residues on the bridge α-helix and in the underlying trigger loop (Bar-

Nahum et al. 2005) may be involved in the translocation mechanism. Amino

acids interacting with the RNA–DNA hybrid are also expected to be important,

because mobilizing the hybrid likely facilitates translocation (Holmes and Erie 2003).

7. Mushrooms and antimicrobials

A deadly mushroom toxin, α-amanitin, blocks translocation by human RNA polymerase II (Bushnell et al. 2002; Gong et al. 2004) (Figure II-4). α-amanitin is structurally similar to a bacterial antibiotic Microcin J25. α-amanitin is a bicyclic, covalently crosslinked, and multiply modified octapeptide (Figure II-4B, left panel). Microcin J25 (with 21 amino acids in total) consists of a cyclic octapeptide (amino acids 1–8) and peptide tail (amino acids 9–21), the C-terminus of which can enter the octapeptide ring to form a bridge reminiscent of the α-amanitin covalent crossbridge (Bayro et al. 2003; Rosengren et al. 2003; Wilson et al. 2003) (Figure II-4B, right panel). α-amanitin is the most potent and specific known inhibitor of human RNA polymerase II, and Microcin J25 inhibits homologous bacterial RNA polymerases (Yuzenkova et al. 2002). Both toxins penetrate the secondary pore, and α-amanitin interacts strongly and specifically with the bridge α-helix (Bushnell et al. 2002) (Figure II-4A), thought to be important in translocation mechanisms (Gnatt et al. 2001; Vassylyev et al. 2002; Bar-Nahum et al. 2005). Similar to its structural analogue α-amanitin, Microcin J25 most likely inhibits translocation. Suggestions that Microcin J25 blocks NTP loading through the secondary pore to inhibit transcription (Adelman et al. 2004; Mukhopadhyay et al. 2004) are not likely correct. Certainly, the cork-in-bottle

model for Microcin J25 action is contrary to the NTP-driven translocation model, which posits an alternate route of NTP entry. This issue will be resolved when transient-state kinetic analyses of Microcin J25 inhibition are reported. Single molecule elongation studies have lacked sufficient resolution to discriminate between an inhibitor that blocks substrate binding and one that blocks translocation (Adelman et al. 2004). As a potent translocation inhibitor, α-amanitin has proven to be an incisive probe with which to unravel the human RNA polymerase II elongation mechanism (Gong et al. 2004, 2005).

8. Fidelity and efficiency

Why did NTP-driven translocation develop as the dominant mechanism for substrate NTP loading into the active sites of multi-subunit RNA polymerases? The answer relates to the requirement for high fidelity and efficiency of RNA synthesis (Nedialkov et al. 2003a; Zhang and Burton 2004). There lingers a textbook ethic that RNA polymerases lack the separate exonuclease activities of DNA polymerases, because RNA is a throw-away copy of the genetic material, and, therefore, errors in transcription are tolerated. Errors in DNA replication, in contrast, are thought to have a more severe consequence, because lesions can be fixed into genetic material as heritable mutations. Based on kinetic studies of elongation, however, it appears that many RNA polymerases are highly accurate enzymes, perhaps matching the fidelity of DNA polymerases,

at least during ongoing synthesis. In that case, how do multi-subunit RNA polymerases maintain high fidelity and efficiency during elongation?

For both DNA and RNA polymerases, high fidelity must be achieved with relatively low-affinity substrate binding, because these enzymes must recognize 4 substrates with different shapes and chemical groups. Accurate base-pairing must provide a primary initial screening of the incoming (d)NTP, but it is also essential to screen for or against the 2'- OH, so that DNA polymerases can exclude ribo-NTPs, and RNA polymerases can exclude deoxy-NTPs. Exclusion of mono- and diphosphates is also important. Recent structural data of simple RNA and DNA polymerases and kinetic studies of multi-subunit RNA polymerases indicate that the screening of substrates begins outside the active site of the enzyme, and that a series of checkpoints must be passed before the substrate tightens into the active site for phosphoryl transfer. Pre-assessment of the NTP substrate largely precludes replication and transcription errors, with low-energy cost, so DNA and RNA polymerases can achieve high fidelity with relatively low-affinity substrate recognition.

For simple RNA and DNA polymerases (i.e., T7 RNA polymerase and *Bacillus stearothermophilus* large-fragment DNA polymerase I), template pre-insertion, (d)NTP pre-insertion, and insertion (active site) positions of the dNMP template have been defined (Johnson et al. 2003; Johnson and Beese 2004; Temiakov et al. 2004; Yin and Steitz 2004) (Figure II-1B). Substrate (d)NTPs are known to associate with the (d)NTP pre-insertion and insertion (active site) positions. Based on the NTP-driven translocation model for human RNA

112

polymerase II, we suggest that accurate (d)NTP binding may be the route favored to drive the elongation complex between the template pre-insertion (pre-translocated) and the (d)NTP pre-insertion (post-translocated) sites for some simple polymerases. If this supposition is correct, then simple RNA and DNA polymerases may use a (d)NTP-driven translocation mechanism similar to that used by multi-subunit RNA polymerases. This mechanism would allow 2 low-affinity pre-assessments of the substrate (d)NTP prior to loading it into the active site: 1 at the template pre-insertion site (Johnson et al. 2003; Johnson and Beese 2004); and 1 at the (d)NTP pre-insertion site (Temiakov et al. 2004). Transfer from the (d)NTP pre-insertion site to the active site requires a protein conformational change before potential (d)NMP mis-incorporation (Temiakov et al. 2004). Pre-screening is an effective fidelity check, because a (d)NTP that is rejected outside the active site cannot result in (d)NMP mis-incorporation. The template pre-insertion site for DNA polymerases is located on the opposite side of the O α- helix (Johnson et al. 2003; Johnson and Beese 2004), far from the active site, so initial pre-assessment of an incoming dNTP might occur far from the active site. In multi-subunit RNA polymerases, initial NTP pre-assessment occurs on the opposite side of the bridge α-helix (Nedialkov et al. 2003a; Zhang and Burton 2004; Zhang et al. 2005) (Figure II-2C), far from the active site, suggesting that NTP loading and assessment strategies may be similar for non-homologous simple and multi-subunit enzymes. T7 RNA polymerase opens the downstream transcription bubble at the N-terminal tip of the O'-α-helix (at Phe644). For T7 RNA polymerase, therefore, only a single template base is

available to pair with an NTP substrate. To use NTP-driven translocation requires that an NTP pair with template as soon as its DNA base becomes single stranded, during DNA passage past Phe644 and the O'-α-helix. T7 RNA polymerase has been shown to load NTPs to an NTP pre-insertion site prior to a conformational change (open to closed transition involving rotation of the O-α-helix finger domain) that brings the substrate NTP into the insertion (active) site (Temiakov et al. 2004).

Kinetic analyses of RNA-dependent RNA polymerase of poliovirus (in the presence of Mn^{2+}) (Arnold and Cameron 2004; Arnold et al. 2004; Castro et al. 2005) are similar to those for human RNA polymerase II (in the presence of Mg^{2+}) (Nedialkov et al. 2003a; Zhang and Burton 2004). Poliovirus RNA polymerase is a homologue of simple RNA and DNA polymerases. This result leads us to believe that simple enzymes, like multi-subunit enzymes, may use (d)NTP-driven translocation mechanisms.

For multi-subunit RNA polymerases, the NTP substrates are prescreened in the main enzyme channel, which is separated from the active site by the bridge α-helix (Nedialkov et al. 2003a; Zhang and Burton 2004) (Figures II-2C, 2D). The incoming NTP is first screened by base-pairing to template at the *i+3* position. It is screened again at the *i+2* position. We believe that the 2'-OH of the ribose ring is assessed at the *i+2* position by Rpb2 Asp505, so dNTPs are normally rejected well outside the RNA polymerase II active site. The space from the main enzyme channel to the active site, which the substrate NTP must traverse in transition from the pre- to the post-translocation position, is

constrained, so only accurately paired NTPs can successfully complete the passage (Nedialkov et al. 2003a) (Figures II-2C, 2D, and 3). Once beyond the bridge α-helix, the incoming dNMP–NTP base pair is again challenged during an isomerization step from an open to closed active site (Zhang and Burton 2004; Zhang et al. 2005). This is an induced-fit step, similar to the conversion from the pre-insertion (open conformation) to the insertion (closed conformation) site for simple RNA and DNA polymerases (Johnson et al. 2003; Johnson and Beese 2004; Temiakov et al. 2004; Yin and Steitz 2004). In kinetic studies, the open to closed transition is indicated by sequestration of active-site metal ions from EDTA chelation (Zhang and Burton 2004; Zhang et al. 2005). For human RNA polymerase II, the apparently slow bond synthesis step may represent a second induced-fit fidelity check within the active site. If the tightened NTP-Mg^{2+} cannot be appropriately deformed to the catalytically competent state, this would signal an error in substrate loading, inducing reversal of isomerization and NTP release. According to the NTP-driven translocation model, pyrophosphate release (bond completion) is coupled to the next NTP-driven translocation (the following bond initiation). If translocation is blocked (through NMP mis-incorporation), bond synthesis might be reversed and the inaccurately loaded NTP expelled, providing a fidelity check even after phosphodiester bond formation but prior to pyrophosphate release. Thus, simple DNA and RNA polymerases and multi-subunit RNA polymerases use a complex pathway of (d)NTP prescreening with multiple fidelity checks and an induced-fit step (or steps) prior to bond formation. These enzymes may run a final fidelity check after bond formation but before

pyrophosphate release. Release of pyrophosphate results in (essentially) irreversible bond completion.

NTP-driven translocation is efficient because it allows for accurate pre-alignment, prescreening, and presorting of NTP substrates (Nedialkov et al. 2003a; Zhang and Burton 2004). Furthermore, the end of one bond-addition cycle (pyrophosphate release) is coupled to the beginning of the next (translocation and NTP loading) (Zhang and Burton 2004; Zhang et al. 2005). NTPs load through the main channel. In the NTP-driven translocation model, the secondary pore is the route of pyrophosphate release. Improperly loaded and rejected NTP (or dNTP) substrates also exit through the secondary pore in cases in which they pass prescreening checkpoints. Only when the elongation complex is stalled and ratcheted to the post-translocated position can the secondary pore be a route for substrate NTP loading.

9. Conclusions

In summary, a new view of translocation mechanisms has been developed, based on kinetic analyses of elongation by human RNA polymerase II. NTP-driven translocation is consistent with available kinetic and structural data, although, in the future, inconsistencies with the Cramer elongation complex structure will have to be resolved. NTP-driven translocation posits that the RNA polymerase secondary pore is not the only, nor is it the major, route of NTP entry

into the elongation complex. NTPs are loaded first in the enzyme main channel. NTP-driven translocation provides a new understanding of fidelity and efficiency, indicating that RNA synthesis is a higher-fidelity reaction than was previously thought. Simple RNA and DNA polymerases may use a shorthand version of (d)NTP-driven translocation, providing multiple fidelity checks of the incoming substrate at a low-energy cost.

Inspection of available RNA polymerase II elongation complex structures suggests that the opening and closing of the downstream transcription bubble is regulated during elongation. We suggest that downstream bubble opening is maintained through NTP- Mg^{2+} loading, and is regulated by transcriptional elongation factors. Downstream bubble collapse is a signal for RNA polymerase to pause transcription. Maintenance of the open bubble configuration is the signal for continued rapid synthesis, and is supported by the weak binding of templated NTP- Mg^{2+} substrates within the main RNA polymerase channel.

Disparate models for elongation will be reconciled using a combination of structural studies, transient-state kinetic analyses, and mutagenic studies. Understanding the coupled mechanisms of translocation and NTP- Mg^{2+} pre-screening is important to the understanding of the fidelity and efficiency of multi-subunit RNA polymerases.

References

Adelman, K., Yuzenkova, J., La Porta, A., Zenkin, N., Lee, J., Lis, J.T. et al. 2004. Molecular mechanism of transcription inhibition by peptide antibiotic Microcin J25. Mol. Cell, **14**: 753–762.

Arnold, J.J., and Cameron, C.E. 2004. Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg2+. Biochemistry, **43**: 5126–5137.

Arnold, J.J., Gohara, D.W., and Cameron, C.E. 2004. Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn2+. Biochemistry, **43**: 5138–5148.

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**: 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**: 17361–17364.

Bayro, M.J., Mukhopadhyay, J., Swapna, G.V., Huang, J.Y., Ma, L.C., Sineva, E. et al. 2003. Structure of antibacterial peptide Microcin J25: a 21-residue lariat protoknot. J. Am. Chem. Soc. **125**: 12382–12383.

Bushnell, D.A., Cramer, P., and Kornberg, R.D. 2002. Structural basis of transcription: α-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. Proc. Natl. Acad. Sci. U.S.A. **99**: 1218–1222.

Castro, C., Arnold, J.J., and Cameron, C.E. 2005. Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. Virus Res. **107**: 141–149.

Foster, J.E., Holmes, S.F., and Erie, D.A. 2001. Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation. Cell, **106**: 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 3.3 Å resolution. Science (Wash. D.C.), **292**: 1876–1882.

Gong, X.Q., Nedialkov, Y.A., and Burton, Z.F. 2004. α-amanitin blocks translocation by human RNA polymerase II. J. Biol. Chem. **279**: 27422–27427.

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: 461–470.

Guajardo, R., and Sousa, R. 1997. A model for the mechanism of polymerase translocation. J. Mol. Biol. 265: 8–19.

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: 35597–35608.

Johnson, K.A. 1992. Transient-state kinetic analysis of enzyme reaction pathways. Enzymes, 20: 1–61.

Johnson, K.A. 1995. Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. Methods Enzymol. 249: 38–61.

Johnson, S.J., and Beese, L.S. 2004. Structures of mismatch replication errors observed in a DNA polymerase. Cell, 116: 803–816.

Johnson, S.J., Taylor, J.S., and Beese, L.S. 2003. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. Proc. Natl. Acad. Sci. U.S.A. 100: 3895–3900.

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: 955–965.

Kireeva, M.L., Komissarova, N., Waugh, D.S., and Kashlev, M. 2000. The 8-nucleotide-long RNA: DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J. Biol. Chem. 275: 6530–6536.

Landick, R. 2004. Active-site dynamics in RNA polymerases. Cell, 116: 351–353.

Mukhopadhyay, J., Sineva, E., Knight, J., Levy, R.M., and Ebright, R.H. 2004. Antibacterial peptide Microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. Mol. Cell, 14: 739–751.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H. et al. 2003a. NTP-driven translocation by human RNA polymerase II. J. Biol. Chem. 278: 18303–18312.

Nedialkov, Y.A., Gong, X.Q., Yamaguchi, Y., Handa, H., and Burton, Z.F. 2003b. Transient state kinetics of RNA polymerase II elongation. Methods Enzymol. 371: 252–262.

Oster, G. 2002. Brownian ratchets: Darwin's motors. Nature (London), **417**: 25.

Patel, S.S., Wong, I., and Johnson, K.A. 1991. Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. Biochemistry, **30**: 511–525.

O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D. G., and Notredame, C. (2004). 3DCoffee: combining protein sequences and structures within multiple sequence alignments. J Mol Biol *340*, 385-395.

Poirot, O., Suhre, K., Abergel, C., O'Toole, E., and Notredame, C. (2004). 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. Nucleic Acids Res *32*, W37-40.

Rosengren, K.J., Clark, R.J., Daly, N.L., Goransson, U., Jones, A., and Craik, D.J. 2003. Microcin J25 has a threaded sidechain-tobackbone ring structure and not a head-to-tail cyclized backbone. J. Am. Chem. Soc. **125**: 12464–12474.

Sousa, R. 2003. On models and methods for studying polymerase translocation. Methods Enzymol. **371**: 3–13.

Sousa, R. 2005. Machinations of a maxwellian demon. Cell, **120**: 155–156.

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**: 381–391.

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 2.6 A resolution. Nature (London), **417**: 712–719.

Wang, H., and Oster, G. 2002. Ratchets, power strokes, and molecular motors. Appl. Phys. A **75**: 315–323.

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

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

Wilson, K.A., Kalkum, M., Ottesen, J., Yuzenkova, J., Chait, B.T., Landick, R. et al. 2003. Structure of Microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. J. Am. Chem. Soc. **125**: 12475–12483.

Yin, Y.W., and Steitz, T.A. 2002. Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science (Wash. D.C.), **298**: 1387–1395.

Yin, Y.W., and Steitz, T.A. 2004. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. Cell, **116**: 393–404.

Yuzenkova, J., Delgado, M., Nechaev, S., Savalia, D., Epshtein, V., Artsimovitch, I. et al. 2002. Mutations of bacterial RNA polymerase leading to resistance to Microcin J25. J. Biol. Chem. **277**: 50867–50875.

Zhang, C., and Burton, Z.F. 2004. Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. J. Mol. Biol. **342**: 1085–1099.

Zhang, C., Yan, H., and Burton, Z.F. 2003. Combinatorial control of human RNA polymerase II (RNAP II) pausing and transcript cleavage by transcription factor IIF, hepatitis δ antigen, and stimulatory factor II. J. Biol. Chem. **278**: 50101–50111.

