Regulation of RNA Transcript Elongation: Mechanisms of RNA Polymerase Pausing and Translocation

By

Pyae P. Hein

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN – MADISON

2013

Date of final oral examination: 12/10/2013

The dissertation is approved by the following members of the Final Oral Committee:

Robert Landick, Professor, Biochemistry
Samuel E. Butcher, Professor, Biochemistry
Aaron Hoskins, Assistant Professor, Biochemistry
James Keck, Professor, Biochemistry
M. Thomas Record, Jr., Professor, Biochemistry & Chemistry

Regulation of RNA Transcript Elongation: Mechanisms of RNA Polymerase Pausing and Translocation

Pyae P. Hein

University of Wisconsin

Madison, Wisconsin

Research completed in the laboratory of Professor Robert Landick

Abstract

The evolutionarily conserved DNA-dependent multisubunit RNA polymerases (RNAPs) are essential cellular enzymes and are required for gene expression and regulation in all cellular organisms. To accomplish these roles, RNAP has evolved into a complex molecular machine in which precisely orchestrated movements in a network of flexible modules mediate the four basic steps of the nucleotide-addition cycle: translocation of RNA and DNA chains through the polymerase; binding of the nucleotide triphosphate (NTP) substrate; catalysis; and pyrophosphate release. During RNA synthesis, RNAP can temporally halt nucleotide addition by pausing. Transcriptional pausing plays a central role in regulation of transcription and may involve conformational changes in RNAP due to interactions of RNAP with intrinsic signals encoded in DNA and RNA, changes in translocation register, or both. This process is further regulated by transcription elongation factors like NusA and RfaH, which enhance or reduce pausing, respectively. However, basic mechanisms of transcriptional pausing and the role of

RNAP translocation in regulation of transcript elongation are poorly understood. In this thesis, I first present an overview of what is currently known about the structural mechanisms of RNAP translocation, pausing, and the effect of regulators in these key processes (Chapter One), and then present work to investigate pause mechanisms, specifically how the pause signal is integrated/transmitted from the RNA exit channel of the enzyme to the enzyme's active site, and the contributions of RNAP translocation to pausing. In Chapter Two, I determined what components of the transcription elongation complex (EC) control the translocation register of RNAP. Our results suggest that the identity of RNA:DNA nucleotides in the active site are strong determinants of translocation bias, with the 3´ RNA nucleotide favoring the pretranslocated state in the order U>C>A>G. Transcript elongation in bacteria is regulated in part by structures that form in the nascent RNA transcript and interact with RNAP in the RNA exit channel of the enzyme. I investigated the effects of different length and type of duplexes in the RNAP exit channel on pausing and response to regulators (Chapter Three). Strikingly, 8-bp RNA duplexes are sufficient to stimulate pausing and those with less than 8-bp duplexes do not give full hairpin effect. NusA further enhances both RNA and DNA-mediated pausing, but preferentially recognizes RNA duplexes for its action. RfaH reduces oligo-mediated pause enhancement; this effect can only be overcome by increasing the complementary oligo concentration at low RfaH concentration.

Finally, I investigated how regulatory communication between the formation of secondary structure in the RNA exit channel and the active site of RNAP modulate the catalytic center function (Chapter Four). I first determined whether formation of the nascent RNA structure is different for ECs reconstituted with various mutant RNAPs by

measuring the binding affinities of short RNA oligos complementary to the exiting RNA. Consistent with predictions from the pause phenotype, an antisense RNA oligo binds to the nascent RNA weakly in ECs reconstituted with mutant RNAPs that inhibit pausing. Blocking clamp opening by cysteine pair disulfide crosslinking of the RNA exit-channel or by binding of RfaH to ECs also decreases the formation of the pause RNA structure as measured by oligo binding. Additionally, I investigated the contributions of RNAP translocation to transcriptional pausing to determine how paused hairpin RNA inhibits translocation and how transcription factors influence pausing through their effects on tranlsocation. The formation of the paused RNA hairpin in the RNA exit channel causes movement of the clamp and the flap domains, which in turn inhibits RNAP translocation and the catalysis of rapid nucleotide transfer. Transcription factors like NusA and RfaH tune and provide additional function to this communication network by affecting the clamp movements, translocation, or both functions.

In the final chapter (Chapter Five), I summarize my major findings regarding the key roles of translocation, the paused RNA hairpin formation, interactions between the paused RNA hairpin and RNAP, conformational changes in ECs that are coupled to PEC formation, and transcription factors in hairpin-stabilized transcriptional pausing. This thesis work significantly impacts the field of transcription, providing insight into mechanistic details of transcriptional pausing and describing the regulatory communication between the active site and the nascent RNA exit channel, which had not previously been established experimentally.

Acknowledgements

I would like to thank my advisor, Dr. Bob Landick, for helping me develop my scientific skills, improve my knowledge of biochemistry, and learn how to think like a real scientist. His constructive challenges, numerous ideas for improvement, and input inspired many of the improvements in this work.

I am especially grateful to my thesis committee members Drs. Jim Keck, M.Tom Record, Sam Butcher, Aaron Hoskins, and Aseem Ansari for their continued guidance, support, and advice. This work has benefited enormously from the combined wisdom and experience of Dr. Bob Landick and my thesis committee members.

I would like to offer my sincere thank to members of the Landick lab for their indispensable advice, help, and company. In particular, I am especially grateful to my very talented undergraduate student, Kellie Elizabeth Kolb, for her unconditional help, contributions, and wonderful friendship throughout the years with my research work. It was a great privilege to work with a talented lab team.

Finally, I express my deepest appreciation to my father, Kyaw Win Aung, and my mother, Nu Nu Swe, who guided me with their strong work ethic and high expectations, endured with extraordinary support, and helped me through the hardest times of my life. Their teachings are powerful tools that have enabled me to tackle many of life's challenges. I could not have done this without their unconditional love, support, and guidance. I feel very blessed for having them as my parents.

It was a very difficult time in my life when I lost my beloved mother to aplastic anemia and a deterioration of bone marrow. Her death was difficult for me to accept since there was nothing I could do to help her. It was because of my mother that I developed a

passion for biomedical research. I hope to become a biomedical researcher so that I may one day make a difference and help others during such a complicated time. Even though I could not save my mother's life, being involved in this biomedical research will give me the opportunity to save others' parents in similar situations. I thank her for instilling in me such a motivation and desire.

I am also grateful to my siblings, Wint Nandar and Naing Ko Ko Lin, who provided constant encouragement and inspiration, and celebrated our achievements in life.

I finally feel that I am on the verge of achieving my goal. To me, research is more than just understanding and applying scientific knowledge, it means having the opportunity to explore and put our dreams and imaginations into practice.

Table of Contents

List of Figures

List of Tables

Chapter One

Introduction

Part of this introduction is adapted from a published paper:

"The bridge helix coordinates movements of modules in RNA polymerase"
Pyae Hein and Robert Landick. 2010. BMC Biol. 8: 141 doi:10.1186/1741-7007-8-141.

The final section of this chapter raises key research questions that will be experimentally addressed in Chapters Two, Three, and Four

Table of Contents

1-1 Transcription cycle

The expression of genes in all living organisms begins with transcription. As a crucial step in gene expression and regulation, the synthesis of nascent RNA transcripts complementary to the template DNA by DNA-dependent RNA polymerases (RNAPs) is highly regulated. Transcription occurs in a repetitive cycle, which comprises three basic sequential steps common to all organisms in three domains of life: initiation, elongation, and termination (Geszvain, 2005). The important events in each step are described below.

1-1-1. Transcription Initiation

At the beginning of the transcription cycle RNAP must recognize a promoter DNA sequence and separate the two DNA strands to expose the transcription start site. Unlike single-subunit RNAPs like T7 RNAP, multi-subunit RNAPs require initiation factors (*e,g.,* σ-factors in bacteria and basal initiation factors such as TBP and TFIIB etc. in archaea and eukaryotes) for sequence-specific promoter recognition, DNA strand separation, and initiation complex formation (Burgess et al., 1969; Haugen et al., 2008; Mooney et al., 2005; Werner and Grohmann, 2011). The formation of the transcriptionally competent initiation complex is a multistep process necessitating multiple intermediates and structural rearrangements (Saecker et al., 2002). In bacteria, σ associates with the core RNAP to form holoenzyme ($\alpha_2\beta\beta'\omega\sigma$). The resulting holoenzyme binds to the promoter DNA and forms a closed initiation complex (RPc). The closed complex isomerizes to the open complex (RPo) in multiple steps in which opening of the promoter DNA strands occurs to expose the transcription start site (Haugen et al., 2008; Saecker et al., 2002). Initial RNA synthesis occurs with the continuous melting of the

downstream DNA and scrunching of DNA in the transcription bubble while maintaining contacts between σ and the promoter elements (Kapanidis et al., 2006). DNA scrunching provides energy to break σ-promoter interactions and facilitates the dissociation of σ during the transition to elongation. When this transition fails, RPo makes short RNA products (up to 15 nucleotide long) in a process called abortive transcription (Revyakin et al., 2006).

1-1-2. Transcription Elongation

After escaping from the promoter, the transcription initiation complex releases the σ subunit and transforms into the remarkably stable transcription elongation complex (EC) (Geszvain, 2005; Krummel and Chamberlin, 1992). The elongation phase of transcription involves the productive RNA synthesis by RNAP while translocating in single-nucleotide steps along the DNA until it reaches a termination site. Elongation proceeds at an average rate of 30-100 nt s^{-1} for tens of kilobases down the DNA template (Abbondanzieri et al., 2005; Vogel and Jensen, 1994). Each round of the nucleotide addition cycle incorporates a single nucleotide complementary to the base in the template DNA. Occasionally the movement of RNAP along the DNA is not continuous. Rather, the nucleotide addition cycle is interrupted by pause signals encoded in the DNA and RNA. Transcriptional pausing regulates RNA synthesis and plays important regulatory functions such as coupling of transcription and translation in bacteria, proper folding of nascent RNA, recruitment of regulators and forming the first step in both Rho-dependent and intrinsic (Rho-independent) termination of transcription in bacteria (Artsimovitch and Landick, 2002; Gusarov and Nudler, 1999; Landick, 2006; Landick et al., 1985; Pan

et al., 1999; Pan and Sosnick, 2006; Peters et al., 2011). For instance, pausing at the *his* site, which occurs in the histidine biosynthetic operon in bacteria, especially *E. coli*, is stabilized by a hairpin structure in the nascent RNA and aids proper loading of ribosomes as part of a transcription attenuation control mechanism (Landick, 2006; Landick et al., 1985). The current mechanistic understanding of transcriptional pausing will be discussed extensively in section 1-4.