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 NTP-driven translocation. Mol. Cell. Biol. **25**: 3583–3595.

CHAPTER THREE

Xiong, Y., and Burton, Z. F. (2007). A tunable ratchet driving human RNA polymerase II translocation adjusted by accurately templated nucleoside triphosphates loaded at downstream sites and by elongation factors. J Biol Chem *282*, 36582-36592. [Originally published In Press as doi:10.1074/jbc.M707014200 on September 17, 2007]

CHAPTER THREE

A TUNABLE RATCHET DRIVING HUMAN RNA POLYMERASE II TRANSLOCATION ADJUSTED BY ACCURATELY TEMPLATED NUCLEOSIDE TRIPHOSPHATES LOADED AT DOWNSTREAM SITES AND BY ELONGATION FACTORS

Abstract

When nucleoside triphosphate (NTP) substrates and α-amanitin are added to a human RNA polymerase II elongation complex simultaneously, the reaction becomes stalled in the core of the bond synthesis mechanism. The mode of stalling is influenced by NTP substrates at the active site and at downstream sites, and by Transcription Factor IIF (TFIIF) and TFIIS. NTP substrates templated at i+2, i+3, and i+4 downstream DNA sites can reverse the previously stable binding of an NTP loaded at the i+1 substrate site. Deoxy-(d)NTPs and NDPs (nucleoside diphosphates) do not substitute for NTPs at the i+2 and i+3 positions (considered together) or the i+4, i+5, and i+6 positions (considered together). The mode of stalling is altered by changing the number of downstream template sites that are accurately occupied by NTPs and by changing NTP concentration. In the presence of the translocation blocker α-

amanitin, a steady state condition is established in which RNA polymerase II stably loads an NTP substrate at i+1 and forms a phosphodiester bond but cannot rapidly complete bond synthesis by releasing pyrophosphate. These observations support a role for incoming NTP substrates in stimulating translocation; results appear inconsistent with the secondary pore being the sole route of NTP entry for human RNA polymerase II; and results indicate mechanisms of dynamic error avoidance and error correction during rapid RNA synthesis.

1. Introduction

A goal of transient state kinetic analyses is to reveal the internal mechanism of an enzyme reaction by observing synchronized millisecond events on a millisecond time scale. In this work, the mushroom toxin α-amanitin is utilized as a transient state inhibitor, locking human RNA polymerase II in the core of the elongation mechanism. By altering the conditions of stalling and by employing two distinct reaction quenching methods, details of the core mechanism of human RNA polymerase II become apparent.

Recent x-ray crystal structures of *T. thermophilus* RNA polymerase elongation complexes indicate a simple thermal ratchet mechanism for elongation (Vassylyev et al., 2007; Vassylyev et al., 2007), with NTP substrates loading through the secondary pore, a solvent accessible channel, to the active site (designated i+1). Very little space is available in these structures to load NTPs through the main enzyme channel, which holds the DNA duplex and RNA-DNA hybrid. The downstream transcription bubble is closed at the i+2 position by base-pairing, so it is difficult to imagine how NTP substrates could interact at i+2 or i+3 template sites. *T. thermophilus* β R422 makes a specific contact to the i+1 DNA template phosphate, and appears to provide a specific mechanism for closing the downstream bubble at the i+2 position. This mechanism for closing the downstream bubble is not conserved in human or yeast RNA polymerase II,

in which the corresponding residue to β R422 is Rpb2 G493 (human) and Rpb2

G506 (yeast). Furthermore, even if the downstream bubble were open in the *T. thermophilus* structure, little space is available to flip a dNMP-NTP base pair from the i+2 to the i+1 active site. *T. thermophilus* structures further indicate that a closed conformation of the "trigger loop-trigger helix" assembly may support the catalytic intermediate, during phosphodiester bond formation.

According to a thermal ratchet mechanism, translocation and pyrophosphate release should be rapid and spontaneous, and neither process can be driven by an incoming, templated NTP substrate. In this paper, however, we present data that appears inconsistent with a simple thermal ratchet mechanism for elongation catalyzed by human RNA polymerase II. Furthermore, the data are most consistent with an open configuration of the trigger loop supporting phosphodiester bond addition. Data indicate that the secondary pore is unlikely to be the sole route of NTP entry into human RNA polymerase II. These experiments indicate a ratchet driving forward translocation that is regulated by incoming NTP substrates and transcriptional elongation factors. Elongation catalyzed by human RNA polymerase II, therefore, indicates features that are not apparent in images of the *T. thermophilus* elongation complex.

Two models have been invoked to describe NTP loading by multi-subunit RNA polymerases. The secondary pore NTP loading hypothesis posits that NTPs load only to the i+1 active site and that loading to i+2, i+3, and i+4 downstream sites is impossible (Bar-Nahum et al., 2005; Batada et al., 2004; Cramer et al., 2001; Gnatt et al., 2001; Landick, 2005; Temiakov et al., 2005;

Zhang et al., 1999). According to the secondary pore NTP loading hypothesis, pyrophosphate release and translocation are expected to be rapid and spontaneous, and NTP loading is likely to be rate-limiting (Batada et al., 2004). The NTP-driven translocation hypothesis, by contrast, posits NTP loading at the i+1 and i+2 sites and potentially at sites further downstream (Burton et al., 2005; Gong et al., 2004; Gong et al., 2005; Nedialkov et al., 2003; Zhang and Burton, 2004; Zhang et al., 2005). The secondary pore is a deep and narrow channel extending from the "funnel" on the surface of the RNA polymerase II molecule to the deeply buried active site (Batada et al., 2004; Landick, 2005). The pore appears to be too narrow to exchange two NTPs, so, misloading of an NTP substrate (the usual case according to the secondary pore NTP loading hypothesis), requires release of the incorrect NTP followed by loading of the correct (or another incorrect) NTP (or NDP or other small molecule). In addition to being about 15 angstroms deep and having a minimum diameter of about 7 angstroms (NTPs have a minimum diameter of about 6 angstroms), the secondary pore has significant negative electrostatic potential (Batada et al., 2004). Because of the negative electrostatics of the pore, Kornberg and colleagues estimated that NTP loading might be rate-limiting (20 to 30 s^{-1}) for NTPs loaded through the pore, even at physiological NTP concentrations: i.e. 3.1 mM ATP, 0.47 mM GTP, 0.28 mM CTP, and 0.57 mM UTP (average values for mammalian cells) (Traut, 1994). NTP loading, however, is not rate-limiting for human RNA polymerase II. Stable NTP loading has been determined to be 1450 +/- 330 s^{-1} (Zhang and Burton, 2004). The secondary pore NTP loading

hypothesis requires substantial competition between templated and non-templated NTPs at i+1, but, for the most part, such competition is not observed.

The secondary pore NTP loading hypothesis does not appear to be consistent with millisecond phase kinetic studies of RNA synthesis by human RNA polymerase II. Kinetic analyses show that RNA polymerase II has two rate-limiting steps in each bond addition cycle, and neither slow step corresponds to NTP loading (Burton et al., 2005; Gong et al., 2004; Gong et al., 2005; Nedialkov et al., 2003; Zhang and Burton, 2004; Zhang et al., 2005). One rate-limiting step appears to include a conformational change associated with phosphodiester bond formation. This step occurs between rapid and stable NTP binding and slow chemistry. The other rate-limiting step occurs within the interval between phosphodiester bond synthesis and the next stable NTP loading. This step includes pyrophosphate release, translocation, and NTP-Mg^{2+} loading. Because stable NTP-Mg^{2+} loading can be very rapid, the rate-limiting component of this interval has been interpreted as translocation coupled to pyrophosphate release. As indicated above, the *T. thermophilus* RNA polymerase elongation complex structures appear most consistent with the secondary pore NTP loading model for bacterial RNA polymerase (Vassylyev et al., 2007; Vassylyev et al., 2007).

The NTP-driven translocation hypothesis requires NTP loading to downstream template sites, and, therefore, posits an unpaired downstream bubble. Based on available x-ray crystal structure data, however, the extent of opening of the downstream bubble remains controversial (Gnatt et al., 2001; Kettenberger et al., 2004; Vassylyev et al., 2007; Vassylyev et al., 2007; Wang et

al., 2006; Westover et al., 2004). All available yeast RNA polymerase II elongation complex structures are open in the downstream region. Some of these structures are open to the i+5 or i+6 downstream positions (Wang et al., 2006), and base pairs that have the capacity to form in the downstream region are not observed to form (Gnatt et al., 2001; Wang et al., 2006; Westover et al., 2004). A *T. thermophilus* RNA polymerase elongation complex was observed to close in the downstream region (closed at i+2), although a key residue for closure, β R422 (corresponding to *S. cerevisiae* Rpb2 G506), is not conserved in RNA polymerase II (Vassylyev et al., 2007). If RNA polymerase II has the capacity to close the downstream bubble, therefore, RNA polymerase II uses different contacts for closure than *T. thermophilus* RNA polymerase. Some chemical reactivity studies support a model of variable downstream template opening for multi-subunit RNA polymerases (Kahl et al., 2000; Zaychikov et al., 1995), and functional studies indicate that the downstream bubble of human RNA polymerase II can interact with NTP substrates (Gong et al., 2005).

According to the NTP-driven translocation model, an NTP loads to the active site by transfer as a dNMP-NTP base pair from the i+2 to the i+1 site. NTP-driven translocation appears to be coupled to pyrophosphate release, linking initiation of each new bond addition cycle to the end of the previous cycle. According to this model, incorrect NTPs rarely load to the active site because NTPs are pre-sorted on the DNA template prior to active site loading, enhancing fidelity. In this model, the secondary pore is the route of pyrophosphate release and may be a route for expulsion of NTP substrates rejected at the active site.

Furthermore, NTP-driven translocation can be utilized as a means to dislodge an NTP from the active site in response to a translocation block, indicating a role for downstream NTPs in dynamic error avoidance and error correction (Burton et al., 2005; Gong et al., 2005). Normally, a translocation block signals a transcription error.

In this paper, we show templated effects of i+2, i+3, and i+4 NTPs on the fate of an NTP substrate loaded at i+1. This result demonstrates the major tenet of the NTP-driven translocation model: that NTPs load to downstream template sites while the active site is occupied. It appears from these studies that translocation pressure can be tuned by limiting the number of downstream sites that are accurately filled by NTPs and by altering the concentrations of accurately templated downstream NTPs. Elongation factors TFIIF and TFIIS also appear to adjust translocation pressure. We present a model for a tunable ratchet driving translocation that responds to accurately templated downstream NTPs and elongation factors, and that couples NTP-driven translocation to NTP substrate tightening at the i+1 active site.

2. Materials and Methods

2.1 Cell culture, extracts and proteins

HeLa cells were purchased from the National Cell Culture Center (Minneapolis, MN). Extracts of HeLa cell nuclei were prepared as described (Shapiro et al., 1988). Recombinant TFIIF was prepared as described (Wang et

al., 1993; Wang et al., 1994). Recombinant human TFIIS was purified by phosphocellulose chromatography followed by MonoS chromatography (our unpublished procedure).

2.2 NTP stocks

Ultra pure NTP sets were purchased from Amersham Pharmacia Inc. Based on our experience using these reagents, CTP and UTP stocks appear to be free of detectable ATP and GTP contamination. The ATP stock is substantially free of GTP contamination. The GTP preparation appears to be lightly contaminated with ATP, which does not complicate our experiment, because GTP is the final addition to the reaction, after prior addition of ATP. In reactions containing TFIIF, we observe some tendency to incorporate AMP or GMP, when none has been deliberately added to the reaction. As we show here, this observation results from stimulation of RNAP II elongation by TFIIF. The probable major source of ATP and GTP contamination in these reactions is the HeLa extract from which we initially derive elongation complexes (ECs) and not the NTP stocks. When TFIIS is added to the reaction in addition to TFIIF, inappropriate AMP and GMP incorporation is suppressed, indicating that the TFIIF preparation is not likely to be a source of ATP or GTP contamination.

2.3 Preparation of Elongation Complexes

^{32}P-labeled C40 (40 nucleotide RNA ending in a 3'-CMP) RNAP II ECs were formed on MagneSphere (Promega, Madison, WI) metal bead-immobilized

templates. The adenovirus major late promoter DNA was synthesized by polymerase chain reaction using a 5'-biotinylated upstream primer and immobilized on streptavidin-coated beads. Factors for initiation were derived from a HeLa extract. RNA synthesis was done in transcription buffer (12 mM HEPES, [pH 7.9], 12% v/v glycerol, 60 mM KCl, 0.12 mM EDTA, 0.12 mM EGTA, 1.2 mM dithiothreitol, and 0.003% IGEPAL CA-630 [Sigma-Aldrich]) containing 8 mM $MgCl_2$. All steps were at 25°C. For these studies, the adenovirus major late promoter had a modified downstream sequence (+1-ACTCTCTTCCCCTTCTCTTTCCTTCTCTTCCCTCTCCTCC-+40-AAAGCCTTT-+49). The purpose of the 39 nucleotide CT cassette was to synthesize C40 with addition of 300 μM ApC dinucleotide, 10 μM dATP, 5 μCi per reaction α-^{32}P-CTP (800 Ci per mmol), and 20 μM UTP (10 min), bypassing the requirement for addition of ATP and GTP. 20 μM CTP was then added for 10 min. C40 ECs were washed with 1% Sarkosyl and 0.5 M KCl buffer to dissociate initiation, elongation, pausing, and termination factors, contributed by the HeLa extract, and re-equilibrated with transcription buffer containing 8 mM MgCl2 and 20 μM CTP and UTP (except where noted). Functionally saturating concentrations of TFIIF (10 pmol) and TFIIS (3 pmol per reaction) were added as indicated. TFIIF and TFIIS were added to reactions to maximize the fraction of the post-translocated EC at A43 (Zhang and Burton, 2004) and to maximize isomerization reversal. TFIIF was added during the pre-incubation. TFIIS was added during the ATP pulse. Steps after ATP addition for A43 synthesis were performed by using the Kintek Rapid Chemical Quench-Flow (RQF-3) instrument at 25°C.

2.4 Running start, two-bond, double-quench protocol

The running start, two-bond protocol has been described (Funk et al., 2002; Nedialkov et al., 2003; Nedialkov et al., 2003; Zhang et al., 2003). Here, we improve the procedure by reporting data from experiments quenched with EDTA or with HCl (Arnold and Cameron, 2004; Arnold et al., 2004). Rapid quench experiments were done using the Kintek Rapid Chemical Quench-Flow (RQF-3) instrument. All steps were done at 25 °C. Elongation was through the sequence 40-CAAAGGCC-47. C40 ECs were incubated with 20 µM CTP and UTP. 10 or 100 µM ATP (depending on the protocol) was added to bead-immobilized C40 ECs on the bench top, to extend the EC to the A43 position. During the next 30–120 seconds (depending on the protocol), ECs were loaded into the left sample port of the RQF-3 instrument. GTP (in transcription buffer) at twice its working concentration was loaded into the RQF-3 right sample port. Programmed, equal volume mixing in the RQF-3 first combined ECs with GTP substrate. Reactions were then quenched with 0.5 M EDTA or 1 M HCl, after a precisely timed delay (>0.002 second). For EDTA-quenched reactions, ECs on beads were collected with a magnetic particle separator and processed by electrophoresis, as described (Funk et al., 2002; Nedialkov et al., 2003; Nedialkov et al., 2003; Zhang et al., 2003). HCl quenching dissociates the transcript from the RNAP II EC. HCl-quenched reactions were delivered into collection tubes containing a sufficient volume of 1 M KOH and 300 mM Tris-base (~70 µl), to neutralize the pH of the solution. Beads were removed using a

magnetic particle separator. The solution was extracted with an equal volume of phenol-chloroform. The aqueous phase was adjusted to 0.3 M sodium acetate containing 20 μg of glycogen carrier and ethanol-precipitated. After vacuum drying, RNA samples were analyzed by electrophoresis. Gel bands were quantified using a Molecular Dynamics Phosphorimager. At the A43 stall point, the EC fractionates into multiple conformational states, which are revealed by their distinct elongation kinetics to G44. Formation of the G44 bond, therefore, provides detailed insight into the mechanism and informs about recovery from the transcriptional stall at A43. The transition from G44→G45 provides information about processive elongation.

2.5 Quality of RNAP II elongation complexes

The complex kinetics we report demonstrate multiple conformers of A43 EC at the time of GTP addition in the running start, two-bond protocol. Because ECs were isolated on bead templates from HeLa nuclear extracts, it is reasonable to consider whether A43 ECs differ in their kinetic properties because of experimental treatments or because of damage to a subset of ECs during preparation. However, A43 conformational states are determined by treatments that occur after EC purification. The initial conformational states detected at A43 are different in the presence of TFIIF or in the absence of an elongation factor, showing that protein factors drive RNAP II between functional modes. Furthermore, increasing GTP concentration blurs the distinction between different kinetic states, indicating that RNAP II changes conformation through

interactions with substrate, as expected for an RNAP. Also, the distribution of A43 states is dependent on the time of stalling at A43, demonstrating reversibility between states (Nedialkov et al., 2003). A43 conformational states, therefore, are selected based on treatments (protein factors, substrate and time of incubation) after purification and are not an artifact of preparation. In the purification scheme, RNAP II molecules are selected for the ability to accurately initiate transcription in concert with the general initiation factors, and all C40 and A43 complexes are active in elongation. Sarkosyl and salt treatment appears to strip all contaminating transcription factors and complicating activities from the EC (Lei et al., 1999).

2.6 Isomerization reversal experiments

Isomerization reversal experiments and quantification of data were done essentially as described (Gong et al., 2005). Briefly, human C40 (40 nucleotide RNA ending in 3'-CMP) RNA polymerase II elongation complexes were formed by accurate initiation from the adenovirus major late promoter, using an ApC dinucleotide primer. The transcribed sequence was modified (Figure III-1) to allow C40 synthesis without substantial overrun. 5'-biotinylated DNA templates were immobilized on streptavidin beads. All reactions were done at 25 °C. C40 complexes were radiolabeled with α-^{32}P-CTP. C40 complexes were extended to A43 by addition of 100 µM ATP for 30 seconds (s). During the 30 s incubation, A43 complexes were transferred into the left sample port of the

KinTek RQF-3 Rapid Chemical Quench Flow instrument and then extended by computer-controlled rapid mixing. Reactions were quenched with EDTA or HCl, as indicated in individual protocols.

To reactions in which TFIIS was added during the ATP running start, 20 µM CTP and UTP were added during the pre-incubation to maintain the C40 stall position. In this protocol, CTP and UTP are diluted to a working concentration of 5 µM by two equal volume mixings. When TFIIS was not present during the pre-incubation, CTP and UTP were also omitted (except as specified in individual protocols). To compensate for inconsistencies in sample recovery or gel loading, transcripts were quantified as a ratio of relevant bands within each gel lane. Gel lanes were analyzed independently for percent of signal present in A43 plus all longer transcripts. The Molecular Dynamics Phosphorimager was calibrated and found to give a nearly linear response over a range of at least 20,000-fold ^{32}P exposure above a detectable background. A linear response over a 100-fold exposure range is sufficient to detect all relevant transcripts.

2.7 Molecular images

Molecular images were prepared using the graphics program Visual Molecular Dynamics (Humphrey et al., 1996). Structures were modified (i+2 and i+3 NTPs placed by modeling, DNA strands extended, non-template DNA placed), as described previously (Burton et al., 2005).

3. Results

3.1 Robust isomerization reversal.

In Figure III-1, we show an isomerization reversal experiment modeled on those of Gong et al. (2005) (Gong et al., 2005). Interestingly, the reaction is observed to advance and then retreat, particularly when accurately templated downstream NTPs are present (Figure III-1A). In the absence of accurately templated downstream NTPs, reversal is diminished and transient (Figure III-1B). Phosphorimager quantification of the experiment is shown in Figure III-1C, with the data represented on two time scales (1 and 0.2 seconds (s)). Normally, enzymatic reactions do not advance and then retreat. So how can such an unusual observation be understood?

"Isomerization reversal" refers to reversal of stable NTP loading to the i+1 active site of RNA polymerase II in response to: 1) a translocation block (α-amanitin); and 2) accurately templated downstream NTPs. Neither α-amanitin nor downstream NTPs induce strong reversal by themselves.

The reaction design is shown at the top of Figure III-1. Human RNA polymerase II is used to synthesize C40 elongation complexes (a 40 nucleotide RNA ending in 3'-CMP). The encoded RNA sequence downstream of C40 is 40-CAAAGCCUUU-49. Complexes are combined with elongation factor TFIIF, which stimulates forward translocation. On the bench top, 100 µM ATP and elongation factor TFIIS, which stimulates RNA dinucleotide cleavage and re-start, are added

Figure III-1. Robust isomerization reversal by human RNA polymerase II induced by accurately templated downstream NTPs. The RNA sequence and reaction protocol are shown at the top of the figure. □A indicates α-amanitin. A) Gel data for the reaction that includes 2.5 mM CTP and UTP, which are accurately templated at downstream sites. No transcripts longer than C45 are detected above background. 5 µM CTP and UTP were included in all reactions to maintain the C40 elongation complex. 0* indicates that ATP and chase NTPs were not added to C40 elongation complexes. 0 indicates that the reaction was stopped after the 30 s ATP pulse, with no NTP chase. Chase times are in seconds (0.002 to 1 s). The A41 transcript is generated by TFIIS-mediated dinucleotide RNA cleavage from the A43 stall position. B) Gel data for the reaction lacking accurately templated downstream NTPs adjacent to G44 (5 mM ATP in place of 2.5 mM CTP and UTP). C) Phosphorimager quantification of the gels shown in A and B. Two time scales (1 and 0.2 s) are shown. G44+ % indicates G44 plus all longer transcripts, indicated as percent of total (A43 and longer transcripts). Synthesis of the G44 bond is strongly and stably reversed in response to both α-amanitin and accurately templated downstream NTPs, using the EDTA quench procedure.

1
AcucuuccuucucucucuuccucuccucCaaAGCCU
40 43 44 45

C40 + TFIIF

2.5 mM GTP, CTP, UTP
or 2.5 mM GTP, 5.0 mM ATP

100 µM ATP
+ TFIIS

1 mM αA

EDTA
Stop

30 s

Δt

Figure III-1.

139
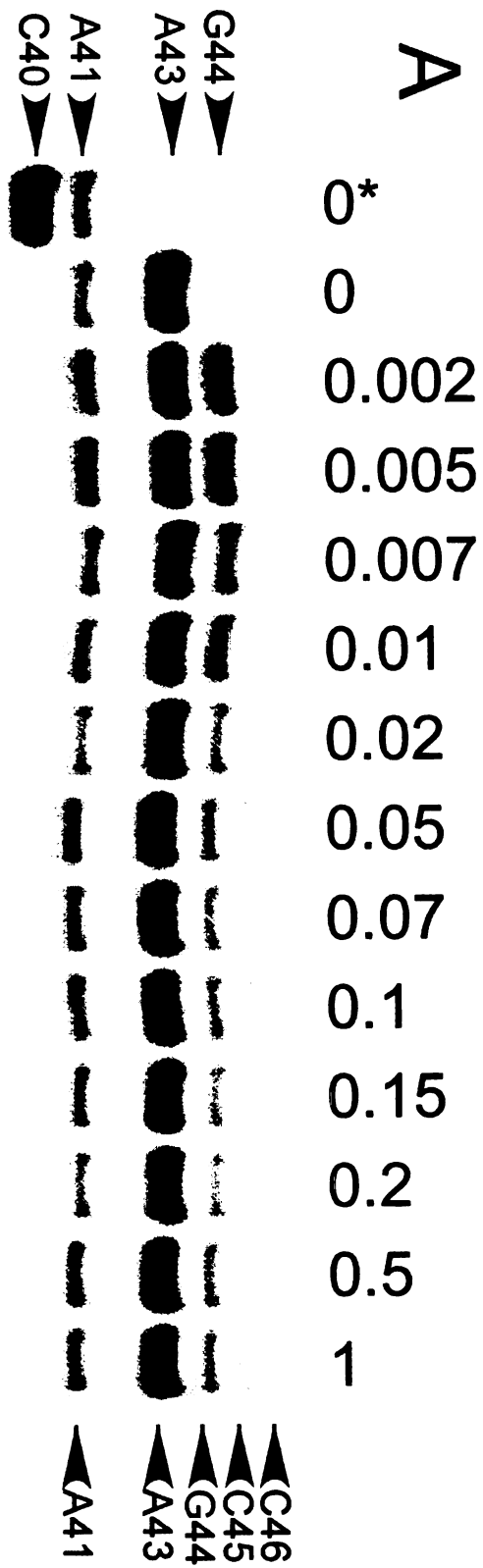
Figure III-1 (continued).

Figure III-1 (continued).

141

Figure III-1 (continued).

C

Figure III-1 (continued).

D

Plot of G44+ (%) versus Time (s). The y-axis (G44+ (%)) ranges from 0 to 40, and the x-axis (Time (s)) ranges from 0.0 to 0.2. Two curves are shown: filled squares for 2.5 mM GTP, CTP, UTP and circles for 2.5 mM GTP, 5.0 mM ATP. Annotations indicate IR=25%, IR=69%, and IR=100%.

Legend:
- ■ 2.5 mM GTP, CTP, UTP
- ⊙ 2.5 mM GTP, 5.0 mM ATP

for 30 s to advance the RNA to A43. The combination of TFIIF and TFIIS promotes forward translocation (TFIIF) and increases the fraction of elongation complexes on the active synthesis pathway (TFIIF and TFIIS), increasing the signal for the experiment. During the 30 s stall at A43, the elongation complex preparation is transferred to the sample port of the KinTek RQF-3 rapid chemical quench flow instrument. Using the rapid mixing device, the sample is combined with 2.5 mM GTP, 2.5 mM CTP, 2.5 mM UTP, and 1 mM α-amanitin for various

times (0.002 to 1 s) before quenching with 0.5 mM EDTA (Figure III-1A). The G44 signal is highest at the 0.002 and 0.005 s time points and then decreases. Despite the abundance of CTP substrate, a very small amount of C45 is detected, starting at 0.05 s. Little C45 is synthesized, and no longer transcripts are detected above background, because α-amanitin is added to the reaction,
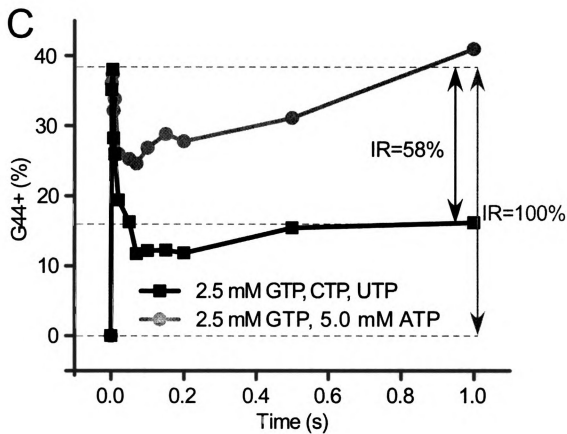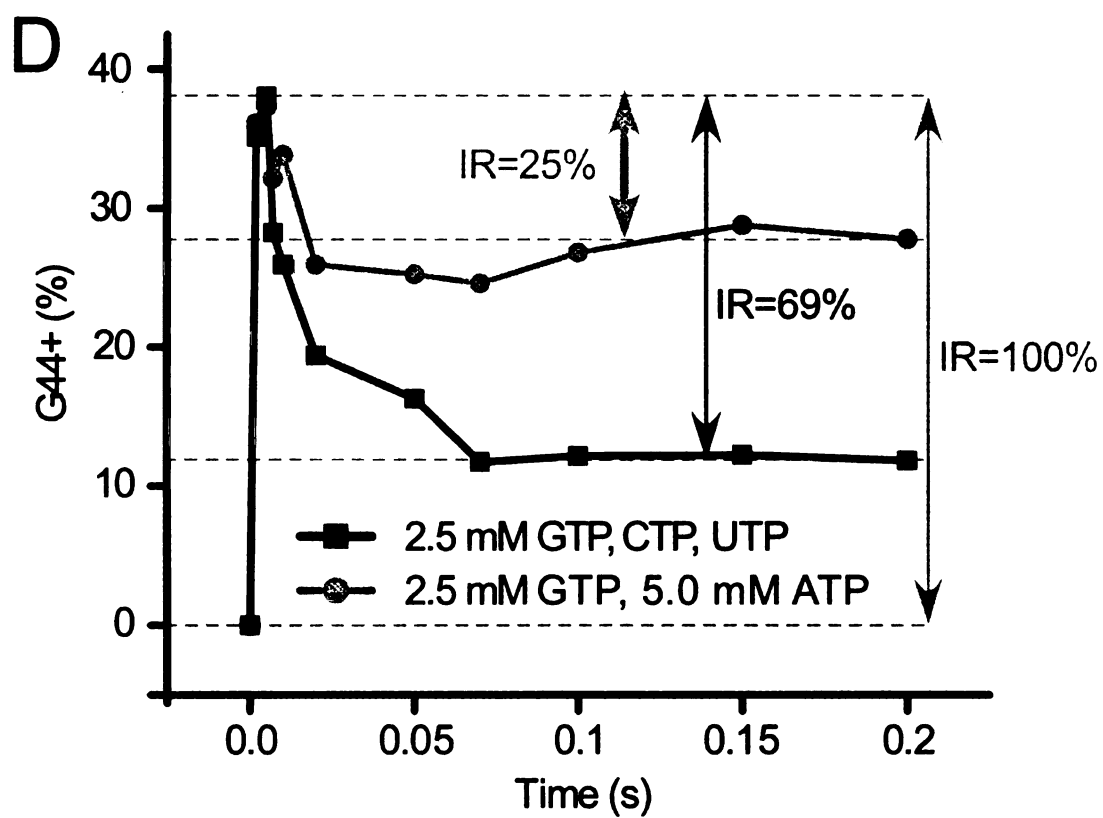
and α-amanitin blocks translocation (Bushnell et al., 2002; Gong et al., 2004;

Gong et al., 2005). α-amanitin appears to allow synthesis of G44 primarily from

those elongation complexes that were initially post-translocated at the A43 stall position. After G44 synthesis, C45 synthesis requires translocation, so little C45 and no longer transcripts are observed, indicating that α-amantin effectively

blocks translocation. While GTP for G44 synthesis occupies the i+1 active site, CTP can potentially occupy the i+2 and i+3 downstream template sites, and UTP can potentially occupy the i+4, i+5, and i+6 downstream sites, depending on the availability of downstream DNA for specific NTP interactions.

This experiment demonstrates 69 % transient (0.2 s) and 58 % sustained (1 s) isomerization reversal with added CTP and UTP, calculated as 100 % - ((G44+ % (0.2 or 1 s))/(G44+ % (burst; 0.002 or 0.005 s)) x 100 %). G44+ indicates G44 plus all longer transcripts (in this case, C45). We observe greater isomerization reversal (IR = 69% or 58% in this experiment) than observed in similar experiments using the same template in Gong et al. (2005) (Gong et al., 2005) (IR = 25%) because we used 1 mM α-amanitin in this experiment and 0.5 mM α-amanitin in comparable experiments in Gong et al. (2005). Also, UTP, which stimulates reversal, was omitted in the experiments of Gong et al. (2005) (Gong et al., 2005).

When 2.5 mM CTP and UTP are substituted with 5 mM ATP (an NTP that is not accurately templated at adjacent downstream sites), isomerization reversal is significantly reduced and is overcome with longer times of incubation (IR = 25% at 0.2 s and IR = 0% at 1 s). Because the extent of reversal with negative control samples varies with the time of EDTA quench addition, reversal appears to be primarily determined by the activity of GTP- and/or ATP-Mg^{2+}. Because of prior addition, all reactions in Figure III-1 contain 5 μM CTP and UTP (added to maintain the C40 stall position in the presence of the RNA cleavage factor TFIIS). 5 μM CTP and UTP may be sufficient to contribute weakly to reversal. We suggest that NTPs that are not accurately templated may interact weakly at downstream sites and promote transient reversal. When accurately

templated downstream NTPs are limiting, GTP-Mg^{2+} becomes stabilized in the i+1 active site and becomes increasingly resistant to reversal with the passage of time.

The observation of reversal with control samples is of interest, because it shows isomerization reversal in response to a translocation block. Furthermore, a mechanism for exchange of the i+1 GTP substrate is indicated that responds to free pools of NTPs. The data indicate that free NTP-Mg^{2+} is necessary to displace the i+1 GTP substrate, because little or no reversal is detected with the earliest times of EDTA addition (0.002 to 0.005 s), a treatment that eliminates free NTP-Mg^{2+} but not GTP-Mg^{2+} tightened in the active site. Isomerization reversal, therefore, (1) requires free NTP-Mg^{2+}, (2) is strongly stimulated by accurately templated downstream NTP-Mg^{2+}, and (3) appears to be a more dramatic phenomenon than can be fully scored using the current reaction design because of rapid GTP-Mg^{2+} re-loading after release, prior to EDTA addition.

Because the experiment in Figure III-1 indicates the activity of accurately templated CTP and possibly UTP at downstream sites, i.e. i+2, i+3, and i+4, on the fate of the GTP loaded and initially tightened in the active site (i+1), this experiment indicates NTP loading at downstream sites while the active site is occupied. EDTA quenching is expected to chelate Mg^{2+} from downstream CTP- and UTP-Mg^{2+} very rapidly. The i+1 GTP-Mg^{2+}, however, can be protected within the active site from EDTA quenching, and can proceed to form the G44 phosphodiester bond after EDTA addition. We posit that, because of NTP-driven translocation, CTP and UTP strain against the α-amanitin translocation block,

resulting in expulsion of a large fraction of the i+1 GTP-Mg^{2+}, which had previously been fated for G44 bond synthesis. Thus, RNA polymerase II could utilize incoming NTPs in dynamic error avoidance and error correction, to relieve a translocation block (Gong et al., 2005). Essentially, RNA polymerase II appears to react to the α-amanitin translocation block by weakening binding of the i+1 GTP in the active site. Because GTP is the accurately templated substrate for G44 synthesis, RNA polymerase II cannot easily reject GTP, limiting the extent of isomerization reversal. As a result, only 69 or 58 % reversal is achieved.