While pausing inhibits nucleotide addition at some sites transiently, ECs can undergo additional structural rearrangements (*e.g.,* secondary structure formation in the nascent RNA or backward translocation of RNAP) resulting in long-lived paused ECs, which can be arrested for up to tens of seconds or more at some sites. The backtracking of RNAP causes the RNA 3′ end to disengage from the active site by inserting it into the RNAP secondary channel located downstream from the active site. Backtracked RNAPs are incapable of nucleotide addition and can be rescued by transcript cleavage, stimulated by regulators such as Gre factors in bacteria (Borukhov et al., 1992; Laptenko et al., 2003; Opalka et al., 2003). Pausing and elongation rate are also influenced by elongation factors (*e.g., E. coli* NusA, NusG, and RfaH) recruited to ECs shortly after promoter escape or at specific sites (Roberts et al., 2008).

1-1-3. Transcription Termination

When transcribing RNAP reaches the termination site, it releases a nascent RNA transcript, dissociates from DNA, and can restart a new round of the transcription cycle. In bacteria, there are two mechanisms of termination, intrinsic termination, which is driven by the sequences in the DNA and RNA and Rho-dependent, which requires the

Rho protein. At intrinsic terminators, the nascent RNA adopts a terminator hairpin structure with a stem rich in G-C base pairs followed by a U rich region that extends to the RNA:DNA hybrid formed in the EC. In Rho-dependent termination, the homohexameric Rho protein (in *E. coli*) binds to the emerging transcript at a rut site (Rho utilization site, usually C-rich and G-poor), translocates along the nascent RNA transcript towards the EC, and dissociates the EC at the termination site most likely by pulling the RNA out of the EC using the energy provided by ATP hydrolysis (Lowery-Goldhammer and Richardson, 1974; Peters et al., 2011; Skordalakes and Berger, 2006).

Transcription regulation can also occur at the termination level. Antitermintors can activate silenced genes by allowing some RNAPs to bypass termination sites. Antitermination systems were discovered through studies of transcription at the bacteriophage λ early promoters (P_L and P_R). At the antitermination site, factors, like bacteriophage λQ and λN proteins, modify ECs into a termination-resistant form (Rees et al., 1997; Rees et al., 1996; Roberts et al., 2008; Whalen et al., 1988; Yarnell and Roberts, 1992). The termination-resistant ECs read through the termination signal and proceed in transcribing downstream genes. Thus, regulation can occur at every step of the transcription cycle to regulate gene expression, allowing for fine-tuning of the entire process of transcription in response to specific environmental or nutritional signals.

1-2. Mechanisms of transcript elongation by multi-subunit RNA polymerases

To elongate efficiently, RNAP has evolved into a complex molecular machine in which precisely orchestrated movements in a network of flexible modules mediate nucleotide addition. This section provides a comprehensive overview of the structural

elements in RNAP responsible for function, including elements associated with the nucleic acids, the active-site (or main) channel, the secondary channel, the downstream DNA channel, and the RNA exit channel. Mechanisms of transcript elongation that are involved in nucleotide addition, catalysis, proofreading will also be discussed in detail.

1-2-1. Structural conservation of multi-subunit RNAPs

Many structural studies of bacterial and eukaryotic RNAPs and ECs reveal important insights into their overall architecture (Cramer et al., 2000; Cramer et al., 2001; Gnatt et al., 2001; Vassylyev et al., 2002; Vassylyev et al., 2007a; Vassylyev et al., 2007b; Zhang et al., 1999). Despite small organism-specific differences, the overall structural framework and function of all multisubunit RNAPs are highly conserved across three domains of life (Cramer et al., 2001; Ebright, 2000; Hirtreiter et al., 2010; Werner and Grohmann, 2011; Zhang et al., 1999). The shape of RNAP resembles a crab-craw, with dimensions of ~150 Å× 100Å ×100 Å (in bacteria) and a central cleft large enough to accommodate double-stranded nucleic acid (Figure 1-1; Hirata et al., 2008; Vassylyev et al., 2002; Zhang et al., 1999). While bacterial RNAP is comprised of five core conserved subunits ($\alpha_2\beta\beta'\omega$), archaeal and eukaryotic RNAPs have additional specialized subunits (Cramer et al., 2008). These auxiliary subunits are located on the periphery of the core structure and may play roles in assembly, stability, regulation, DNA melting and early transcription by interacting with core subunits and regulatory factors (Cramer et al., 2001; Werner and Grohmann, 2011). Overall, the relative positions and orientations of the conserved subunits within bacteria match the positions of their counterparts in eukaryotic RNAPs. Unlike eukaryotic systems, which use several classes of RNAPs

(RNAP I, RNAP II, and RNAP III) to synthesize different types of RNA, bacteria use a single type of RNAP for making all cellular RNAs (Cramer et al., 2000; Ebright, 2000).

Two major subunits of bacterial RNAP (β and β′) are homologous to those of eukaryotic RNAPs (Rpb1 and Rpb2 respectively, in yeast RNAP II). The β and β′ subunits contain highly conserved elements that cluster around the active site in three-dimensional core RNAP (Allison et al., 1985; Geszvain, 2005; Sweetser et al., 1987) with the variable elements on the periphery (Figure 1-1). The two α subunits of bacterial RNAP (homologous to RNAP II Rpb3 and Rpb11 subunits) aid in RNAP assembly and regulation. Each α subunit has two functional domains, the amino-terminal domain (NTD) and the carboxy-terminal domain (CTD), connected by a flexible linker. The α-NTD functions in RNAP assembly by providing assembly platform for binding of β and β′ subunits whereas the α-CTD facilitates transcription initiation by interacting with transcription activators and promoter UP element (Browning and Busby, 2004; Haugen et al., 2008). The smallest subunit, ω, (~6 kDa) is homologous to Rpb6 in eukaryotic RNAP II and may be involved in efficient RNAP assembly (Minakhin et al., 2001).

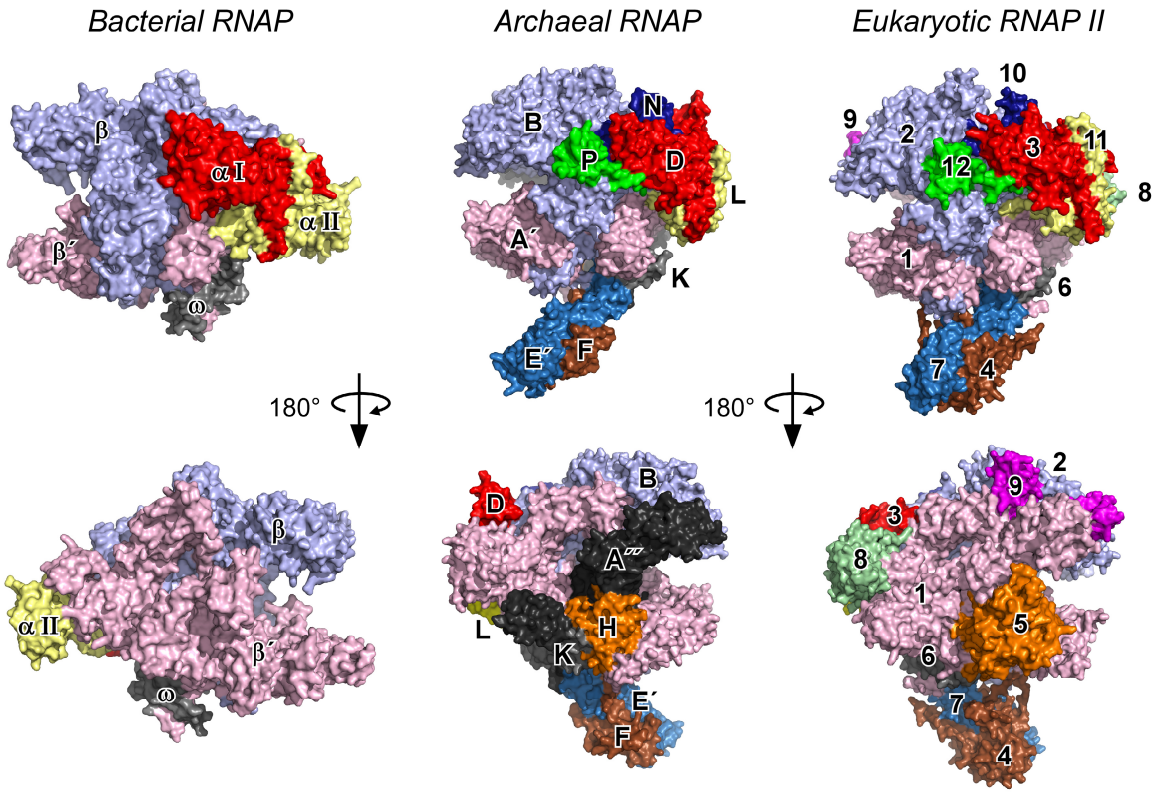


Figure 1-1. The structural conservation of RNA polymerase (RNAP) in all three domains of life

RNAP structures are built based on crystal structures of bacterial RNAP (PDB: 1I6V) (Campbell et al., 2001), archaeal RNAP (PDB: 2PMZ) (Hirata et al., 2008), and eukaryotic RNAP (PDB: 1WCM) (Armache et al., 2005). Five core conserved RNA polymerase subunits are shown as solid surfaces labeled with the subunit names. (Bacteria, α_I, red; α_{II}, yellow; β, light blue; β′, light pink; and ω, gray; archaea, D, L, B, A′, and K; and eukaryotic RNA polymerase II, Rpb3, Rpb11, Rpb2, Rpb1, and Rpb6. The same color-coding is used to compare the bacterial subunits that are homologous to those in archaea and eukaryote. All structural graphics in this thesis are prepared using PyMol (DeLano Scientific LLC, Palo Alto, CA)

1-2-2. Structural features of a bacterial elongation complex

The elongation complex of RNAP associated with nucleic acids differs structurally from the free RNAP mainly in the position of the clamp domain, which closes down near the RNA:DNA hybrid to hold the nucleic acid scaffold tightly, thereby increasing processivity of the EC (Figure 1-2). In the EC, approximately 35 bases of the DNA are protected inside the RNAP based on DNase I and hydroxyl radical cleavage (Figure 1-2A; Geszvain, 2005). The DNA duplex enters the downstream channel formed by the β′ jaw, β′ clamp, β lobe, and β protrusion, and forms the transcription bubble by melting ~ 12-17 bases to expose the template DNA to base pair with RNA (Figure 1-2B). The two DNA strands rewind at the upstream edge of the transcription bubble. The duplex DNA makes a ~90° turn and exits the RNAP. One of the key features of the EC is the formation of an 8-9 base pair RNA:DNA hybrid in the active-site cleft (the main channel) while maintaining the transcription bubble. The growing RNA strand (> 9 nt) is separated from the RNA:DNA hybrid by the RNAP lid domain located at the junction between the main channel and the RNA exit channel (Gnatt et al., 2001; Vassylyev et al., 2007a). The RNA exit channel can accommodate ~ 5 nucleotide (nt) of single-stranded RNA upstream of the RNA:DNA hybrid (Figure 1-2). The nascent RNA emerges from the RNA exit channel when the strand is longer than 15-nt.