Chelation of Mg^{2+} at early times (0.002 and 0.005 s) results in robust G44 synthesis, because of rapid inactivation of CTP- and UTP-Mg^{2+}. Chelation of Mg^{2+} at a later time (i.e. 0.02 s) results in dramatically reduced G44 synthesis, apparently because of the activities of CTP and UTP acting at downstream template positions. In Supplementary Material, a similar isomerization reversal experiment is shown in which CTP and UTP are replaced by CTP and ATP. Because substituting UTP with ATP slightly reduces isomerization reversal, it appears that i+4 UTP contributes to reversal.

An important feature of the isomerization reversal experiment is that α-amanitin is used on a millisecond time scale as a transient state inhibitor. In this reaction design, α-amanitin is added together with accurately templated substrate and downstream NTPs, and α-amanitin exerts its effects within milliseconds. RNA polymerases catalyze an elongation step on about a 5 to 50 millisecond

time scale, underscoring the need for millisecond time resolution in analysis of single bond additions (Gong et al., 2004; Gong et al., 2005; Nedialkov et al., 2003; Zhang and Burton, 2004; Zhang et al., 2005). In isomerization reversal experiments, conditions for forward synthesis and inhibition by α-amanitin are established together, and what results is a glimpse of the internal mechanism of RNA synthesis, as the reaction is driven forward by substrate and blocked at the translocation step by inhibitor.

3.2 NTPs occupy template at least to i+4.

In previous work, accurate NTP interactions were detected at the i+2 and i+3 downstream sites, while the i+1 substrate site was occupied (Gong et al., 2005). Here, we report templated interactions of NTPs at least to the i+4 position, while the i+1 site is occupied. The assays shown in Figure III-2 differ from those shown in Figure III-1 in that TFIIS was omitted, reducing the magnitude of the G44 transcription signal, but, perhaps, increasing the translocation force generated by TFIIF (in the absence of TFIIS). Furthermore, α-amanitin was added at 0.5 mM, instead of 1 mM. Because we wished to test many conditions and many time points, we needed to reduce the cost incurred from adding the higher concentration of α-amanitin. Also, by slightly slowing the rate of α-amanitin binding by limiting concentration, some elongation complexes may be able to further advance through the bond addition mechanism before inhibition is fully

148

established. One goal of the transient state inhibition experiment is to reveal the inner mechanism of RNA polymerase II elongation.

Figure III-2 shows isomerization reversal experiments with addition of different downstream NTPs and analogues to test the requirements for occupying the i+2 and i+3 CTP sites (considered together) and the i+4, i+5, and i+6 UTP sites (considered together), while the i+1 active site is occupied by GTP. Precisely how many downstream DNA template sites are bound by NTP substrates has not been determined, so we wondered whether the i+4 site is accurately occupied. For this experiment, concentrations of downstream NTPs were reduced 5-fold compared to Figure III-1, from 2.5 to 0.5 mM, close to physiological levels, to show that physiological concentrations of downstream NTPs are sufficient to induce isomerization reversal. Reducing the concentration of CTP (i+2 and i+3) in the reversal experiment is also expected to accentuate any effects of UTP (i+4, i+5 and i+6). Because TFIIS was omitted from these reactions, there was no need to include traces of CTP and UTP to maintain the C40 stall position, before the addition of substrate NTPs, so CTP and UTP were not added to these reactions during the pre-incubation.

Depending on the combinations of NTPs added, reversal curves separate into 3 classes (Figure III-2). In every case, 2.5 mM GTP is included as the substrate for G44 synthesis because this concentration allows very rapid binding of GTP at i+1. α-amanitin is added at 0.5 mM. Downstream NTPs are added at 0.5 mM. Reactions are stopped by addition of EDTA. The combination of GTP, CTP, and UTP induces the most robust reversal (IR = 50% at 0.2 s). Omission of

Figure III-2. NTPs that are accurately templated at adjacent downstream positions (i+2, i+3, and i+4, at a minimum) are required for robust isomerization reversal at i+1 (G44). The protocol is shown at the top of the figure. RNA sequence is the same as shown in **Figure III-1**. TFIIS was omitted from reactions. CTP is required to observe robust G44=>A43 isomerization reversal. UTP (in the presence of CTP) strongly stimulates isomerization reversal. Close analogues of CTP and UTP do not contribute to isomerization reversal. IR% is calculated between 0.002 or 0.005 s and 0.2 s. Error bars indicate standard deviation of three independent experiments performed on different days. For some data points, error bars are obscured by the graph symbol. G, A, C, U, dC, dT, dU indicates the NTP or dNTP. 2A, 2U indicates 1 mM ATP or UTP.

C40 + TFIIF

100 µM ATP

30 s

2.5 mM GTP
+/- 0.5 mM CTP
+/- (d)NTP or NDP
0.5 mM αA

Δt

EDTA
Stop

G, C, U
G, C
G, C, A
G, C, dT
G, C, dU

G, C, UDP
G, 2A
G, 2U
G, dC, U
G, CDP, U

Figure III-2.

151

Figure III-2 (continued).
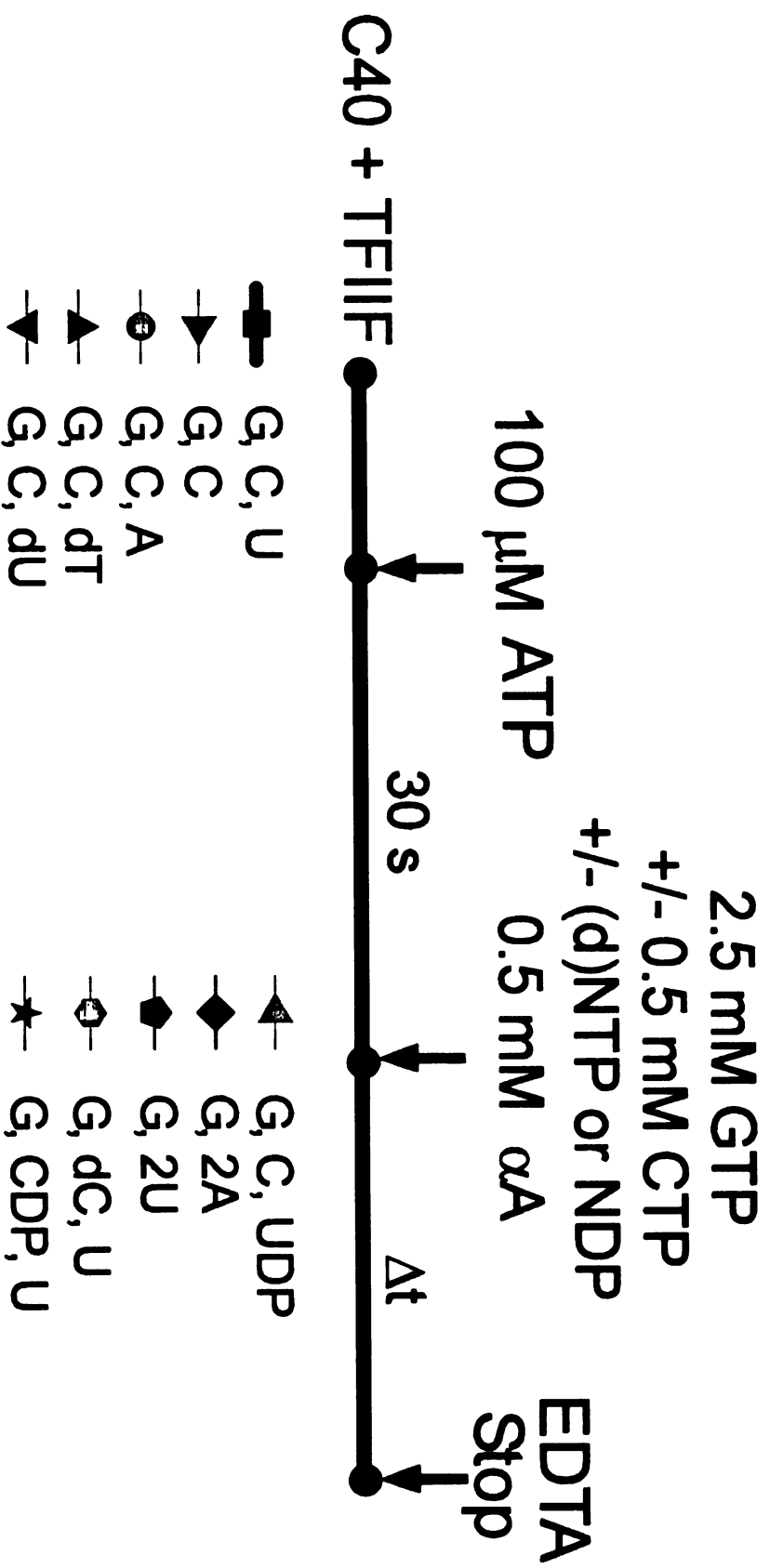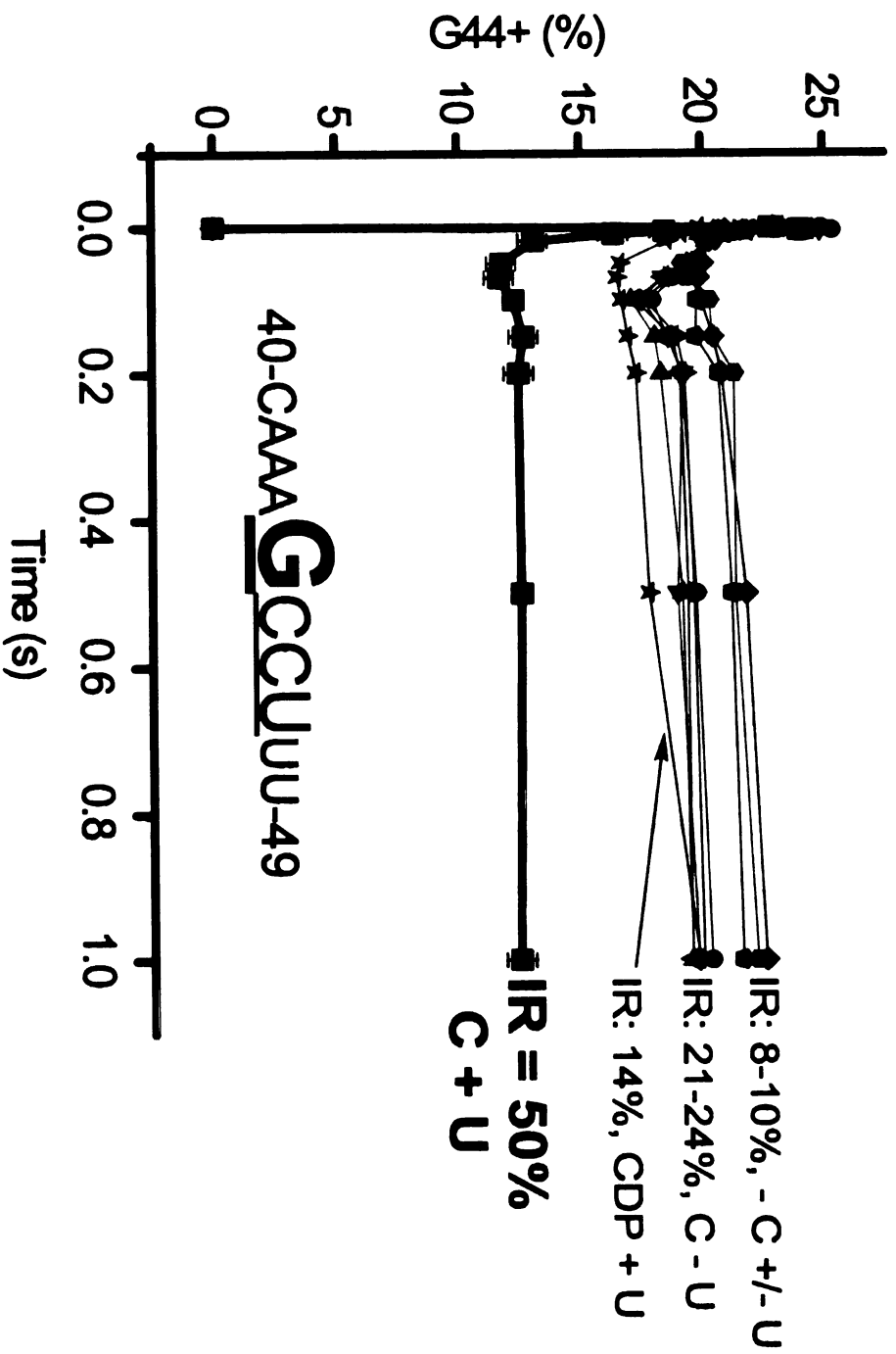
40-CAAA**G**CCUUUU-49

IR = 50%
C + U

IR: 14%, CDP + U

IR: 21-24%, C - U

IR: 8-10%, - C +/- U

CTP results in almost no sustained reversal, because the i+2 and i+3 sites must be accurately occupied by CTP to strongly stimulate reversal. Reactions lacking CTP, therefore, result in limited, transient reversal, and serve as controls for the importance of downstream, accurately templated NTPs.

GTP, UTP, ATP, and dCTP will not substitute for CTP, at the i+2 and i+3 positions (IR = 8-10 % at 0.2 s) (Figure III-2). CDP added in place of CTP also does not stimulate robust reversal, although the curve does not precisely overlay with the other control curves lacking CTP (IR = 14% at 0.2 s; green stars; the curve has a reduced burst height). The slight, transient reversal observed in control samples lacking CTP can be attributed to effects of non-templated NTP-Mg^{2+} and α-amanitin. Probably, inaccurately templated NTPs contribute to reversal in control experiments, because different times of EDTA addition result in different responses, indicating that the mechanism for reversal involves EDTA chelating Mg^{2+} from NTP-Mg^{2+}. It is not possible to exclude non-templated NTPs from the reaction, because i+1 substrate GTP must be added for G44 bond synthesis.

Surprisingly, UTP, which is accurately templated at the i+4, i+5, and i+6 downstream sites, stimulates isomerization reversal at G44 (Figure III-2). In the presence of CTP, which is required for reversal, withholding UTP from the reaction reduces reversal. This result indicates that NTPs can simultaneously and accurately occupy the i+1 (GTP; active site), i+2 (CTP), i+3 (CTP), and i+4 (UTP) sites, at a minimum. Whether NTPs can interact at the i+5 (UTP) and i+6 (UTP) sites has not been determined. Because dTTP and dUTP (2'-deoxy UTP)

do not stimulate reversal (in the presence of CTP), dTTP and dUTP cannot substitute for UTP at the i+4, i+5, and i+6 positions (considered together). GTP, ATP, CTP, and UDP also fail to substitute for UTP in the reversal experiment.

These results demonstrate the high selectivity of downstream NTP binding. Binding is accurately templated, but the capacity for accurate base-pairing to template is not sufficient. NTPs cannot be substituted with NDPs or dNTPs, even when these analogues have comparable base-pairing specificity. This result demonstrates selectivity of NTP interactions at downstream sites, consistent with the high fidelity of RNA synthesis catalyzed by multi-subunit RNA polymerases.

In Figure III-2, we demonstrate strong suppressive effects of UTP on synthesis of G44, in the presence of CTP and α-amanitin. As predicted by the NTP-driven translocation hypothesis, UTP also suppresses synthesis of C45 and C46 in the presence of the α-amanitin translocation block (Figure III-3). In order to provide a fair comparison between those samples that contain CTP and UTP to those samples that contain CTP but lack UTP, the synthesis of C45 and C46 was normalized, respectively, to synthesis of the previous G44 and C45 bonds. Only those samples containing CTP are considered, because these are the only samples that can synthesize C45 and C46. Remarkably, inclusion of UTP in the reaction strongly suppresses C45 synthesis and (within detection) abolishes C46 synthesis. For C45 synthesis, UTP is accurately templated at the i+3, i+4, and i+5 downstream positions. For C46 synthesis, UTP is accurately templated at the

Figure III-3. UTP, templated at downstream sites, suppresses synthesis of C45 and C46. C45+/G44+ % (A) and C46+/C45+ % (B) were plotted versus time. In this way, C45 and C46 synthesis rates were normalized for available G44 and C45 substrate, respectively. The protocol is shown at the top of the figure. dUTP, dTTP, UDP, ATP, GTP, and CTP do not substitute for UTP for suppression of C45 and C46 synthesis. UTP strongly suppresses detection of the C46 transcript.
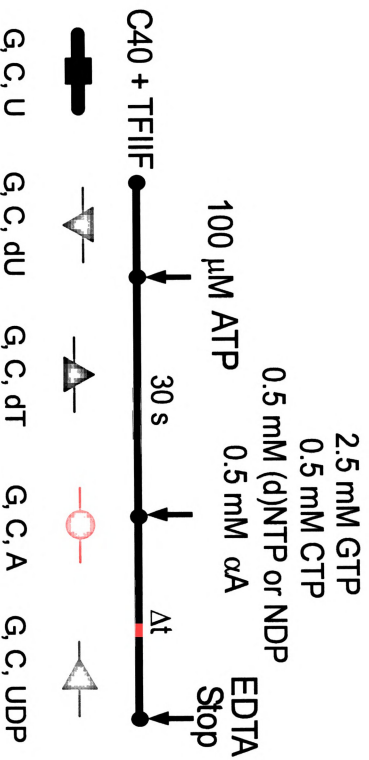
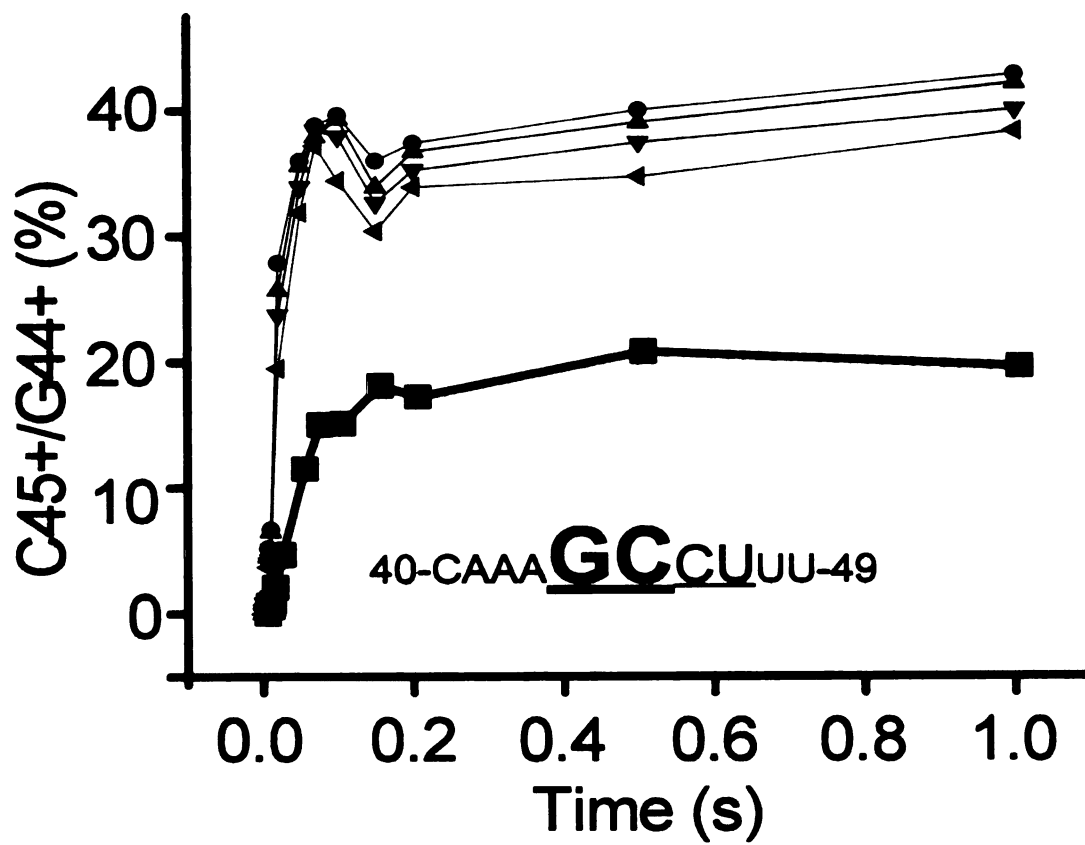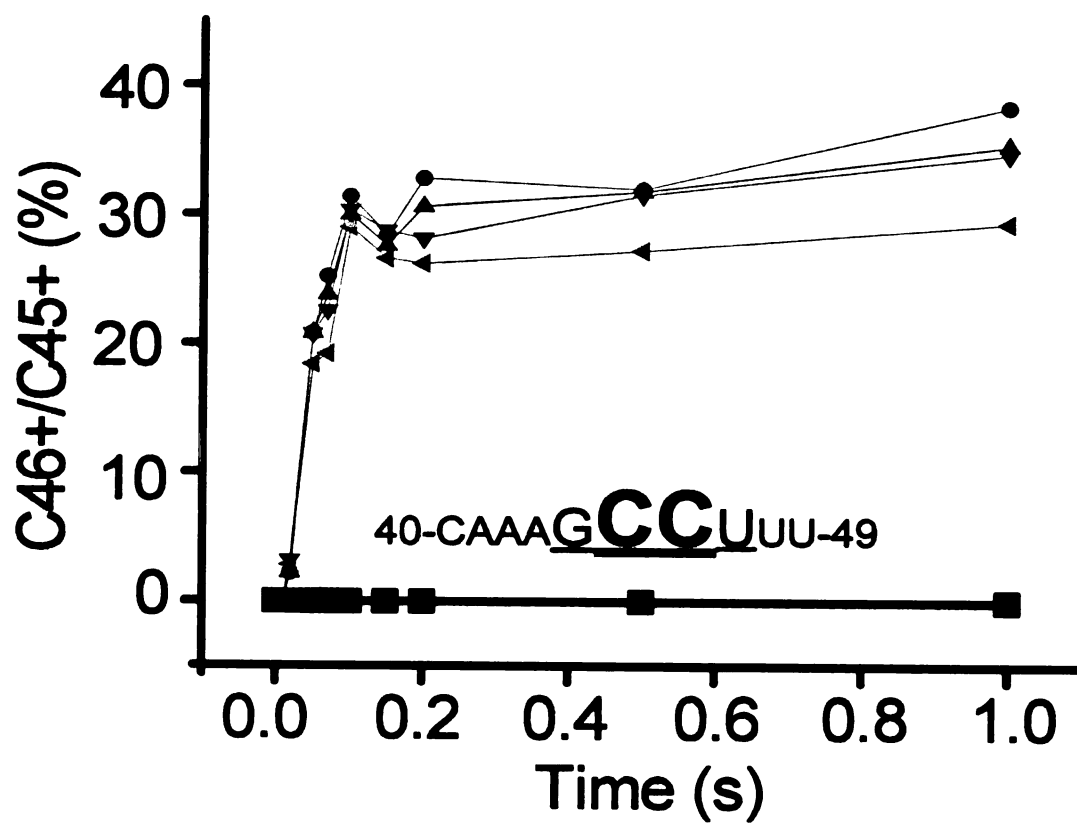**Figure III-3.**

Figure III-3 (continued).

Figure III-3 (continued).

i+2, i+3, and i+4 positions. ATP, GTP, CTP, dUTP, TTP, and UDP fail to replace UTP in reversal reactions. Once again, this result is most consistent with the NTP-driven translocation model. This result does not appear consistent with the secondary pore NTP loading hypothesis.

3.3 The core mechanism of RNA synthesis.

In order to gain more insight into the reversal experiment, we compared EDTA quenched reactions to reactions quenched with HCl. In addition, to gain insight into the effects of accurately templated downstream NTPs, the concentrations of CTP and UTP were varied (Figure III-4). Stopping the reaction with acid shows the timing of synthesis of the phosphodiester bond (Arnold and Cameron, 2004; Arnold et al., 2004; Gong et al., 2005; Zhang and Burton, 2004; Zhang et al., 2005). Bond synthesis, however, may remain reversible until pyrophosphate is released. If the phosphodiester bond has been formed, however, the bond cannot reverse after acid addition. If the bond has not yet formed, the bond cannot be formed after HCl addition. By contrast, after EDTA quench addition, two outcomes are possible: 1) phosphodiester bond formation; and 2) phosphodiester bond reversal. Because an NTP-Mg^{2+} can be sequestered from EDTA chelation within the i+1 active site, the NTP-Mg^{2+} can progress to synthesize a phosphodiester bond after EDTA addition. Alternatively, the NTP-Mg^{2+} can potentially be destabilized in the active site after EDTA addition, as in the isomerization reversal experiment.

An ongoing issue in RNA synthesis is the dynamics of the pyrophosphate release step and its link to translocation. Based on transient state

159

kinetic analyses of human RNA polymerase II elongation, the Burton laboratory posits that pyrophosphate release is coupled to NTP-driven translocation, explaining the surprisingly high NTP dependence of the rate-limiting step during elongation that corresponds to both translocation and pyrophosphate release (Gong et al., 2005; Zhang and Burton, 2004; Zhang et al., 2005). By contrast, the secondary pore NTP loading hypothesis seems to require that pyrophosphate release and translocation be rapid, spontaneous and NTP-independent. Without these features in a secondary pore NTP loading model, RNA polymerase would pause after synthesis of every bond, independent of substrate NTP concentration. According to estimates of the Kornberg laboratory, for NTPs to enter the active site through the secondary pore by diffusion, NTP loading would likely be rate-limiting for RNA synthesis (20 to 30 nucleotides s^{-1}) (Batada et al., 2004). So far as we can discern from rapid kinetic studies, NTP loading is not rate-limiting for RNA synthesis (>1450 +/- 330 s^{-1}) and pyrophosphate release is not rapid and spontaneous (about 30 s^{-1} and coupled to the NTP-driven translocation step) (Zhang and Burton, 2004).

In the presence of α-amanitin, RNA polymerase is stalled at the core of the bond addition mechanism, and the position of stalling is dependent on accurately templated downstream NTPs (Figure III-4). Stalling is indicated by the failure of the HCl quench (phosphodiester bond formation) and EDTA quench (stable GTP-Mg^{2+} loading) curves to converge. With addition of GTP, CTP, and UTP to the reaction, a long-lived steady state is established in which the HCl quench curve rises above the EDTA quench curve, but does not rise to the level

160

Figure III-4. Evidence for a tunable ratchet driving translocation adjusted by altering the concentrations of accurately templated downstream NTPs. A comparison of HCl quench (red circles) and EDTA quench (black squares) curves is shown for reactions done at different CTP and UTP concentrations (0, 0.25, 0.5, and 2.5 mM). A) 2.5 mM CTP and UTP. The gray line indicates the negative control experiment with 0 mM CTP and UTP (also shown for reference in B and C). B) 0.5 mM CTP and UTP. Error bars indicate standard deviation of three independent experiments done on different days. Some error bars are obscured by the graph symbols. C) 0.25 mM CTP and UTP. D) 0 mM CTP and UTP. In A-C, in which downstream NTPs are present, the HCl quench curve rises above the EDTA quench curve. In D, in which adjacent downstream NTPs are absent, the EDTA quench curve rises above the HCl quench curve.

Figure III-4.

C40 + TFIIF

100 μM ATP

2.5 mM GTP
0-2.5 mM CTP, UTP
0.5 mM αA   EDTA or HCl

30 s

Δt

Stop

40-CAAA**G**CCUUU-49

Figure III-4 (continued).

A

Graph: G44+ (%) versus Time (s)

Legend:
- 2.5 mM CTP, UTP, EDTA
- 2.5 mM CTP, UTP, HCl
- No CTP, UTP, EDTA

Figure III-4 (continued).

B

Legend:
- ■ 0.5 mM CTP, UTP, EDTA
- ● 0.5 mM CTP, UTP, HCl
- ▲ No CTP, UTP, EDTA
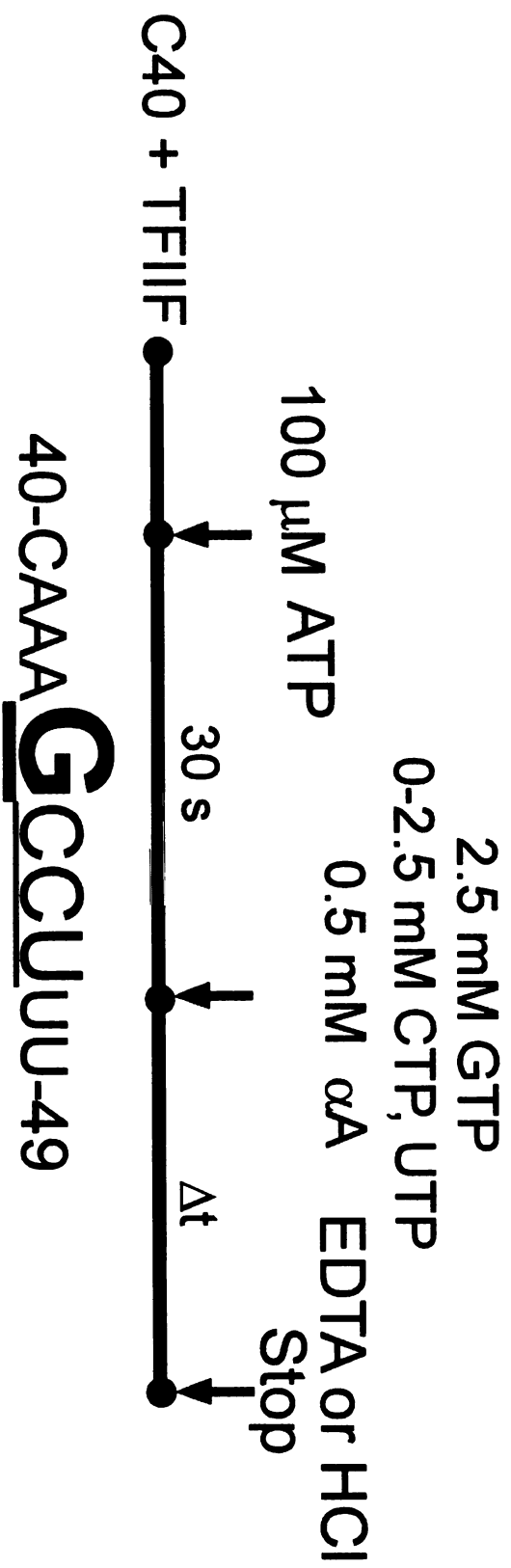
X-axis: Time (s)
Y-axis: G44+ (%)
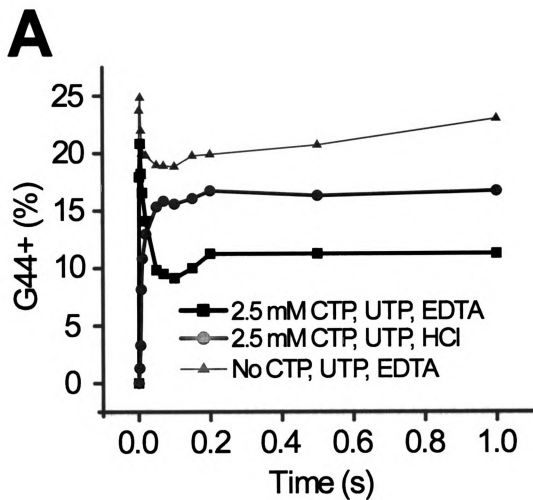
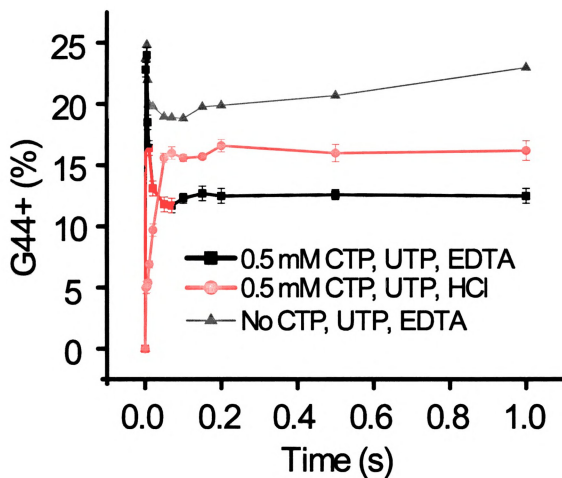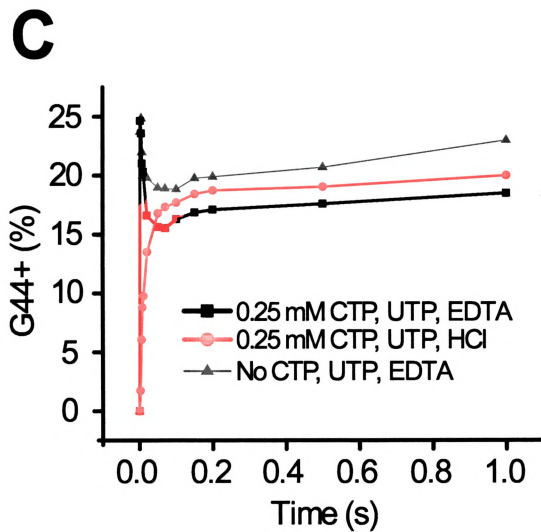**Figure III-4 (continued).**

Figure III-4 (continued).
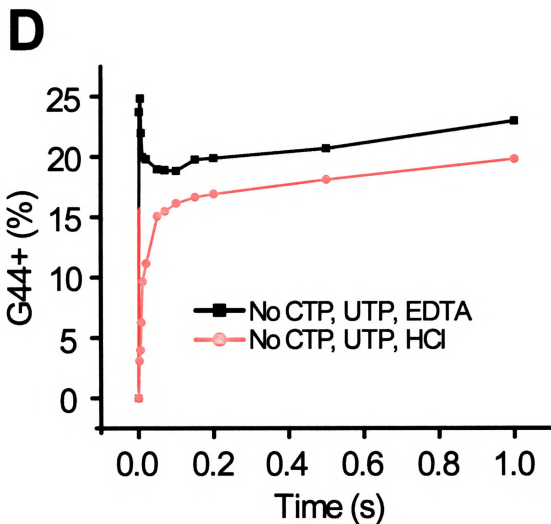
of control samples (gray symbols) lacking CTP and UTP (Figures III-4A-D). Higher concentrations of CTP and UTP appear to develop increased translocation strain on the elongation complex, and therefore increase the persistent separation of the HCl and EDTA quench curves (compare Figures III-4A-C). Long-lived elevation of the HCl quench curve above the EDTA quench curve indicates that the G44 phosphodiester bond has been formed, but pyrophosphate release has been prevented by the translocation block. In the absence of α-amanitin, convergence of HCl and EDTA quench curves indicates that pyrophosphate has been released and that bond synthesis is complete. Having the HCl quench curve rise above the EDTA quench curve indicates that, after EDTA addition, synthesis of the G44 phosphodiester bond is reversed by endogenous pyrophosphorolysis (G44.PPi<=>A43.GTP=>A43 + GTP).