Because temporary dissociation of RNAP from the EC leads to the premature termination and the collapse of the transcriptional bubble, RNAP has evolved to contain many elements that prevent this early termination by enclosing the DNA with its clamp and jaw domains. In stable ECs, RNAP binds tightly to the nucleic acid scaffold through non-specific interactions between positively charged amino acid residues in the binding

pockets and the negatively charged phosphate backbone of DNA or RNA (Geszvain, 2005; Gnatt et al., 2001; Vassylyev et al., 2007a). Other specific interactions between RNAP and nucleic acid sequences occur at the promoter or pause sites and cause RNAP to bind to these sequences, hindering its escape, while the non-specific interactions allow RNAP to move freely along the DNA without stalling by lowering the activation energy required for translocation. Furthermore, the nascent RNA transcript in the RNA exit channel also contributes to the overall stability of the EC.

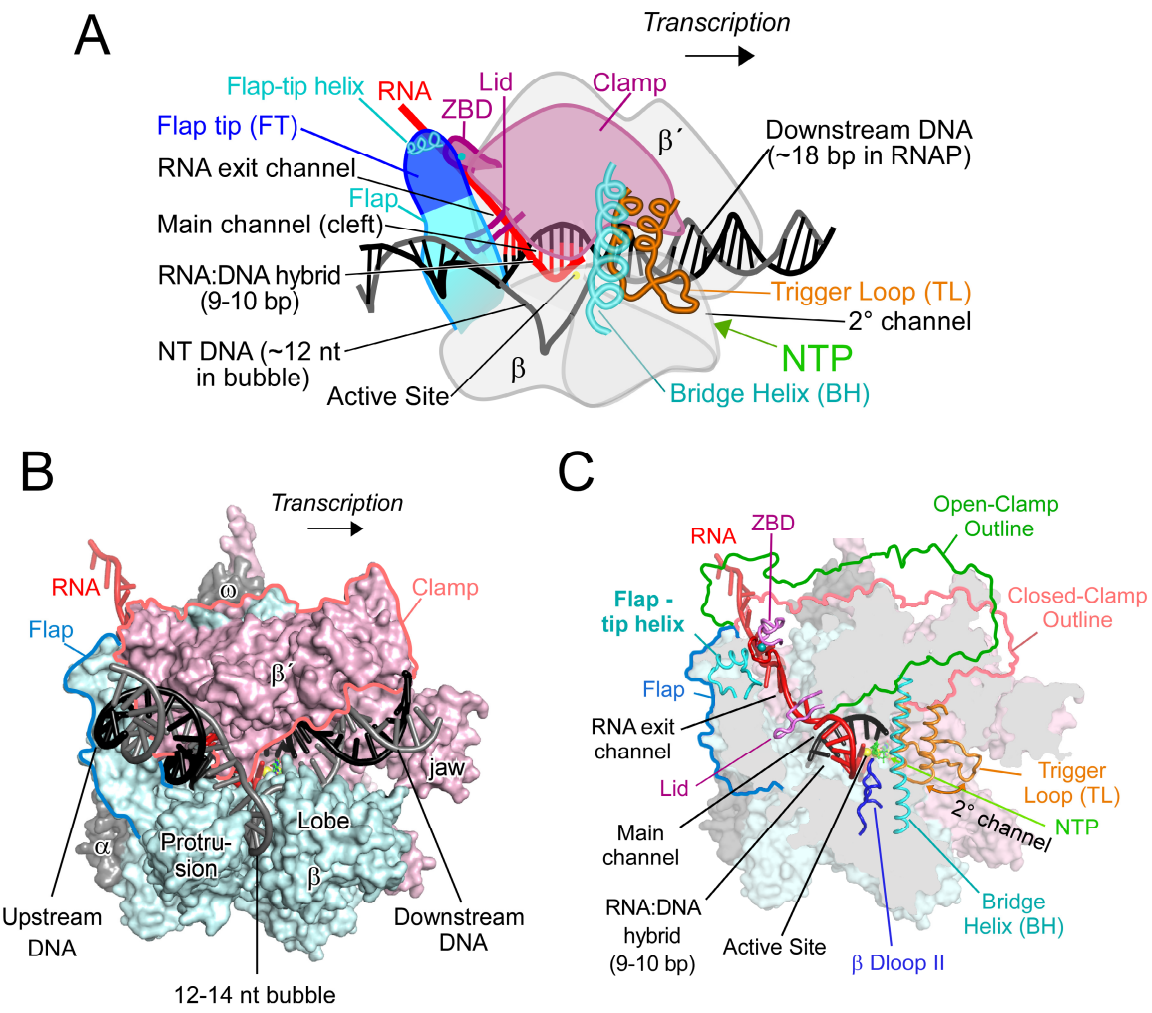


Figure 1-2. Structure of a transcription elongation complex formed by multi-subunit RNA polymerase

(A) Cartoon of a transcription elongation complex (EC) showing clamp (magenta), trigger loop (TL, orange), bridge helix (BH, cyan), lid domain (purple), β flap-tip (dark blue) that make direct contact with the exiting RNA, β flap-tip helix and β flap (light blue), zinc-binding domain (ZBD, dark purple). The features of EC, including RNA:DNA hybrid, transcription bubble, RNA exit channel, active-site channel, and secondary channel are shown.

Figure 1-2 legend. (cont.)

(B) Structure of a bacterial elongation complex based on the crystal structure of a NTP-bound RNA polymerase from *Thermus thermophilus* (PDB: 2o5j) (Vassylyev et al., 2007b) DNA (black and gray), and RNA polymerase subunits (α,ω,β,β′; gray, gray, blue, and pink, respectively) are indicated. Subunits are depicted as solid surfaces with several structural elements that are important for enzyme function including elements associated with the nucleic acid (front DNA-binding β′ clamp domain, β lobe, β protrusion, and β′ jaw) and those that cover the RNA exit channel (β flap) are indicated. The active site Mg^{2+} is shown in yellow.

(C) Cutaway view of an elongation complex with flap (blue), closed-clamp (salmon), and open-clamp (green) outlines. DNA (black) is melting into a transcription bubble that allows template-strand pairing with RNA (red) in a 9-10 base pair RNA-DNA hybrid. Components of RNAP important for function including β Dloop II (blue) are colored as in A. The bridge helix (cyan) and trigger loop/helices (orange) lie on the downstream side of the active site. Interconversion of the trigger loop and trigger helices is indicated by the orange curved arrow. The presumed path of NTP entry is indicated by the straight arrow. The active site Mg^{2+} ions are shown as yellow spheres, and α,β-methylene-ATP in green and red.

The main (active-site) channel

The main channel is highly conserved among all RNAPs and is formed by the cleft between β and β′ in bacteria. It holds nucleic acid scaffold containing the RNA:DNA hybrid. At the center of the main channel lies the active site, which is marked by two magnesium ions essential for the catalysis of nucleotide addition. The Mg^{2+} I is firmly chelated by three universally conserved aspartate residues (*E. coli* β′ D460, β′ D462, and β′ D464), and remains bound at the active site throughout the transcription cycle (K_d = ~100 μM) (Sosunov et al., 2003; Sosunov et al., 2005; Zhang, 2009). However, Mg^{2+} II is a transient ion, loosely co/ordinated by the incoming substrate nucleotide triphosphate (NTP), pyrophosphate (PPi), or transcript cleavage factors like GreB in bacteria (K_d > 10 mM; Figure 1-3). The active center also contains two moveable parts: a flexible yet conserved loop called the trigger loop (TL) and a long bridge helix (BH) that spans the active site cleft (Figures 1-2 and 1-3). The TL plays a key role in NTP loading, nucleotide addition, and regulation of transcript elongation through a major conformation change from unfolded to folded state, in which it is sometimes referred to as the trigger helices (TH). The BH plays a secondary role by facilitating the folding of the TL and translocation (described below; Toulokhonov et al., 2007; Zhang, 2009).

Another loop called the βDloop II (*E. coli* residues from 558 to 575) abuts the active site, immediately adjacent to the β′ bridge helix (Figure 1-2C; Vassylyev et al., 2002; Vassylyev et al., 2007a; Zhang et al., 1999). In paused elongation complexes (PECs), the RNA transcript 3′ end moves out of the active site (sometimes described as being frayed) so it is no longer correctly positioned for nucleotide addition. Substitutions in the βDloop II loop inhibit PEC formation by retaining the 3′ end of RNA in the active

site, resulting in pause-resistant RNAPs (Toulokhonov et al., 2007). Therefore, the βDloop II is important for EC function by modulating the pausing behavior of RNAP in bacteria (Landick et al., 1990; Toulokhonov et al., 2007). The carboxy-terminal β′ clamp domain forms one of two pincers that can close down in ECs to grasp the DNA and RNA:DNA hybrid more tightly, providing significant stability to the EC. The other pincer is composed of two mobile β domains called the lobe and protrusion. The movement of the β lobe and protrusion located opposite site of the front DNA-binding β′ clamp contribute to the opening and closing of the active-site channel (Vassylyev et al., 2002; Vassylyev et al., 2007a; Zhang et al., 1999).