Persistent straining of accurately templated downstream NTPs against the α-amanitin translocation block appears to force RNA polymerase into a more highly strained conformation than that assumed with addition of non-templated NTPs. When EDTA is added to strained complexes, and Mg^{2+} is chelated, disabling downstream CTP- and $UTP-Mg^{2+}$, it appears that phosphodiester bond reversal through endogenous pyrophosphorolysis becomes a dominant event. When accurately templated downstream NTPs are withheld from the reaction, this strained conformation of the elongation complex is not as strongly adopted, and robust and sustained bond reversal is not observed (Figure III-4D).

Another way of viewing the somewhat unexpected observation of HCl quench curves rising above EDTA quench curves is this. At the time of EDTA addition, the extent of phosphodiester bond synthesis is that shown by the HCl quench curve. After EDTA addition, the reaction reverses (in response to EDTA addition) to the position indicated by the EDTA quench curve. Loss of G44 signal requires reversal of phosphodiester bond synthesis, in this case, after addition of the EDTA quench. The bond reversal reaction requires endogenous pyrophosphate, once again indicating the difficulty in releasing pyrophosphate and also indicating coupling of pyrophosphate release to NTP-driven translocation, a step that is effectively blocked by α-amanitin. Addition of 10 mM exogenous pyrophosphate has no effect on isomerization reversal reactions (data not shown), and pyrophosphate was not an added component of these reactions. Only very small amounts of pyrophosphate are generated from elongation reactions, so exogenous pyrophosphorolysis does not contribute to reversal.

Having HCl quench curves rise above EDTA quench curves provides further evidence of phosphodiester bond reversal in response to addition of EDTA. According to our interpretation of isomerization reversal, G44.PPi transcripts, which score as G44 transcripts by HCl quenching, revert to A43, in response to EDTA addition. Because G44 transcripts are detected by HCl quenching and lost after EDTA quenching, G44 transcripts are not lost in the quantification by extension to C45 (or beyond) and, thus, escaping detection. Bond reversal in response to EDTA addition indicates that chelation of CTP- and

UTP-Mg^{2+} may inhibit accurately templated NTP-driven translocation, which is required to maintain formation of the G44 bond. Furthermore, accurately templated NTP-driven translocation appears to be coupled to phosphodiester bond synthesis and pyrophosphate release. Comparing Figures III-4A-C shows that reducing CTP- and UTP-Mg^{2+} concentration reduces separation of the HCl and EDTA quench curves, indicating that lower concentrations of incoming NTPs may generate reduced translocation pressure against the α -amanitin translocation block.

The experiment shown in Figure III-4, in which HCl quench and EDTA quench curves fail to converge within 1 s, was done in the presence of TFIIF and the absence of TFIIS. In previously published experiments (Gong et al., 2005), which were done in the presence of both TFIIF and TFIIS, HCl quench and EDTA quench curves converged within 0.2 to 0.4 s when accurately templated downstream NTPs were present. By contrast, in the absence of accurately templated downstream NTPs, convergence required about 2 s. Slow convergence of HCl and EDTA quench curves was explained because accurately templated NTP-driven translocation was necessary for bond completion or bond reversal, particularly in the presence of a translocation block. Comparing the results in Figure III-4 with previous results indicates that TFIIS will interact with elongation complexes during active synthesis and help in some way to cause either bond completion (pyrophosphate release) or isomerization reversal (or both).

To confirm the previous result, we did the experiment shown in Figure III-5, in which we compare the convergence of HCl and EDTA quench curves in the presence and absence of TFIIS. As expected, TFIIS causes rapid convergence of the HCl quench and EDTA quench curves (within about 0.1 s) (Figure III-5). In Figure III-5, reactions that include TFIIS are identical to those that lack TFIIS, except that, in this case, TFIIS was added with GTP, CTP, UTP, and α-amanitin. If TFIIS is added to the reaction during the ATP pulse, the results are similar to those previously observed in Gong et al. (2005) (Gong et al., 2005). Convergence of HCl quench and EDTA quench curves, in the presence of TFIIS, demonstrates a previously unknown function of TFIIS. TFIIS interacts with stalled elongation complexes that have loaded an NTP into the active site and are engaged in forming a phosphodiester bond. TFIIS affects isomerization reversal, and reversal is an apparent fidelity mechnanism, involving rejection of the substrate NTP in response to a translocation block. It was previously unknown that TFIIS could interact with actively transcribing complexes. Furthermore, because TFIIS contributes to convergence of HCl and EDTA quench curves, separation of HCl and EDTA quench curves observed in Figures III-4 and 5 is not due to an artifact of quantification. The observation of phosphodiester bond reversal in Gong et al. (2005) (Gong et al., 2005) is confirmed, indicating that accurately templated NTP-driven translocation is coupled to isomerization reversal, bond formation and pyrophosphate release.

Figure III-5. Evidence for a tunable ratchet driving translocation that is regulated by TFIIS. TFIIS helps to resolve translocation blocked RNA polymerase II elongation complexes, without inducing RNA dinucleotide cleavage. In the absence of TFIIS, EDTA and HCl quench curves fail to converge within 1 s (red symbols). In the presence of TFIIS (black symbols), EDTA and HCl quench curves converge rapidly (within 0.1 to 0.2 s).
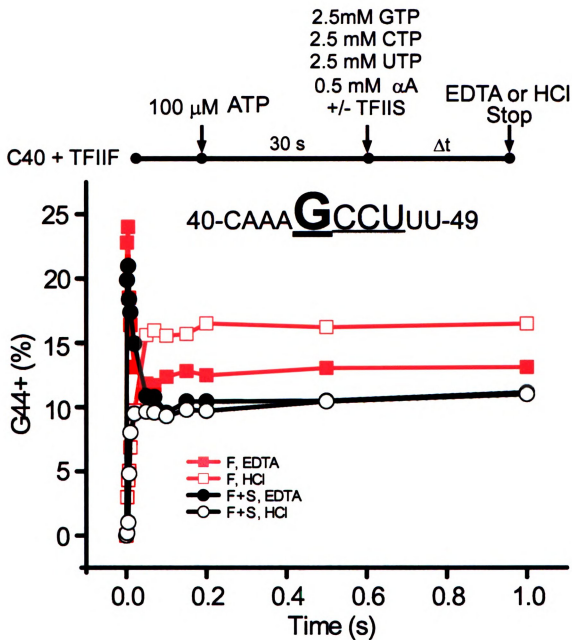
Figure III-5.

4. Discussion

4.1 NTP driven translocation.

In the presence of a translocation block (α-amanitin), we demonstrate robust effects of NTPs that are accurately templated at adjacent downstream positions (i+2, i+3, and i+4, at a minimum) on the fate of a substrate NTP loaded and initially tightened in the i+1 active site of human RNA polymerase II. Neither improperly templated NTPs, nor accurately templated NDPs, nor accurately templated dNTPs show these effects, revealing chemical selectivity at multiple downstream positions. Because these results demonstrate simultaneous NTP occupancy of the active site and downstream sites, these results support the NTP-driven translocation model and appear inconsistent with the secondary pore being the sole route of NTP loading for human RNA polymerase II. Based on available structures (Gnatt et al., 2001; Kettenberger et al., 2004; Wang et al., 2006; Westover et al., 2004; Westover et al., 2004), it does not appear that NTPs can access the i+2, i+3, and i+4 sites through the secondary pore.

Recent x-ray crystal structures of *T. thermophilus* RNA polymerase elongation complexes indicate that NTP-driven translocation may not be general for all multi-subunit RNA polymerases. Because *T. thermophilus* structures have a closed downstream bubble (closed at i+2), this raises the question of whether bacterial RNA polymerase can bind an NTP substrate at i+2. It should be noted that merely because *T. thermophilus* has a closed downstream bubble in some elongation complex structures, this does not show that the downstream bubble is

always closed. For instance, allosteric effects of substrate NTPs with *E. coli* RNA polymerase may indicate interactions of an NTP substrate at or near the i+2 site (Foster et al., 2001; Holmes and Erie, 2003). Single molecule studies of elongation catalyzed by *E. coli* RNA polymerase are also most consistent with NTP substrates binding at or near i+2 (Abbondanzieri et al., 2005).

4.2 Isomerization reversal and the "trigger loop" hypothesis.

In X-ray crystal structures, multiple placements of the trigger loop (Rpb1 1081 to 1093, *S. cerevisiae* residues) have been observed, including "open" and "closed" conformations (Vassylyev et al., 2007; Vassylyev et al., 2007; Wang et al., 2006). The trigger loop appears to close over the i+1 NTP substrate as part of an induced fit mechanism, tightening the catalytic intermediate to promote catalysis and ensure transcriptional fidelity. The trigger loop, however, is not required for phosphodiester bond formation by multi-subunit RNA polymerases. A radical deletion of the *E. coli* RNA polymerase trigger loop is slow for elongation, but it remains active (Toulokhonov et al., 2007). A four amino acid substitution mutant designed to block trigger loop dynamics is also very slow in elongation (Vassylyev et al., 2007). Based on the *T. thermophilus* elongation complex structures, *E. coli* β' M932, R933, and H936 are predicted to be key residues to support elongation and to contribute to transcriptional fidelity. Although mutant proteins in these residues were tested for pausing half-life, activities in elongation compared to wild type RNA polymerase were not reported (Toulokhonov et al., 2007). *E. coli* β' T934M, found in this critical region, is rapid

in elongation compared to wild type (Weilbaecher et al., 1994). Mutations in the trigger loop increase the duration of pauses at an RNA hairpin-dependent pause site from the histidine operon, demonstrating that the trigger loop has an important role in regulating transcriptional pausing (Toulokhonov et al., 2007).

Trigger loop closing, however, appears inconsistent with α-amanitin (Bushnell et al., 2002; Wang et al., 2006) and with TFIIS (Kettenberger et al., 2003; Kettenberger et al., 2004) binding to the active elongation complex. Because the data presented in this paper indicate that α-amanitin and TFIIS affect NTP-Mg^{2+} isomerization and pyrophosphate release, it appears that α-amanitin and TFIIS can interact with human RNA polymerase II during catalysis. Results in the human RNA polymerase II system, therefore, are most understandable if an open trigger loop conformation supports catalysis. At this time, we do not have a clear explanation for this apparent discrepancy with the *T. thermophilus* elongation complex structure. Yeast RNA polymerase II elongation complexes with closed trigger loop conformations do not appear to be precise catalytic intermediates because of a bend in the bridge α-helix and distortions in base pairs at the i and i+1 positions (Vassylyev et al., 2007; Wang et al., 2006). To obtain x-ray crystal structures that approach a catalytic intermediate requires blocking NMP incorporation, so non-incorporatable NTP analogues or a 3'-deoxy chain terminator at the RNA 3' end are used. Structures with a closed trigger loop conformation, therefore, could represent a response to the block to NMP

incorporation. Critical mutations in the trigger loop are so far characterized most completely for transcriptional pausing, rather than for catalysis or transcriptional fidelity, so it is not yet clear whether a closed trigger loop conformation is normally utilized for bond synthesis (Toulokhonov et al., 2007).

4.3 A tunable ratchet driving translocation.

To explain the effects of α-amanitin, TFIIF, TFIIS, and downstream NTPs in adjusting translocation pressure, we propose a model for a tunable ratchet driving translocation (Figure III-6). Figure 6A shows a schematic of a tunable translocation ratchet, with the bridge α-helix forming its core. Figure III-6B shows a structural image of the bridge α-helix and a proposed conformational switch (the "funnel"; *S. cerevisiae* Rpb1 673 to 763), proposed to help roll the ratchet domain forward and back. A primary feature of the tunable ratchet model is that it posits conformational coupling between the RNA polymerase II active site (i+1) and the NTP-driven translocation mechanism, which involves templated NTP binding to downstream sites (i+2, i+3, and i+4). In this model, the bridge α-helix forms the center of a cylindrical domain that rolls forward to translocate the DNA and then slides back relative to DNA, during each bond addition step.

Tightening of the i+1 NTP substrate is posited to place strain on the DNA template stimulating forward translocation. A right angle bend in the template DNA over the bridge α-helix is suggested to link catalysis to

Figure III-6. A model for a tunable ratchet driving translocation by human RNA polymerase II. A) The bridge α-helix (yellow) is depicted as forming the core of a cylindrical domain (light blue) that rolls forward and then slides back, relative to the DNA template (green; phosphates indicated as yellow dots) during each translocation step. The "funnel" (pink) connects active site and rolling ratchet functions. Rpb2 R1020 and Rpb1 K752 are proposed to be sensors of progression of each bond addition. RNA and NTP substrates are red. Mg^{2+} atoms are green dots. Tilting of the funnel (pink), gray lines and letters indicate rotation of the ratchet domain. Rpb1 Y836 (S. cerevisiae residue number) is purple. Y836 is posited to be important in switching the bend point in the DNA template strand as it passes over the bridge α -helix. a) relaxed elongation complex; b) isomerized (tightened) elongation complex; c) complex after phosphodiester bond formation; d and e) maximally strained elongation complexes, in which NTP-driven translocation is coupled to pyrophosphate release; and f) relaxed elongation complex. α-amanitin (α A) primarily blocks the dominant translocation step, but limited translocation is suggested to be associated with multiple steps in the mechanism. B) Structural model of the rolling ratchet mechanism. Top panels: three views of the bridge α-helix (Rpb1 810 to 845) (yellow) and C-terminal cap (Rpb1 846 to 868) (yellow) and the funnel region (Rpb1 526 to 763) (mauve) are shown in relation to the transcription bubble. The proposed cylindrical ratchet domain location is indicated by white circles (left and right images) and a white rectangle (center image). Bottom panel: a detail

of the right angle bend in the DNA template as it passes over the bridge α-helix is

shown with basic residue contacts to phosphates. NTP substrates are shown at the

i+1 (orange), i+2 (pink), and i+3 (lime) positions. The i+2 and i+3 NTPs were placed

by modeling. Mg^{2+} atoms are green; the DNA template strand is green; the non-

template DNA strand is white; RNA is red; phosphorous atoms are indicated as tan

spheres; basic amino acids (Rpb1 R326, K330, K332, R337, K752, R839, R1386,

R1391, and Rpb2 R1129) are blue; Rpb1 Y836 is purple; phenylalanine residues

found at the N-terminal end of the bridge α-helix (Rpb1 813 to 815) are white. DNA

missing from available structures was placed by modeling (Burton et al., 2005).

Multiple crystal structures were used to generate this figure (Gnatt et al., 2001;
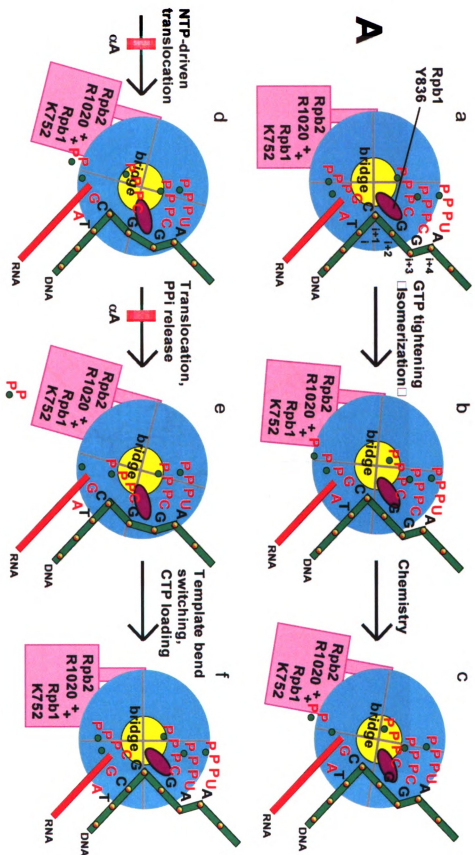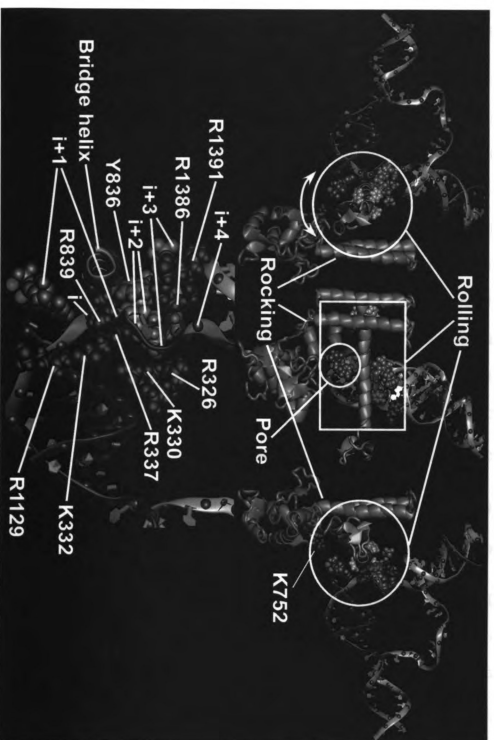
Westover et al., 2004).

Figure III-6.

Figure III-6 (continued).

translocation (Bar-Nahum et al., 2005; Gnatt et al., 2001). We posit that the

bridge α-helix is at the core of a domain that rolls forward, stimulating forward

translocation. Tightening of the i+1 NTP is posited to increase contact between

the triphosphate and basic residues Rpb2 R1020 and Rpb1 K752 (residue

numbers are for S. cerevisiae). Located on the highly conserved 750-GSKG-753

loop, which is linked to the "funnel" (Rpb1 663 to 763), Rpb1 K752 is posited to

be a major sensor of the state of the catalytic cycle (Wang et al., 2006). As

catalysis progresses, a rocking motion is developed in the funnel that stimulates

rolling of the rolling ratchet domain. Movement of accurately templated

downstream NTPs contributes to translocation force. NTP-driven translocation is

thought to be coupled to pyrophosphate release, so bond completion would be

coupled to accurate NTP driven translocation for the next bond addition. As the

next NTP is transferred into the RNA polymerase II active site, the bridge α-helix

domain rolls back, sliding relative to the DNA template (Figure III-6A).

Kornberg and colleagues suggested that the bridge α-helix might cycle

between bending and straightening to drive forward translocation (Gnatt et al.,

2001; Wang et al., 2006). Nudler, Ruckenstein and colleagues suggested a

similar dual ratchet mechanism for elongation, involving conformational changes

of the bridge α-helix and associated "G" or "trigger" loop (Bar-Nahum et al.,

2005). We favor a mechanism in which the bridge α-helix lies at the core of a

larger domain that rolls (relative to template) more than bends to drive forward

181

translocation. Rolling of the bridge α-helix forward causes Rpb1 Y836 to clash

with template DNA, helping to induce a change in the location of the 90 degree

bend in the DNA template strand over the bridge (Figure III-6A). Changing the

position of the bend in the DNA is suggested to be an important part of a

switching mechanism to alter contacts between basic groups (Rpb1 R326, K330,

K332, R337, R839, R1386, and Rpb2 R1129) and DNA phosphates (Figure III-

6B, lower panel). Altering the DNA bend point is posited to cause first releasing

and then re-setting phosphate contacts to the next downstream position,

supporting stepped translocation of the DNA.

Breaking phosphate contacts results in a reverse roll of the ratchet

domain, causing the DNA template to slide relative to the bridge α-helix. In this

way, the bridge α-helix is at the core of a molecular ratchet driving forward

translocation. We propose that larger motions of RNA polymerase II may support

the ratchet roll of the bridge domain. For instance, the funnel may rock toward

the bridge α-helix, to enhance the forward roll, and then rock back during

template sliding, during the reverse ratchet roll (Figure III-6). Because the funnel

includes Rpb1 K752, which is poised to interact with the i+1 NTP triphosphate,

this establishes a connection that is mediated through the funnel structure

between the i+1 active site and translocation driven by accurately templated

NTPs at i+2, i+3, and i+4. A functional link between the funnel and the bridge α-

helix is indicated by the structure of α-amanitin bound to RNA polymerase II,

because translocation is blocked when α-amanitin is wedged between the funnel and the bridge α-helix, where the toxin would be expected to inhibit funnel rocking and bridge helix rolling (Bushnell et al., 2002). Available genetics is consistent with the proposed "rock and roll" mechanism for RNA polymerase II translocation, because mutagenic analyses demonstrate the importance of the funnel, the cleft (proposed to be important for i+3 NTP binding), and fork loop 2 region (proposed to be important for i+2 NTP binding) in translocation (Trinh et al., 2006).

A reversibly rolling ratchet that is tightened by induced fit at the i+1 active site and by NTP-driven translocation at i+2, i+3, and i+4 downstream sites fulfills the requirements for a mechanism in which translocation tension can be adjusted. The ratchet is wound forward by the i+1 substrate NTP as it tightens in the active site. Inaccurately templated substrates fail to tighten at i+1 and, therefore, are expected to be rapidly replaced before they can engage in chemistry. The ratchet is also proposed to be driven by accurately templated NTP substrates bound at downstream sites, explaining how NTP-driven translocation can be coupled to pyrophosphate release. Elongation factors appear to increase (TFIIF) or decrease (TFIIS) tension on the ratchet through allosteric interactions with RNA polymerase II. Whether or not bacterial RNA polymerases utilize NTP-driven translocation, we posit that an induced fit, "rock and roll" conformational change may be a feature of translocation by all multi-subunit RNA polymerases.

4.4 NTP-driven translocation and fidelity.

The rock and roll mechanism supports very deliberate, stepwise translocation of the DNA template, limited to single-base pair steps. Free reversibility of RNA polymerase II between pre- and post-translocated elongation complexes has not been observed in millisecond phase kinetic studies, although a rapidly reversible thermal ratchet and spontaneous pyrophosphate release are predicted features of the secondary pore NTP loading model and the dual ratchet model. In this paper, we show that human RNA polymerase II does not release pyrophosphate spontaneously. Isomerization reversal experiments indicate that pyrophosphate release is linked to translocation, rendering pyrophosphate release coupled to accurate NTP-driven translocation a potential fidelity check point in each step of RNA synthesis. In the presence of the translocation blocker α-amanitin, the reaction can be halted after phosphodiester bond synthesis but prior to pyrophosphate release. Accurately templated downstream NTPs seem to drive forward translocation, because they appear to adjust the set point of the ratchet. These results support the NTP-driven translocation mechanism. No model requiring either rapidly reversible translocation or secondary pore NTP loading appears consistent with these data.

The primary issue in considering NTP-driven translocation is not whether NTPs enter the structure through the secondary pore, as has been generally assumed, or through the main enzyme channel, as appears to be the case for human RNA polymerase II. The essential issue is the fidelity of RNA

184

synthesis. NTP-driven translocation induced by accurately templated downstream NTPs is a high fidelity mechanism, because it provides a mechanism for pre-sorting and repeated confirmation of accurately loaded NTP substrates. For human RNA polymerase II, if NTPs are retained in the queue, NTP substrates are confirmed versus template at the i+4, i+3, i+2, and i+1 sites, prior to NMP incorporation at i+1. Remarkably, although necessary, the capacity for accurate base-pairing is not sufficient for retention in the queue. Because they fail to support isomerization reversal, very close NTP analogues such as dNTPs and NDPs appear to be rejected downstream of the active site, even when they are accurately templated. In addition to preventing inaccurate loading of NTPs to the active site, accurately templated downstream NTPs appear to generate translocation force that is used either: 1) to complete synthesis of correct bonds; or 2) to prevent and reverse transcription errors in progress. In this manner, accurately templated NTP-driven translocation appears to be coupled to fidelity mechanisms.

References

Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R., and Block, S. M. (2005). Direct observation of base-pair stepping by RNA polymerase. Nature *438*, 460-465.

Arnold, J. J., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg2+. Biochemistry *43*, 5126-5137.

Arnold, J. J., Gohara, D. W., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn2+. Biochemistry *43*, 5138-5148.

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*, 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*, 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 bubble opening by multi-subunit RNA polymerases. Biochem Cell Biol *83*, 486-496.

Bushnell, D. A., Cramer, P., and Kornberg, R. D. (2002). Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. Proc Natl Acad Sci U S A *99*, 1218-1222.

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

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

Funk, J. D., Nedialkov, Y. A., Xu, D., and Burton, Z. F. (2002). A key role for the alpha 1 helix of human RAP74 in the initiation and elongation of RNA chains. J Biol Chem *277*, 46998-47003.

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 3.3 A resolution. Science *292*, 1876-1882.

Gong, X. Q., Nedialkov, Y. A., and Burton, Z. F. (2004). Alpha-amanitin blocks translocation by human RNA polymerase II. J Biol Chem *279*, 27422-27427.

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*, 461-470.

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*, 35597-35608.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J Mol Graph *14*, 33-38, 27-38.

Kahl, B. F., Li, H., and Paule, M. R. (2000). DNA melting and promoter clearance by eukaryotic RNA polymerase I. J Mol Biol *299*, 75-89.

Kettenberger, H., Armache, K. J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347-357.

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*, 955-965.

Landick, R. (2005). NTP-entry routes in multi-subunit RNA polymerases. Trends Biochem Sci *30*, 651-654.

Lei, L., Ren, D., and Burton, Z. F. (1999). The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. Mol Cell Biol *19*, 8372-8382.

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*, 18303-18312.

Nedialkov, Y. A., Gong, X. Q., Yamaguchi, Y., Handa, H., and Burton, Z. F. (2003). Assay of transient state kinetics of RNA polymerase II elongation. Methods Enzymol *371*, 252-264.

Shapiro, D. J., Sharp, P. A., Wahli, W. W., and Keller, M. J. (1988). A high-efficiency HeLa cell nuclear transcription extract. Dna *7*, 47-55.

Temiakov, D., Zenkin, N., Vassylyeva, M. N., Perederina, A., Tahirov, T. H., Kashkina, E., Savkina, M., Zorov, S., Nikiforov, V., Igarashi, N., *et al.* (2005). Structural basis of transcription inhibition by antibiotic streptolydigin. Mol Cell *19*, 655-666.

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. Mol Cell *27*, 406-419.

Traut, T. W. (1994). Physiological concentrations of purines and pyrimidines. Mol Cell Biochem *140*, 1-22.

Trinh, V., Langelier, M. F., Archambault, J., and Coulombe, B. (2006). Structural perspective on mutations affecting the function of multisubunit RNA polymerases. Microbiol Mol Biol Rev *70*, 12-36.

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*, 157-162.

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*, 163-168.

Wang, B. Q., Kostrub, C. F., Finkelstein, A., and Burton, Z. F. (1993). Production of human RAP30 and RAP74 in bacterial cells. Protein Expr Purif *4*, 207-214.

Wang, B. Q., Lei, L., and Burton, Z. F. (1994). Importance of codon preference for production of human RAP74 and reconstitution of the RAP30/74 complex. Protein Expr Purif *5*, 476-485.

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*, 941-954.

Weilbaecher, R., Hebron, C., Feng, G., and Landick, R. (1994). Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. Genes Dev *8*, 2913-2927.

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*, 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*, 481-489.

Zaychikov, E., Denissova, L., and Heumann, H. (1995). Translocation of the Escherichia coli transcription complex observed in the registers 11 to 20: "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble". Proc Natl Acad Sci U S A *92*, 1739-1743.

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

Zhang, C., Yan, H., and Burton, Z. F. (2003). Combinatorial control of human RNA polymerase II (RNAP II) pausing and transcript cleavage by transcription factor IIF, hepatitis delta antigen, and stimulatory factor II. J Biol Chem *278*, 50101-50111.

Zhang, C., Zobeck, K. L., and Burton, Z. F. (2005). Human RNA Polymerase II Elongation in Slow Motion: Role of the TFIIF RAP74 {alpha}1 Helix in Nucleoside Triphosphate-Driven Translocation. Mol Cell Biol *25*, 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 3.3 A resolution. Cell *98*, 811-824.

CHAPTER FOUR

SUMMARY AND FUTURE PLAN

Multi-subunit RNAPs are among the largest and most dynamic enzymes known. RNAPs represent a unique class of enzyme with a characteristic subunit structure conserved in eubacteria, archaea, eukarya, and some viruses. Persistence of this structure through evolution is evidence of its importance in supporting complex life. While transcribing a DNA template, a multi-subunit RNAP behaves as a molecular motor with a translocation step that can be largely sequestered from external forces exerted on the DNA or RNA (Abbondanzieri et al., 2005; Adelman et al., 2002; Dalal et al., 2006; Davenport et al., 2000; Greenleaf and Block, 2006; Herbert et al., 2006; Shaevitz et al., 2003; Vassylyev and Artsimovitch, 2005; Wang et al., 1998; Yin et al., 1995). Yeast RNAP II is a close homologue of human RNAP II, and elongation complex structures are available for the yeast enzyme (Gnatt et al., 2001; Kettenberger et al., 2004; Wang et al., 2006; Westover et al., 2004; Westover et al., 2004). Many yeast mutant RNAPs with defects in elongation are known, providing potential insight into the transcriptional mechanism (Trinh et al., 2006). Bacterial RNAP is a more distant evolutionary homologue of the yeast and human enzymes. Relevant X-ray crystal structures are available for bacterial RNAPs (Vassylyev et

al., 2007; Vassylyev et al., 2007), and many mutant proteins have been identified in the bacterial system (Trinh et al., 2006). In studies of human RNAP II, the Burton laboratory has developed evidence for previously unknown fidelity mechanisms in transcriptional elongation related to translocation mechanisms (Gong et al., 2005; Xiong and Burton, 2007).

X-ray crystal structures have recently become available for yeast RNAP II and bacterial *T. thermophilus* RNAP elongation complexes (Gnatt et al., 2001; Kettenberger et al., 2004; Vassylyev et al., 2007; Vassylyev et al., 2007; Wang et al., 2006; Westover et al., 2004; Westover et al., 2004). In particular, *T. thermophilus* structures indicate a possible mechanism for bacterial RNAPs, although the predicted mechanism has not yet been fully vetted by functional dynamic studies. At this time, we council caution in the interpretation of these static X-ray images, particularly for such a large and dynamic protein as RNAP. Compared to other enzymes, RNAPs contain a large proportion of water, which in some cases is removed by dehydration of crystals. *T. thermophilus* is a hyper thermophile, and its RNAP is not functional at the temperatures at which structures are obtained. Some aspects of the crystal structure, therefore, may not fully represent the functional enzyme. Because of high homology among multi-subunit RNAPs, the core mechanisms of eukaryotic and prokaryotic RNAPs are expected to be very similar. Because our kinetic analysis show that human RNAP II utilizes an NTP-driven translocation mechanism (Burton et al., 2005; Gong et al., 2005; Nedialkov et al., 2003; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2003), we wonder whether bacterial enzymes also

utilize this mechanism, despite some features of structures that might be invoked to argue against this conclusion.

The inferred bacterial RNAP model could be described as a reversible isomerization, thermal ratchet translocation mechanism. "Isomerization", or substrate tightening, appears to involve closing of the trigger loop over the i+1 NTP. Based on structures, translocation and pyrophosphate release are expected to be rapid and spontaneous. NTPs have been hypothesized to load to the deeply buried active site past the "funnel", and through the "secondary pore", a solvent accessible channel. In order for an NTP to load through the secondary pore, the "trigger loop" is likely to be in an open conformation, because a closed trigger loop conformation appears to occlude the pore (Vassylyev et al., 2007; Vassylyev et al., 2007). In available *T. thermophilus* RNAP elongation complexes structures, there does not appear to be space to load an NTP substrate through the main enzyme channel, so the pore has been thought to be the sole route of NTP entry. Furthermore, the downstream DNA template is within a double helix with the non-template DNA strand in *T. thermophilus* structures, and a specific contact of β R422 to the i+1 template strand DNA phosphate maintains bubble closure (Vassylyev et al., 2007; Vassylyev et al., 2007). In *T. thermophilus*, therefore, loading of NTP substrates to downstream template sites, as postulated for human RNAP II, seemed unlikely. The i+1 substrate NTP, therefore, might be expected to load only to the post-translocated elongation complex, not the pre-translocated complex. The NTP may initially bind in a pre-insertion state that may have an open trigger loop conformation (Vassylyev et al., 2007; Vassylyev et al.,

2007). Closing of the trigger loop tightens the NTP within the active site. This transition is also described as conversion of a partially disordered trigger loop to a trigger helix configuration. Closing of the trigger loop might provide an induced fit fidelity mechanism to confirm the identity of the dNMP-NTP base pair in the active site. Mis-pairs are expected to fail to tighten, resulting in release of the rejected NTP and replacement of the NTP in the active site. Replacement of incorrectly loaded NTPs must be very rapid, because little competition by non-templated NTPs is observed for multi-subunit RNAPs, and stable NTP acquisition is surprisingly fast (Gong et al., 2005; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2005).

Yeast RNAP II elongation complex structures indicate that NTPs can occupy either an "E" (for "entry") site position or an "A" (for "acceptor") site position (Wang et al., 2006; Westover et al., 2004). Because NTPs in the E site are oriented so that they cannot form accurate base pairs to the DNA template, it is hypothesized that NTPs may flip from the E site into the A site as they enter the active site through the secondary pore. Because of the static nature of X-ray structures, however, it is difficult to be certain that an NTP is entering the active site from the E site.