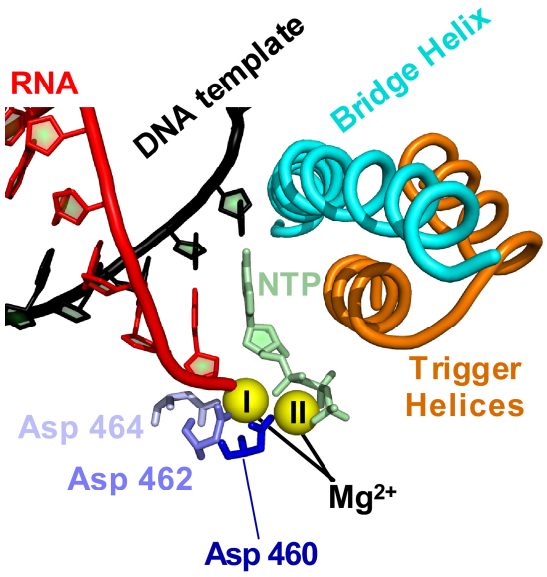


Figure 1-3. The active site of RNA polymerase showing its two Mg^{2+} ions

Mg^{2+} I is coordinated by aspartic acid residues 460, 462, and 464 (*E. coli* numbering) of β′ subunit and remains bound throughout the transcription cycle. Mg^{2+} II is a transient ion, coordinated by the substrate NTP and aspartic acid residue β′460 (Sosunov et al., 2003; Sosunov et al., 2005; Sosunova et al., 2003). Template DNA is shown in black, RNA in red, substrate NTP in light green, trigger loop in orange, bridge helix in cyan, Mg^{2+} in yellow, and Asp triad in blue, respectively.

The RNA exit channel

The RNA exit channel is located at the upstream edge of the active site channel and is approximately 25 Å long and 15 Å wide. The newly synthesized RNA emerges from the TEC through the RNA exit channel, which is formed by the β flap domain on one side and the N-terminal β′ clamp domain on the other (Ebright, 2000; Vassylyev et al., 2002; Vassylyev et al., 2007a; Zhang et al., 1999). The 14 amino acid residue loop called the β′ lid protrudes from the clamp domain at the junction between the main channel and the RNA exit channel (Figure 1-2C). Because of its location, the lid has been proposed to maintain the upstream edge of the transcription bubble and separate the exiting RNA from the RNA:DNA hybrid, contributing to regulation of transcription initiation and elongation (Toulokhonov and Landick, 2006). The β flap makes making direct contact with exiting nascent RNA and provides a docking site for extrinsic regulatory factors such as NusA, allowing it to play a major role in regulation of hairpin-dependent pausing and termination (Toulokhonov and Landick, 2003).

The Downstream DNA channel

At the downstream side of the main channel, the β lobe and β′ jaw form the downstream DNA channel, separated from the main channel by the bridge helix (Figure 1-2). The downstream DNA channel can accommodate up to 20 bp of the downstream duplex DNA based on the protection from nuclease digestion (Geszvain, 2005). This interaction between downstream DNA and RNAP contributes to EC stability and function. For example, ECs with less than 9bp of duplex in this channel are very unstable (Nudler et al., 1996), and changes in sequence of the downstream DNA duplex can affect the elongation rate (Holmes and Erie, 2003) as well as the response to pause and termination signals (Lee et al., 1990; Reynolds and

Chamberlin, 1992; Telesnitsky and Chamberlin, 1989). Furthermore, deletion of the β′ jaw causes defects in the open complex formation during initiation and in bacterial growth (Ederth et al., 2006). The Δ jaw RNAP mutant exhibits a decrease in transcriptional pausing and intrinsic termination suggesting the jaw-downstream DNA duplex interactions modulate the activity of RNAP's active site.

The secondary channel

Since the active center of the enzyme is buried 30-40 Å deep from the surface, access to the active site is limited. Furthermore, the RNA:DNA hybrid fills the main channel and the duplex DNA occupies the downstream DNA channel, thereby preventing substrate NTPs from entering to the active site through these channels. The discovery of the narrower 20 Å long and 10 Å wide secondary channel, connecting the active center to the surface of the enzyme, provides a possible solution as the potential entry route for NTPs (Figures 1-2A and 1-2C). The secondary channel is gated by opening and closing of the TL; these TL conformational dynamics may control substrate delivery and binding of transcript cleavage factors (Vassylyev et al., 2007a; Vassylyev et al., 2007b). Mis-incorporations of incorrect NTPs or any dNTPs into the nascent RNA cause RNAP to halt nucleotide addition because of the distortion introduced into the RNA:DNA hybrid (Wang et al., 2009). This instability in the RNA:DNA hybrid would force RNAP to backtrack, separating the transcript 3′ end of RNA from the RNA:DNA hybrid and threading it into the secondary channel (Cheung and Cramer, 2011). These arrested ECs can be rescued by transcript cleavage factors such as GreA and GreB in bacteria, which bind near the entrance of the secondary channel and insert their functional domain through the secondary channel up to the active site (Figure 1-4D; Borukhov et al., 1992; Borukhov et al., 1993;

Sosunova et al., 2003). Therefore, the secondary channel is also important for holding the backtracked RNA and providing the docking site for factors that maintain transcription fidelity by stimulating the cleavage of misincorporated nucleotides.

1-2-3. Nucleotide addition cycle

The addition of nucleotide to the 3′ end of the elongating RNA involves the nucleotide addition cycle, which is divided into four steps: (i) translocation of RNA and DNA chains through RNAP; (ii) binding of the NTP substrate; (iii) catalysis; and (iv) pyrophosphate (PPi) release (Figure 1-4A). The active site of RNAP is comprised of two sub-sites; the product site (also known as *i* site) and the nucleotide insertion site (*i+1* site) (Zhang, 2009). At the beginning of the nucleotide addition cycle the 3′ nucleotide (nt) is in the *i+1* site at the active center (pre-translocated). Translocation of RNA and DNA chains through RNAP results in the post-translocated state, positioning the 3′ nt into *i* site (Hein et al., 2011). The incoming cognate NTP can enter the *i+1* site together with loosely bound Mg^{2+}II and align with the 3′ OH of the growing RNA transcript, followed by S_N2-type nucleophilic attack of the RNA 3′ OH on the α – phosphorous atom of the incoming NTP occurs (Zhang et al., 2010), a mechanism shared by all polynucleotide polymerases (Steitz, 1998). The catalysis extends the RNA transcript by one nt and generates PPi, the release of which marks the end of the cycle. The re-establishment of the pre-translocated state sets the stage for the next round of nucleotide addition.

1-2-4. Reverse reactions catalyzed by RNAP

Not only does RNAP direct RNA synthesis, but it can also catalyze reactions that shorten the RNA transcript from its 3′ end (pyrophosphorolysis and transcript cleavage).

Pyrophosphorolysis

Pyrophosphorolysis is the reverse nucleotide addition reaction and can occur when the complex is in the pre-translocated register (3′ nt is in *i+1* site) by sequentially removing the 3′ nt in the form of NTP (Figure 1-4B). Pyrophosphorolysis reactions are unlikely to occur in the cell because the cellular concentrations of NTPs (~1-3 mM; apparent K_d of NTPs for the RNAP active site ~50-200 μM) and low concentrations of pyrophosphates (~1 μM; K_m of Pyrophosphorolysis ~ 0.5-3 mM) make the reaction highly unfavorable (Erie et al., 1992; Zhang et al., 2010; Zhang, 2009). However, this reaction provides a powerful tool to detect the translocation register of ECs *in vitro* because only pre-translocated complexes are sensitive to pyrophosphorolysis (Hein et al., 2011). In chapter two, I present experiments in which I used pyrophosphorolysis to determine the translocation bias of RNAPs, allowing insight into RNAP translocation on DNA and its regulation during the nucleotide addition cycle.