Based on our current understanding, we believe some aspects of the currently dominant model for elongation might not be correct. We posit that NTPs load through the main enzyme channel, not the secondary pore, for human and bacterial RNAPs. The secondary pore has been shown to have negative electrostatic potential, and, based on electrostatics, it was calculated that NTP

loading through the secondary pore would be expected to be a rate-limiting step in elongation. NTP loading, however, is not rate-limiting for multi-subunit RNAPs. Trigger loop closing appears to be important for catalysis, but, if NTPs load through the main enzyme channel to pre-translocated template positions, the trigger loop need not open after each bond addition to admit the next NTP. According to our kinetic model, trigger loop opening would tend to occur in response to NTP mis-loading in the active site. In this way, trigger loop opening would initiate transcriptional pausing.

None of the yeast RNAP II structures obtained to date appear to precisely recapitulate a catalytic intermediate. Structures with a closed trigger loop conformation (2E2H and 2NVZ) (Wang et al., 2006; Westover et al., 2004) have a distinct bend in the bridge α-helix, which induces base pair distortions at the i (RNA 3' end) and i+1 (substrate NTP) sites. Because precise geometry is essential to accurately catalyze addition of four different substrates, these distortions do not appear to be consistent with phosphodiester bond formation (Vassylyev et al., 2007). Furthermore, these structures do not closely resemble the *T. thermophilus* elongation complex structure, with a closed trigger loop, which has characteristics expected of a catalytic intermediate. Unlike the yeast structures, the *T. thermophilus* structure has a straight conformation of the bridge α-helix and reasonable geometry of the DNA template, RNA and the NTP-Mg^{2+} for catalysis. This suggests that a closed trigger loop and a straight conformation of the bridge α-helix may be required to support catalysis. Poor geometry in yeast elongation complex structures may be due to the block to catalysis, which,

although necessary to obtain a crystal structure with a loaded NTP substrate or NTP analogue, may be sensed by yeast RNAP II as a block to isomerization (i.e. trigger loop closing) and translocation. For human RNAP II, blocks to isomerization and/or translocation appear to signal transcription errors (Gong et al., 2005; Xiong and Burton, 2007).

Yeast RNAP II elongation complex structures also appear to differ from *T. thermophilus* in the configuration of the downstream transcription bubble. All yeast RNAP II elongation complexes obtained to date have unpaired bases in the downstream region, but opening is variable (Gnatt et al., 2001; Kettenberger et al., 2004; Wang et al., 2006; Westover et al., 2004; Westover et al., 2004). One structure from the Cramer laboratory is closed at i+3. The structure is open at i+2 because it was forced open by mis-pairing the DNA template (Kettenberger et al., 2004). Other structures appear to be open to i+4 or to i+6, whether or not the non-template strand has a greater capacity to pair (Gnatt et al., 2001; Wang et al., 2006; Westover et al., 2004). At this time, the capacity of the downstream bubble to open or close has not clearly been determined for eukaryotic RNAP II. *T. thermophilus* relies on β R422 located on "fork loop 2" to close the downstream bubble (Vassylyev et al., 2007; Vassylyev et al., 2007). β R422 aligns with G493 in human and G506 in yeast RNAP II, showing that, if the downstream bubble closes for human and yeast enzymes, it utilizes distinct (and unknown) amino acid contacts. An open downstream bubble appears necessary to support NTP-driven translocation, for which there is mounting evidence in the

human system (Gong et al., 2005; Xiong and Burton, 2007; Zhang and Burton, 2004; Zhang et al., 2005).

Mutagenic studies are incomplete for analysis of the role of the trigger loop in catalysis. A recent paper describes the effects of a large number of critical mutations in the *E. coli* trigger loop that affect pause site recognition without clearly defining their activities in elongation (Toulokhonov et al., 2007). A radical deletion of the trigger loop and also a four amino acid substitution support elongation at very slow rates, showing that the trigger loop has a role in catalysis, but it is dispensable. To date, it has not been reported whether trigger loop residues predicted to contact the i+1 active site NTP (i.e. *E. coli* β' M932, R933, and H936) have an effect either on catalysis or transcriptional fidelity, as predicted if the closed trigger loop normally accompanies catalysis. This omission was unexpected because elongation is necessary to obtain data on pausing, so data on elongation rates are available, but not reported. β' T934M, from this critical region, supports wild type elongation rates, but has defects in pause and terminator recognition (Weilbaecher et al., 1994). So far, published data, therefore, supports a role of the trigger loop in transcriptional pausing, but not a direct role in catalysis. At this time, therefore, a complete picture of trigger loop function is not available. We would suggest that the trigger loop may be in a closed conformation during NTP substrate loading through the main enzyme channel. Unless a mis-pair is loaded, trigger loop contacts to the substrate NTP do not become very important for bond formation.

Based on functional dynamic studies from the Burton laboratory, the trigger loop hypothesis presents some difficulties. For human RNAP II, it appears that an open trigger loop conformation may be consistent with phosphodiester bond synthesis (the forward reaction) and endogenous pyrophosphorolysis (the reverse reaction) (Gong et al., 2005; Xiong and Burton, 2007). From overlaying structures, it does not appear that α-amanitin can bind to a structure with a closed trigger loop conformation. In experiments in which α-amanitin and NTP substrates are added to the reaction simultaneously, however, events unfold that are most consistent with α-amanitin binding to active transcription complexes that first form a bond and then may reverse the newly formed bond.

In "isomerization reversal" experiments, α-amanitin and NTP substrates are added to RNAP II elongation complexes simultaneously. RNAP II forms a single bond and then appears to be blocked during the following translocation step (Gong et al., 2005; Xiong and Burton, 2007). Very little second bond is formed, which was interpreted to mean that α-amanitin effectively blocks translocation. The presence of α-amanitin does not appear to strongly inhibit NTP loading, isomerization (NTP tightening), or bond synthesis, although all subsequent events are strongly affected, including bond completion, which requires pyrophosphate release. If isomerization requires trigger loop closing, then α-amanitin could not bind to the catalytic complex, because the closed trigger loop would block α-amanitin binding. If α-amanitin is excluded from binding, however, how does α-amanitin interfere with bond completion (pyrophosphate release)? Opening of the trigger loop after chemistry would be

expected to induce pyrophosphate release, because opening of the trigger loop would disrupt the enclosure of the active site. Furthermore, elongation complexes reverse isomerization and phosphodiester bond formation in response to the α-amanitin block (Gong et al., 2005; Xiong and Burton, 2007). If α-amanitin can only bind to an elongation complex with an open trigger loop structure, such a structure appears to support endogenous pyrophosphorolysis, which is reverse chemistry. TFIIS-mediated dinucleotide cleavage must proceed in the presence of an open trigger loop conformation, because TFIIS and a closed trigger loop would clash (Kettenberger et al., 2003; Kettenberger et al., 2004; Wang et al., 2006). This is another example of RNA polymerase II active site catalytic function in the presence of an open trigger loop. The other example is the trigger loop deletion, which is active in elongation. At this time, it is difficult to fully reconcile rapid chemical quench flow data obtained with human RNAP II with the mechanism most strongly indicated by *T. thermophilus* elongation complex structures. At a minimum, isomerization reversal experiments with human RNAP II appear to place constraints on the trigger loop hypothesis that will require further investigation to resolve.

Based on rapid chemical quench flow studies of *E. coli* RNAP and human RNAP II, done in the absence of inhibitors, further constraints are placed on the trigger loop hypothesis. For instance, isomerization must be very rapid, because NTP-Mg^{2+} becomes resistant to EDTA chelation faster than can be accurately measured using rapid chemical quench flow (<0.002 s; rate constant >1000 s^{-1}) (Nedialkov et al., 2003; Zhang and Burton, 2004; Zhang et al., 2003;

Zhang et al., 2005). If isomerization is trigger loop closing, this step must be very rapid and essentially irreversible, because EDTA quench curves (isomerization) remain above HCl quench curves (chemistry) (Xiong and Burton, 2007). If isomerization were readily reversible, EDTA and HCl quench curves must coincide, because stable isomerization cannot be maintained until chemistry is complete. For an example, compare studies of elongation catalyzed by RNA-dependent RNAP from poliovirus, done in the presence of Mg^{2+} or Mn^{2+} as the metal co-factor. In the presence of Mg^{2+} EDTA quench and HCl quench curves are indistinguishable (isomerization is readily reversible). In the presence of Mn^{2+}, the EDTA quench curve rises above the HCl quench curve (NTP-Mg^{2+} isomerization is stable prior to chemistry) (Arnold and Cameron, 2004; Arnold et al., 2004; Gohara et al., 2004).

If isomerization is rapid and essentially irreversible and represents trigger loop closing, then additional rate-limiting steps detected during elongation require explanation (Zhang and Burton, 2004). There are two rate-limiting steps identified in each bond synthesis by human RNA polymerase II: 1) phosphodiester bond synthesis (or an associated conformational change); and 2) the interval that includes translocation, pyrophosphate release, and stable NTP-Mg^{2+} loading. But NTP-Mg^{2+} loading cannot be simultaneously slow and reversible and fast and irreversible, and it is clear from EDTA quench experiments that NTP-Mg^{2+} loading and isomerization can be fast and essentially irreversible (Nedialkov et al., 2003; Zhang and Burton, 2004; Zhang et al., 2003; Zhang et al., 2005). Translocation coupled to pyrophosphate release, therefore,

appears to be a rate-limiting step in human RNA synthesis. Furthermore, this step is highly NTP concentration dependent, indicating that human RNAP II utilizes an NTP-dependent translocation mechanism. The *T. thermophilus* RNAP structure, by contrast, has been interpreted to indicate a simple thermal ratchet mechanism with NTP loading only through the secondary pore (Vassylyev et al., 2007). These conclusions are not consistent with the functional dynamics of elongation catalyzed by human RNAP II.

In a thermal ratchet mechanism, every elongation complex passes through a pre-translocated intermediate in each bond addition cycle (EC.PPi ←→ EC(pre)+PPi). By contrast, in an NTP-driven translocation mechanism, product complex (EC.PPin) must load the next templated NTP substrate (EC.NTPn+1.PPin) in order to advance (EC.NTPn+1.PPin →EC.NTPn+1+ PPi). Because pyrophosphate release and translocation are coupled in the NTP-driven translocation mechanism, the pre-translocated elongation complex (EC(pre)) without bound pyrophosphate (EC.PPi) does not form, unless the elongation complex is delayed (i.e. by NTP limitation).

Exogenous pyrophosphate, therefore, is a probe to discriminate the NTP-driven translocation mechanism from the thermal ratchet mechanism. Exogenous pyrophosphate binds to the pre-translocated elongation complex but not the post-translocated elongation complex. In a thermal ratchet mechanism, NTP substrates bind only to the post-translocated elongation complex. Therefore, in a thermal ratchet mechanism, exogenous pyrophosphate should halt or delay elongation at each bond formation, essentially independent of incoming NTP

concentration. In the NTP-driven translocation mechanism, by contrast, the elongation complex is protected from inhibition by exogenous pyrophosphate through formation of multiple bonds. Addition of exogenous pyrophosphate to elongation reactions, therefore, appears to discriminate between different models for elongation. Because an RNAP can elongate by a mixed mechanism, i.e. by NTP-driven translocation at high NTP and by a thermal ratchet at limiting NTP, pyrophosphate may permit robust elongation at high NTP concentrations and inhibit elongation at low NTP concentrations.

Using α-amanitin as a transient translocation blocker, we found that NTPs interact accurately with templated downstream DNA sites (i+2, i+3, and i+4) and result in "isomerization reversal": release of the i+1 NTP-Mg2+ substrate from the active site in response to the presence of the translocation block and accurately templated downstream NTPs (Gong et al., 2005; Xiong and Burton, 2007). In the absence of downstream templated NTPs, the i+1 NTP substrate remains largely committed to engage in chemistry. This experiment appears to be a simple proof of the NTP-driven translocation model, and, for human RNAP II, appears to falsify the secondary pore NTP loading hypothesis: that NTPs can only enter the active site to the post-translocated elongation complex through the secondary pore. For human RNAP II, NTPs appear to load primarily through the main enzyme channel to downstream DNA sites.

We have initially concentrated effort on the DNA template 5'-40-CAAAGCCTTT-49-3' (non-template strand), observing the influence of CTP and UTP on incorporation of GMP at the G44 position, utilizing both EDTA and HCl

quenching (Xiong and Burton, 2007). Both CTP (i+2 and i+3) and UTP (i+4, i+5, i+6) appear to stimulate isomerization reversal at G44 (Xiong and Burton, 2007). Effects of UTP on isomerization reversal are completely dependent on the presence of CTP, as expected from the NTP driven translocation model. We were surprised to observe UTP effects on synthesis of G44, because this indicates templated NTP binding at least to the i+4 downstream DNA position (Xiong and Burton, 2007). This observation will be confirmed using pyrophosphate block experiments, described above.

Next, we will do a similar experiment with a template such as 40-CAAAGCTTT-48 (non-template strand), to specifically discriminate effects of CTP (i+2) and UTP (i+3, i+4, i+5) on GMP incorporation at G44 (GTP is templated at i+1). The advantage of substituting this template is that i+2 and i+3 NTPs can be varied independently. We wish to determine the concentration dependence of CTP and UTP required to stimulate isomerization reversal at G44. In principle, this experiment provides information about the dissociation constant for templated NTP binding at downstream sites. We wish to determine what CTP and UTP analogues can substitute for the natural bases at the i+2 and i+3 positions. Using the 40-CAAAGCCTTT-49 template, only CTP and UTP influenced incorporation of GMP at G44 (Xiong and Burton, 2007). No analogue, however similar, would suffice to replace either CTP (templated at i+2 and i+3) or UTP (templated at i+4, i+5, and i+6) (Xiong and Burton, 2007).

In summary, the Burton laboratory has developed the most potent approaches available to elucidate the functional dynamics of human RNAP II.

Our studies have revealed details of the internal mechanism of RNA synthesis catalyzed by human RNAP II. These studies do not appear to be consistent with the idea that human RNAP II must function identically to *T. thermophilus* RNAP, which has been inferred from structural studies to utilize a simple secondary pore NTP loading, trigger loop opening and closing, thermal ratchet mechanism for elongation. Based on functional dynamic studies, however, human RNAP II might not utilize this mechanism. The Burton laboratory has developed pyrophosphate as a potent probe of the RNAP II mechanism. This approach demonstrates the NTP-driven translocation mechanism for human RNAP II and will be used in the future to further define the mechanism of elongation.

References

Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R., and Block, S. M. (2005). Direct observation of base-pair stepping by RNA polymerase. Nature *438*, 460-465.

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

Arnold, J. J., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg2+. Biochemistry *43*, 5126-5137.

Arnold, J. J., Gohara, D. W., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn2+. Biochemistry *43*, 5138-5148.

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*, 486-496.

Dalal, R. V., Larson, M. H., Neuman, K. C., Gelles, J., Landick, R., and Block, S. M. (2006). Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. Mol Cell *23*, 231-239.

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*, 2497-2500.

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 3.3 A resolution. Science *292*, 1876-1882.

Gohara, D. W., Arnold, J. J., and Cameron, C. E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): kinetic, thermodynamic, and structural analysis of ribonucleotide selection. Biochemistry *43*, 5149-5158.

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*, 461-470.

Greenleaf, W. J., and Block, S. M. (2006). Single-molecule, motion-based DNA sequencing using RNA polymerase. Science *313*, 801.

Herbert, K. M., La Porta, A., Wong, B. J., Mooney, R. A., Neuman, K. C., Landick, R., and Block, S. M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell *125*, 1083-1094.

Kettenberger, H., Armache, K. J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347-357.

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*, 955-965.

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*, 18303-18312.

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

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. Mol Cell *27*, 406-419.

Trinh, V., Langelier, M. F., Archambault, J., and Coulombe, B. (2006). Structural perspective on mutations affecting the function of multisubunit RNA polymerases. Microbiol Mol Biol Rev *70*, 12-36.

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

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*, 157-162.

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*, 163-168.

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*, 941-954.

Wang, H. Y., Elston, T., Mogilner, A., and Oster, G. (1998). Force generation in RNA polymerase. Biophys J *74*, 1186-1202.

Weilbaecher, R., Hebron, C., Feng, G., and Landick, R. (1994). Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. Genes Dev *8*, 2913-2927.

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*, 481-489.

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*, 1014-1016.

Xiong, Y., and Burton, Z. F. (2007). A tunable ratchet driving human RNA polymerase II translocation adjusted by accurately templated nucleoside triphosphates loaded at downstream sites and by elongation factors. J Biol Chem.

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

Zhang, C., and Burton, Z. F. (2004). Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. J Mol Biol *342*, 1085-1099.

Zhang, C., Yan, H., and Burton, Z. F. (2003). Combinatorial control of human RNA polymerase II (RNAP II) pausing and transcript cleavage by transcription factor IIF, hepatitis delta antigen, and stimulatory factor II. J Biol Chem *278*, 50101-50111.

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*, 3583-3595.