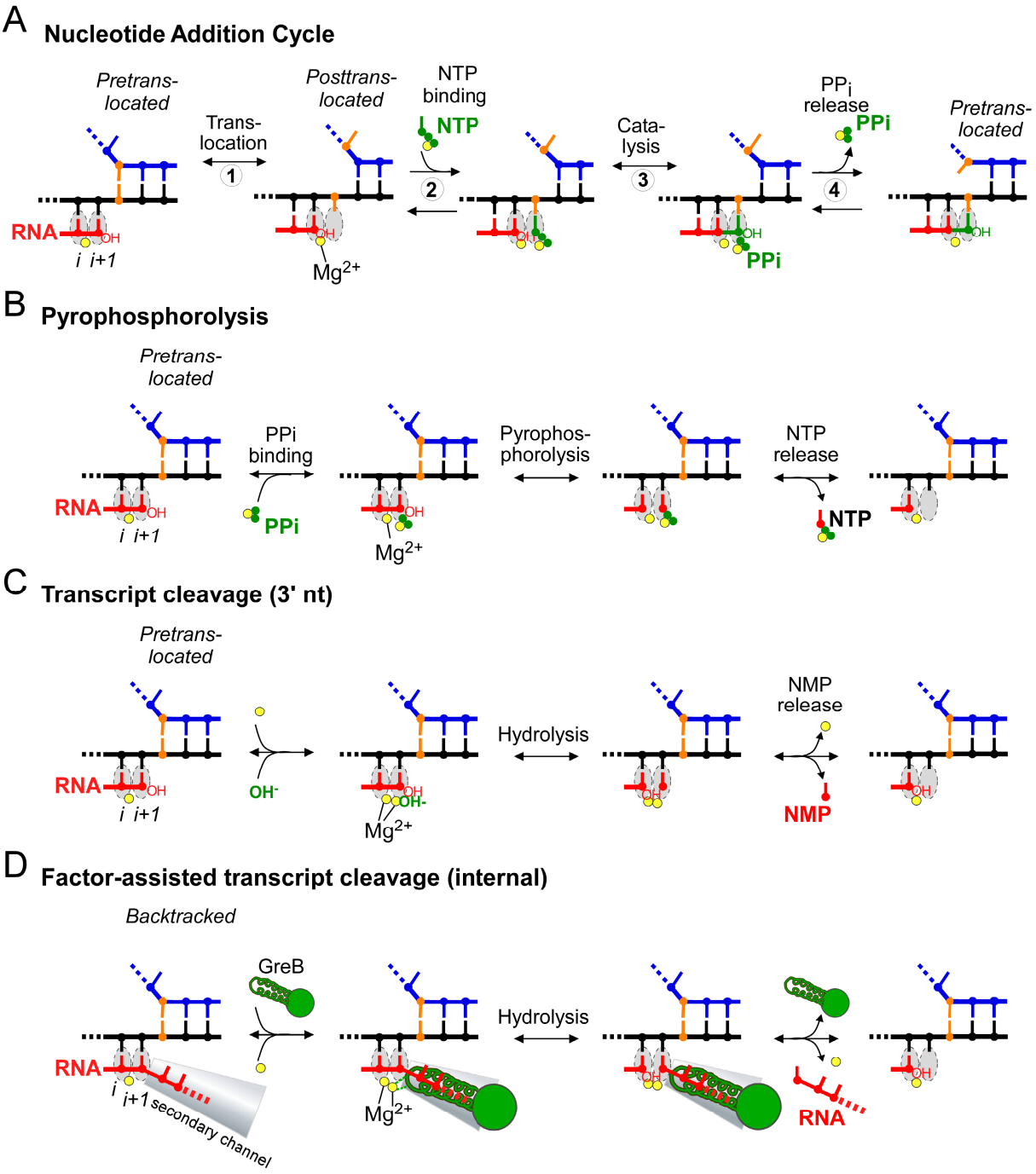


Figure 1-4. Reactions catalyzed by RNA polymerase

Figure 1-4 legend.

(A) Nucleotide Addition Cycle: translocation of RNA and DNA chains through the polymerase; binding of the nucleotide triphosphate (NTP); catalysis; and pyrophosphate (PPi) release. RNAP alternates between pre-translocated and post-translocated states during the nucleotide addition cycle. The 3′ end of the RNA has to be in the product subsite, *i* (post-translocated register) for the incoming NTP to bind in the substrate site and undergo catalysis. Black, template DNA; blue, non-template DNA; red, RNA; yellow sphere, Mg^{2+}; green, incoming NTP; orange, incoming DNA base (located at +2 position in pre-translocated state, at +1 in post-translocated state; a base at the *i+1* site is designated as +1, and those preceding +1 base are assigned negative numbers and those downstream from the +1 base are given positive numbers).

(B) Pyrophosphorolysis, the reverse reaction of nucleotide addition, can occur when the complex is pre-translocated (3′ nt is in *i+1* site), sequentially removing the 3′ nt as NTP. This reaction is triggered by the presence of high PPi concentrations and low NTP concentrations.

(C) Transcript cleavage of the 3′ nt occurs at high concentrations of OH^- (high pH) and in the presence of high concentrations of Mg^{2+}. The OH^- group initiates nucleophilic attack on the scissile phosphate. The hydrolysis of phosphodiester bond cleavage reaction releases NMP from the 3′ end of the RNA.

(D) Factor-mediated polynucleotide hydrolytic activity of RNAP. Backtracked RNA placed in the secondary channel is removed by factor-assisted transcript cleavage reaction. Bacterial Gre factors (*e.g.,* GreB; green) bind near the entrance of the secondary channel and insert a functional domain in the channel. These factors stimulate cleavage reactions by providing two highly conserved acidic residues that participate in coordination and stabilization of the weakly bound Mg^{2+} II.

Transcript Cleavage

The RNAP active center can perform another type of reverse reaction in addition to pyrophosphorolysis, transcript cleavage. Transcript cleavage differs mechanistically from pyrophosphorolysis because it hydrolytically removes one or more nucleotides from the RNA transcript 3′ end. The transcript cleavage reaction can be divided into two types depending on the position of cleavage: exonuclease activity, which removes the terminal nucleotide monophosphate (NMP) from the 3′ end of the RNA, and endonuclease activity, which cleaves the transcript internally and releases 3 to 18 nts RNA 3′- cleavage products (Figure 1-4C; Zhang, 2009). In transcript cleavage reactions, the nucleophile is the hydroxyl (OH^-) that attacks the α-phosphate group of the RNA 3′ terminal NMP. Both forward (nucleotide addition or polymerization) and reverse (pyrophosphorolysis and transcript cleavage) reactions share the same two-Mg^{2+}-mediated mechanism involving stabilization of a trigonal-bipyramidyl transition state by two Mg^{2+} ions (Sosunov et al., 2003). Because Mg-II is loosely bound at the enzyme's active center, transcript cleavage reactions typically require higher concentrations of Mg^{2+} or elevated pH (providing more OH^-). Transcript cleavage factors (GreA or GreB in bacteria and TFIIS in eukaryotes), can accelerate the rate of cleavage by more than 3000 fold (Izban and Luse, 1992; Sosunova et al., 2003). These factors binds RNAP near the entrance of the secondary channel and insert a functional domain into the channel providing the two acidic residues (Asp and Glu) to stabilize and coordinate the weakly bound Mg-II (Figure 1-4D; Izban and Luse, 1992; Laptenko et al., 2003; Sosunova et al., 2003).

Transcript cleavage reactions remove incorrect nucleotides and convert arrested complexes back to active ECs. Remarkably, the misincorporation rate of RNAP is only about once per 100,000 bases despite an elongation rate of ~30-100 s^{-1} (Erie et al., 1992; Ninio, 1991).

Misincorporation disrupts base pairing at the 3′ end of the RNA:DNA hybrid, causing RNAP to pause and then backtrack, leading to the formation of arrested complexes (Komissarova and Kashlev, 1997). Thus, transcript cleavage both provides an intrinsic proof-reading mechanism important for maintaining transcription fidelity and prevents transcriptional arrest.

1-2-5 Roles of the bridge-helix and the trigger loop in nucleotide addition cycle

Components of the RNAP active site identified as important for RNAP catalytic activities and regulation include the bridge helix (β′ 780-815) and the trigger loop (TL; β′ 928-942 and β′ 1131-1145) (Figure 1-3). Recent crystallographic, biochemical, and simulation studies have postulated that movements of the TL and bridge helix occur during the nucleotide addition cycle (Seibold et al., 2010; Toulokhonov et al., 2007; Vassylyev et al., 2002; Vassylyev et al., 2007b; Wang et al., 2006). The TL folds into the trigger helices that contact the NTP substrate in an NTP-bound elongation complex. Different TL conformations have been observed in various EC structures containing or lacking substrate NTPs, leading to the proposal that TL folding is required for nucleotide loading and catalysis supported by biochemical studies showing that restricting TL folding with double proline substitutions or locking the TL into an unfolded state by disulfide crosslinking severely inhibits nucleotide addition (Nayak et al., 2013; Toulokhonov et al., 2007).

The role of bridge helix in nucleotide addition cycle

The bridge helix is a metastable α-helix that spans the leading edge of the enzyme active-site cleft. After the first crystal structures of RNA polymerases emerged, the bridge helix garnered immediate attention as a possible effector of translocation, both because of its central

location and because in RNAPs that are not bound to DNA, it partially unfolds to form a loop that clashes with the position of the template DNA base in DNA-containing RNAP structures (Figure 1-5; (Gnatt et al., 2001). Formation of this loop is proposed to act as the pawl in a ratchet-like translocation mechanism to move DNA through the enzyme, described in more detail in section 1-3-2 (Bar-Nahum et al., 2005). Subsequent crystal structures of NTP-bound elongation complexes suggested that the bridge helix might also play a role in catalysis (Vassylyev et al., 2007b; Wang et al., 2006; Westover et al., 2004). In such structures, a continuous bridge helix forms a three-helix bundle with the neighboring TL. Because positioning of the NTP substrate by trigger-helix contacts is required for efficient catalysis, even small movements of the bridge helix, not necessarily involving unfolding, may modulate catalysis by favoring or disfavoring formation of the trigger helices (Toulokhonov et al., 2007; Vassylyev et al., 2007b).

The bridge helix as a coordinator of conformational changes in RNA polymerase

Systematic mutagenesis of the bridge helix by Weinzierl *et al.* identified critical roles of bridge-helix segments that contact flexible loops in the polymerase on either side of the active site, the downstream DNA channel, and the secondary channel, through which NTPs enter the active site. These BH segments are designated as amino- and carboxy-terminal hinges (HN and HC), on the basis of the effects of the proline substitutions that increase polymerase activity (Cα spheres in Figures 1-5). HN and HC are adjacent both to highly conserved glycines that are likely to facilitate bridge-helix distortions and to regions that do not tolerate alteration (blue in Figures 1-5). Like Pro-Gly sequences that occur at the hinge points of the trigger loop-trigger helix transition, these hinge regions may facilitate helix distortions important to RNA

polymerase function. Recently, Seibold *et al*. (Seibold et al., 2010) also proposed that helix bending at HN facilitates catalysis.

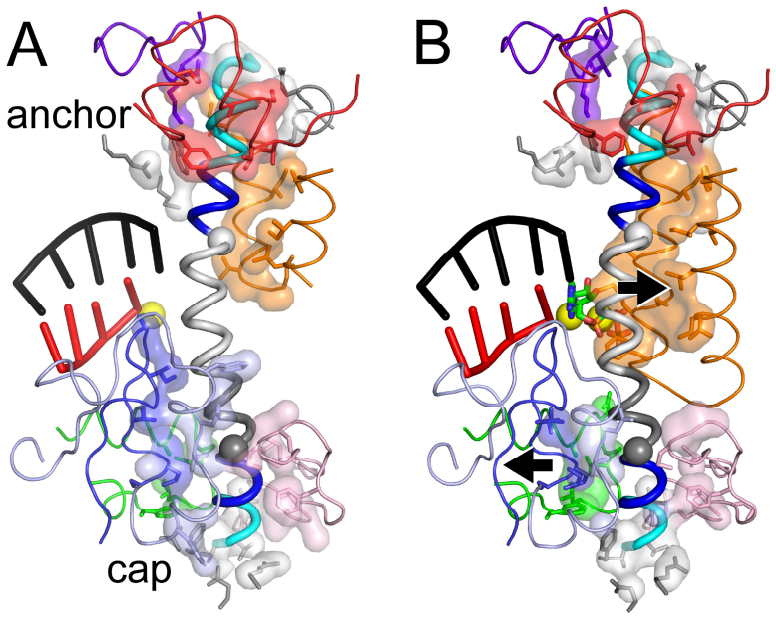


Figure 1-5. RNA polymerase residues that contact the bridge helix

RNAP active site with the DNA downstream of the active site is omitted for clarity. (A) Contacts in a *T. thermophilus* elongation complex lacking NTP (PDB: 2o5i). (B) Contacts in a *T. thermophilus* elongation complex bound by α,β-methylene-ATP (PDB: 2o5j). Residues that lie within 4 Å of the bridge helix (contacts) are shown as a semi-transparent surface and as sticks. Contacts occur principally in two regions, a cap that contacts the amino-terminal portion of the bridge helix and, in the NTP-bound complex, the trigger helices. Contacts made by polymerase loops or modules that change upon bridge-helix movements are color-coded: blue, βDloop II; light blue, β fork loop; green, β link loop and helix; light pink, β′ F-loop; red, β′ switch 1; purple, β′ switch 5 and 11 adjacent residues; orange, β′ trigger loop or trigger helices. A portion of the trigger loop in the NTP-free elongation complex that does not contact the bridge helix is not shown and was not ordered in the structure. Other segments or individual side chains contacting the bridge helix are shown but not colored. Arrows indicate small movements of the bridge helix, D-loop, and fork loop (each approximately 1.5 Å) that occur upon substrate binding coupled to a larger movement of the β lobe/ Rpb2 (Vassylyev et al., 2007b).

The HN-proximal bridge-helix segment contacts four conserved loops in the polymerase that form a cap to the helix and that, in turn, make critical contacts to: the trigger helices (β′/RPB1 F-loop; light pink in Figure 1-5) (Miropolskaya et al., 2009); the downstream fork junction of duplex and melted DNA (β/RPB2 fork loop; light blue in Figure 1-5) (Seibold et al., 2010; Vassylyev et al., 2007b); the NTP substrate (β/RPB2 D-loop; blue in Figure 1-5) (Vassylyev et al., 2007b); and the nascent RNA, especially backtracked RNA (a β/RPB2 helix and loop termed the 'link domain' by Weinzierl (Weinzierl, 2010); (Vassylyev et al., 2007b; Wang et al., 2009); green in Figure 1-5). The HC-proximal bridge-helix segment contacts the clamp domain and switch regions 1 and 5 (red and purple in Figure 1-5) in an anchor that changes conformation when the clamp changes position or upon formation of the trigger helices (Figure 1-5). When the trigger helices form, contacts of the bridge helix to the cap are reduced, consistent with movement of the central portion of the helix toward the trigger helices by 1.5 Å (Figure 1-5) (Vassylyev et al., 2007b). Although this movement is modest, larger movements of the bridge helix occur in a wedged intermediate generated by α-amanitin binding to RNA polymerase II (Brueckner and Cramer, 2008). Furthermore, it is likely that these regions undergo other, and quite possibly larger, changes during steps of the nucleotide addition cycle, including translocation, that remain to be captured by crystal structures. Thus, kinking of the bridge helix at HN and HC may allow it to coordinate conformational coupling between the two sides of the polymerase cleft in ways that remain to be elucidated.

The central role of the trigger loop dynamics in nucleotide addition cycle

The trigger loop is an evolutionarily conserved mobile element located near the bridge helix at the active center of the RNAP (Figures 1-2C, 1-3, and 1-6). The TL is disordered in

earlier RNAP crystal structures, reflecting its high degree of flexibility (Cramer et al., 2001; Vassylyev et al., 2002). However, the folded TL conformation (a folded α-helical hairpin form) has been observed in structures of bacterial and yeast ECs with substrate NTP analogs, supporting the idea that the TL undergoes dramatic conformational changes upon binding of the correct NTP (Vassylyev et al., 2007b; Wang et al., 2006). The discovery of four intermediate TL conformations in crystal structures of bacterial RNAP bound to streptolydigin (Temiakov et al., 2005; Tuske et al., 2005), TFIIS bound to yeast RNAP II (Kettenberger et al., 2004), yeast RNAPII bound to dUTP (Wang et al., 2006), and yeast RNAP II in a one nucleotide backtracked EC (Wang et al., 2009) strongly favor the view of folding/unfolding TL dynamics (Figure 1-6). Based on these observations, the TL has been proposed to oscillate between folded and unfolded conformations, which are thought to play key roles in NTP loading, substrate selection, translocation, catalysis, and regulation of elongation. TL folding brings highly conserved residues into direct contact with the substrate NTP and the bridge helix, and facilitates catalysis by properly aligning the incoming NTP for S_N2-type nucleophilic attack of the RNA 3′OH on the NTP α-phosphate atom.

Together with the crystallographic observations, biochemical experiments provided additional evidence that TL folding occurs in ECs. For example, deletion of the TL in *E. coli* or in *T. aquaticus* reduces catalysis by a factor of ~10,000, and causes defects in bacterial growth and in transcriptional pausing (Temiakov et al., 2005; Toulokhonov et al., 2007). Single amino acid substitutions within the TL profoundly affect nucleotide addition, pyrophosphorolysis, and transcript cleavage (Bar-Nahum et al., 2005; Toulokhonov et al., 2007; Vassylyev et al., 2007b; Zhang et al., 2010). Interestingly, a substitution of two consecutive prolines in the TL of *E. coli* RNAP predicted to destabilize the α-helical (or folded) conformation of TL dramatically reduces

nucleotide addition as much as ~10,000 fold, similar to the rate observed for the mutant RNAP lacking TL (Toulokhonov et al., 2007). More recent disulfide crosslinking experiments have produced a consistent picture of TL dynamics in the nucleotide addition cycle. In these experiments, TL folding and thus nucleotide addition can be drastically inhibited by crosslinking two unnatural cysteine (Cys) residues; one introduced within TL and the second in a position capable of stabilizing the TL into an unfolded or folded conformation (Nayak et al., 2013).

The TL also plays a role in the formation of the paused EC (PEC) in which the TL movement is severely restricted (detailed description is provided in section 1-4). Various alterations in TL conformation have a significant effect on pausing. For instance, removal of the TL results in an RNAP that fails to recognize the pause site and completely eliminates the ~100 fold difference in nucleotide addition rates at pause sites and non-pause sites. A Cys-crosslink reporter system used to probe different conformations of TL also shows that TL adopts a partially folded conformation in the PEC (Nayak et al., 2013). More recent single-molecule studies with yeast RNAP II showed that TL promotes transcriptional fidelity by preventing misincoporation, aiding in the recognition of misincorporation events, or the combination of both (Larson et al., 2012). Substitution of the key TL residue E1103G (yeast RNAP II numbering) causes mutant RNAP II to incorporate more incorrect NTPs into the transcript. In conclusion, TL folding and unfolding dynamics control nucleotide addition, affect transcriptional pausing, and promote fidelity. The conservation of TL dynamics in RNAPs from all three domains of life suggests that all multi-subunit RNAPs share this fundamental mechanism of nucleotide addition.

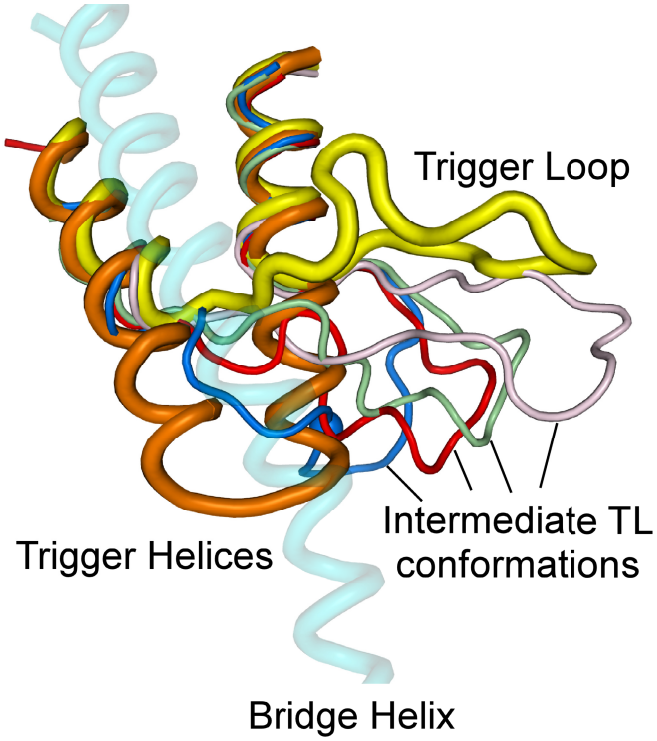


Figure 1-6. Conformations of trigger loop in an elongation complex

The folded (TH; orange) and unfolded trigger loop (yellow) from crystal structures of a NTP-bound bacterial elongation complex (PDB: 205j) (Vassylyev et al., 2007b) and of a *Thermus thermophilus* RNA polymerase lacking nucleic acids (PDB: 1iw7) (Vassylyev et al., 2002), respectively. Four intermediate TL conformations observed in crystal structures of *Thermus thermophilus* RNAP bound to streptolydigin (PDB: 1zyr; pale pink) (Tuske et al., 2005), TFIIS bound to yeast RNAPII (1y1v, pale green) (Kettenberger et al., 2004), yeast RNAPII bound to dUTP (2nvq, red) (Wang et al., 2006), and yeast RNAPII in -1 backtracked EC (3gtg, blue) (Wang et al., 2009) are shown.

1-3. The role of translocation in transcript elongation

Each round of nucleotide addition requires a translocation step in which DNA is threaded through RNAP by the distance of a single base pair, ~3.4 Å (Abbondanzieri et al., 2005). In the active site, forward translocation causes the movement of the RNA 3′ nt from the *i+1* site to the *i* site, melting of one base pair in the downstream DNA duplex, and rewinding one base pair in the upstream DNA duplex (Figure 1-4). The translocation step is one of the key steps in the nucleotide addition cycle and contributes to the rate of nucleotide addition. A study by Bar-Nahum *et al* provides insight into the effect of altering the translocation equilibrium on transcription elongation rate (Bar-Nahum et al., 2005). Two substitutions in the TL were made with opposite phenotypes: one of which (*E. coli* β′ G1136S) causes an increase in elongation rate while the other (*E. coli* β′ I1134V) decreases the rate of nucleotide addition. The alteration of the translocation equilibrium in the complexes was proposed to explain their altered elongation behaviors; ECs formed by the fast mutant appears to favor the more post-translocated state and the slower mutant favors the more pre-translocated state.

Although biochemical, structural, single-molecule, and modeling studies have yielded important insights into the mechanism of translocation, the details of the molecular mechanism and kinetic characteristics of this step are unknown, and it remains controversial whether the translocation step is rate-limiting in the nucleotide addition cycle. Moreover, it is unclear which components of the EC control translocation equilibrium and the oscillation between the pre- and post-translocation states. In chapter two, I will describe our recent finding that the sequence at the RNA 3′ end dictates the translocation bias of ECs. In the next section, I describe the models of translocation: the power-stroke model, Brownian ratchet models, allosteric translocation models, and NTP-driven translocation models (Kireeva et al., 2010; Zhang, 2009).

1-3-1. The power-stroke model

The power-stroke model for RNAP translocation was initially proposed for the single-subunit bacteriophage T7 RNAP on the basis of several crystallographic analyses (Yin and Steitz, 2004). In this model, translocation requires the chemical energy derived from the phosphodiester bond cleavage (NTP hydrolysis) and is tightly coupled with irreversible pyrophosphate (PPi) release (Figure 1-7A). The key feature of the power-stroke model is that it operates unidirectionally by converting chemical energy into the mechanical work of translocation. Thus, this model posits that translocation occurs irreversibly. However, this irreversibility is not consistent with the process of backtracking (Komissarova and Kashlev, 1997), pyrophosphorolysis (Hein et al., 2011) and transcript cleavage (Izban and Luse, 1992; Sosunov et al., 2003), all of which indicate that translocation is reversible. Although the power-stroke model is consistent with the T7 RNAP EC structures, it is currently disfavored in transcription field due to lack of support from solution biochemistry and single-molecule studies. Instead, these studies favor Brownian ratchet models (Abbondanzieri et al., 2005; Bai et al., 2004; Bar-Nahum et al., 2005; Guajardo and Sousa, 1997; Guo and Sousa, 2006).

1-3-2. Brownian ratchet models

Brownian ratchet models postulate that RNAP oscillates between pre- and post-translocated states driven by thermal fluctuations (Figure 1-7B); this conversion between pre- and post-translocated states is faster than nucleotide addition. Consistent with these models, . evidence suggests that the incoming complementary NTP can shift translocation bias toward the post-translocated state in multi-subunit RNAPs (Bar-Nahum et al., 2005; Kireeva and Kashlev,

2009; Kireeva et al., 2008; Tuske et al., 2005), T7 RNAP (Guo and Sousa, 2006), and in the HIV-1 reverse transcriptase (Marchand and Gotte, 2003)

One structural study in RNAP II EC, which captured the pretranslocated state lacking NTP substrates and NTP hydrolysis consistent with a model in NTP binding shifts EC to the post-translocated state (Gnatt et al., 2001). Another RNAP structure shows equilibrium between the pre- and post-translocated states in the crystal matrix (Brueckner and Cramer, 2008) again arguing against a power-stroke model. Many single-molecule studies support Brownian ratchet models (Abbondanzieri et al., 2005; Thomen et al., 2005; Wang et al., 1998). Abbondanzieri *et al.* used single-molecule force-clamp transcription assays to test translocation models, and observed that NTP binding affects the force dependence of translocation, consistent with models in which NTP binding is coupled to translocation. The analysis of multiple single-molecule transcription traces suggested that the measured force-velocity fits well to a modified Brownian ratchet model where the NTP binding occurs in either a pre- or post-translocated states and shifts ECs toward the post-translocated state (Figure 1-7C). Similar single-molecule studies conducted with RNAP II are in agreement with a modified Brownian ratchet model in which NTP binding can occur in a pre-translocated EC but an obvious steric clash would result (Larson et al., 2012). This discrepancy was resolved when initial NTP binding in the *E* site located downstream of the *i+1* site was observed in yeast RNAP II (Batada et al., 2004; Westover et al., 2004). An allosteric site proposed by Erie and coworkers would also satisfy the observation that pre-translocated ECs can accommodate NTP; however there is no structural evidence for an allosteric NTP-binding site.

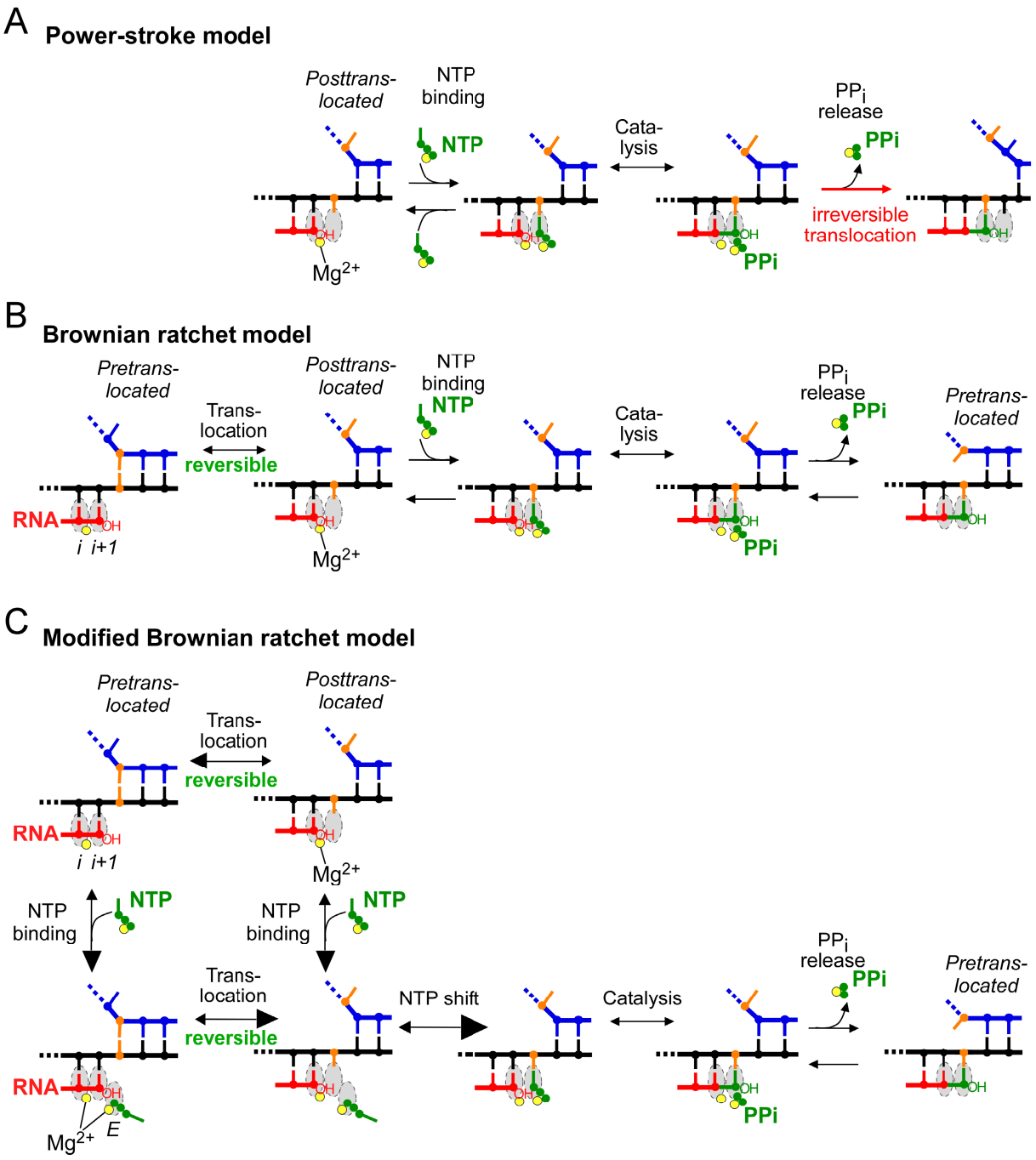


Figure 1-7. Models of RNAP translocation

Figure 1-7 legend.

(A) A power-stroke model showing irreversible translocation, which is tightly coupled to PPi release. Color coding: template DNA, black; non-template DNA, blue; RNA, red; NTP, green; the next NMP in template and non-template stands, orange.

(B) Brownian ratchet model of translocation where the oscillation of RNAP between pre- and post-translocated driven by thermal fluctuations occurs and NTP binding stabilizes the post-translocated state.

(C) A modified Brownian ratchet translocation model where NTP could bind to either the pre or posttranslocated EC. This model requires a secondary NTP binding site (*e.g., E* site located downstream from the *i+1* site) to accommodate the incoming NTP in the pre-translocated state. The size of arrow heads indicates the directionality of the equilibrium.

1-3-3. Allosteric translocation model

Earlier pre-steady-state studies established that the rate of NTP incorporation by bacterial RNAP exhibited a biphasic nature at intermediate NTP concentrations (Foster et al., 2001). This suggested the existence of two kinetically distinct species (a fast species and a slower species that do not interconvert rapidly). According to the Brownian ratchet model where the incoming complementary NTP binds only to the post-translocated ECs, a non-hydrolyzable NTP analogue would inhibit the incorporation of the substrate NTP by acting as a competitor. Instead, Erie and colleagues found that a non-hydrolyzable NTP increases transcription elongation rate. Importantly, these experiments were done with natural substrate ATP in the presence or absence of the ATP analog, AMPcPP, on a DNA template encoding two consecutive AMP residues (Foster et al., 2001). These observations led them to propose that an NTP analogue can bind to an allosteric site in RNAP and cause conformational changes of the EC from an unactivated state where nucleotide addition is slow compared to an activated state, where nucleotide addition is fast. Therefore, the fast fraction observed in the biphasic kinetic of nucleotide addition represents activated complexes and is assumed to be in a post-translocated state while the slower fraction corresponds to unactivated complexes that bias towards the pre-translcoated state (Figure 1-8A).

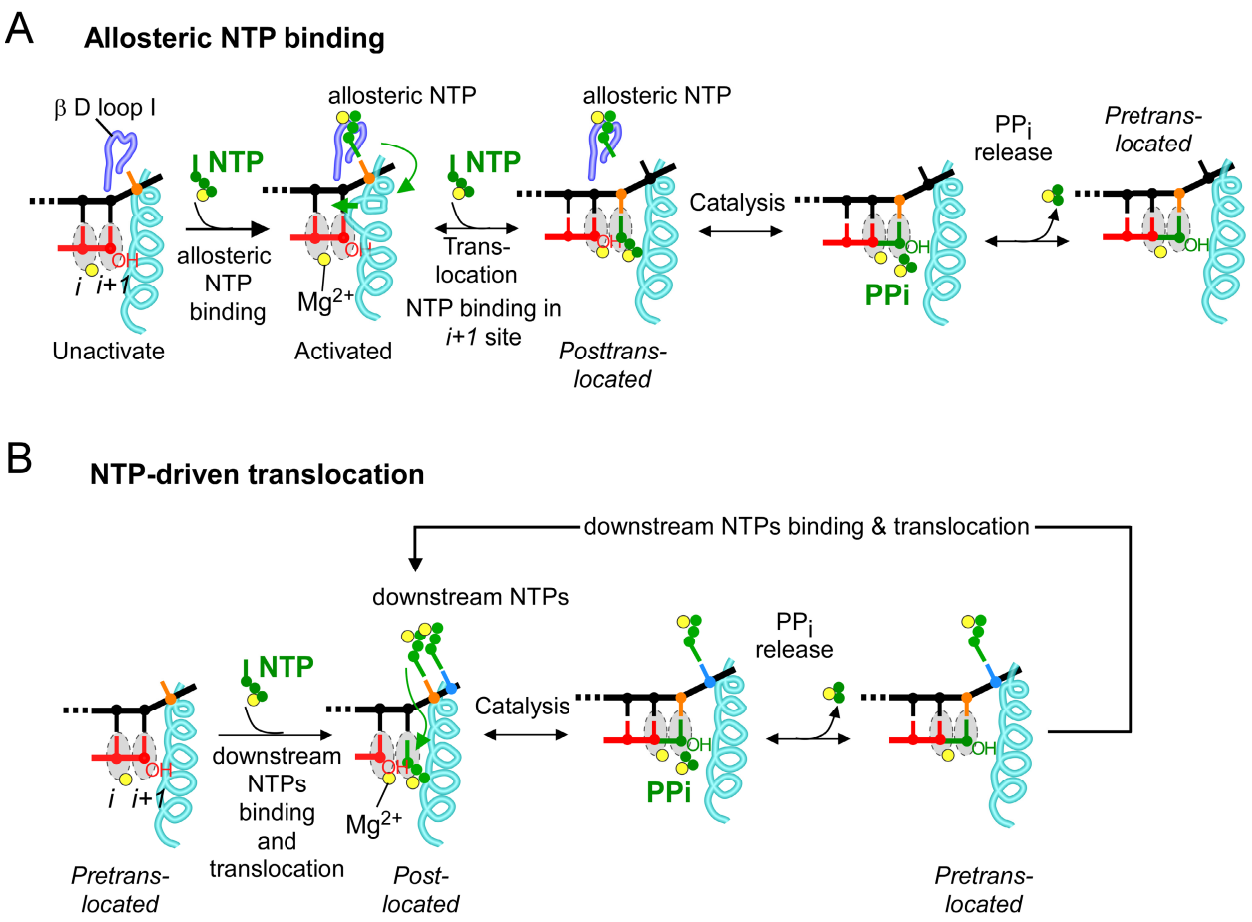


Figure 1-8. Allosteric NTP binding and NTP-driven models of RNAP translocation

(A) Allosteric NTP binding model where two NTPs can bind simultaneously to two distinct sites, the *i+1* site and the allosteric site in βDloop I (fork loop 2; blue). Allosteric NTP binding triggers local unfolding of the bridge helix (cyan) and unfolded bridge helix clashes with the RNA:DNA hybrid promoting the forward translocation of RNAP. Green arrows indicate conformational changes that are thought to occur upon allosteric NTP binding. At low concentrations of NTPs, ECs are unactivated because of the lack of allosteric NTP binding events. At high NTP, NTP binding to the allosteric site switch the EC to an activated state where nucleotide addition is fast presumably by favoring the post-translocated state. RNA is shown in red, template DNA in black, NTPs in green, next incoming template NMP in orange (Foster et al., 2001).

Figure 1-8 legend. (cont.)

(B) NTP-driven translocation model. In this model, base pairing of NTPs to bases in the downstream DNA template at positions *i+1* (orange), *i+2* (light blue) occurs in the main channel when concentrations of NTP are high. Unlike the model in (A), the NTPs bound at downstream DNA are not allosteric. The pre-hybridization of these substrate NTPs in a template-dependent manner favors translocation and increases the rate of previous NTP incorporation by facilitating the NTP loading to the *i+1* site.

Subsequent work from the Erie group established that the presence of NTP that is complementary to the downstream template base at the *i+2* position also gives similar stimulation of nucleotide addition at the *i+1* position (Holmes and Erie, 2003). According to the allosteric model, two NTPs can bind RNAP simultaneously at two different sites: an *i+2* NTP can bind an allosteric site while the incoming *i+1* NTP enters the active center to be incorporated next. An NTP bound in the allosteric site facilitates translocation and thus increases incorporation of the next incoming NTP in the *i+1* site (of the preceding NTP in the *i+1* site) (Figure 1-8A). The NTP binding to an allosteric site can occur in either pre- or post-translocated ECs because an allosteric site is distinct from the active site. An allosteric model is also consistent with the prediction from the modified Brownian ratchet model in which NTP binds RNAP regardless of translocation states requiring the existence of a secondary NTP binding site other than catalytic center (Zhang, 2009).

The conserved loop in the β subunit called the β Dloop I (533-541 *E. coli* numbering, also known as fork loop 2 or the streptolydigin-binding loop) has been suggested as a candidate for an allosteric site because of its location and resemblance to P-loops found in various nucleotide binding proteins (Figure 1-8A). Because the β Dloop I is located ~5-6 Å away from the *i+2* template DNA base, it may make direct contact with an *i+2* NTP, which maintains base pairing interaction with the *i+2* base of the template DNA. Upon binding at the allosteric site, the *i+2* NTP causes the conformational changes in the β Dloop I, which could affect the translocation because of its extensive contacts with the bridge helix and RNA:DNA hybrid (Holmes and Erie, 2003; Zhang, 2009). Although this proposal is attractive, no structure of a multisubunit RNAP with bound NTP at the proposed allosteric site has been reported, nor has

mutational analysis of the allosteric site been performed to test the idea that an NTP can bind this site and affect binding and incorporation of incoming *i+1* NTP.

1-3-4. NTP-driven translocation model

Similar to NTP incorporation kinetics by *E. coli* RNAP, the rate of nucleotide addition by human RNAP II exhibits a biphasic nature at intermediate NTP concentrations (Nedialkov et al., 2003). The fast and the slower fractions identified in nucleotide addition were simply interpreted as pre- and post-translocated ECs, respectively. Interestingly, transcription factor TFIIF, hepatitis delta antigen, and the NTP complementary to *i+2* base in the template DNA increased the amplitude of the fast fraction observed in the elongation kinetics, suggesting that these factors promote forward translocation (Nedialkov et al., 2003; Zhang et al., 2003). Similar to the effect of *i+2* NTP on *i+1* NTP addition seen with bacterial RNAP (and described above), the presence of the *i+2* NTP complementary to the *i+2* base of the template DNA accelerates the incorporation of the preceding *i+1* NTP. Based on these observations, Burton and coworkers proposed the NTP-driven translocation model (Burton et al., 2005; Gong et al., 2004; Gong et al., 2005; Nedialkov et al., 2003). As opposed to the allosteric model, NTPs base paired to the bases in the downstream DNA template at positions *i+1*, *i+2*, *i+3*, and *i+4* in the main channel are not allosteric. Instead, the pre-hybridization of these substrate NTPs in a template-dependent manner favors translocation and increases the rate of previous NTP incorporation by facilitating the NTP loading to the *i+1* site (Burton et al., 2005; Gong et al., 2005; Nedialkov et al., 2003). This model could also explain the biphasic nature of NTP incorporation kinetics, and challenges the hypothesis of NTP loading through the secondary channel because NTPs that enter the active site through the secondary channel would not have access to the downstream DNA bases.

Therefore, NTP loading is proposed to occur in a template-dependent manner by hybridizing the incoming NTPs with the downstream DNA bases in the main channel and delivering them to the *i+1* site in the active site sequentially, consistent with the template-dependent effects of incoming NTPs (Figure 1-8B). However, the crucial direct demonstration (*e.g.,* structural analysis) of NTP loading through the main channel is still lacking.

1-3-5. Translocation as a multi-step process: capturing the intermediates of a translocation step

The translocation process has been proposed to occur in a multi-step fashion because RNAP breaks and creates many networks of interactions with substrate NTP, DNA, and RNA during this step. However, it is a major challenge to detect such small movements and all intermediate steps in the real time scale of tens of microseconds. Nonetheless, high-resolution crystallographic and extensive biochemical studies with a mushroom-derived transcription elongation inhibitor, α-amanitin, support a multi-step translocation process (Brueckner and Cramer, 2008; Gong et al., 2004; Gong et al., 2005).

The report on the pre-steady-state kinetics of nucleotide addition by human RNAP II with α-amanitin established that α-amanitin inhibits translocation without affecting NTP binding (Gong et al., 2004). In agreement with results obtained from biochemical approaches, a translocation intermediate was observed in the crystal of yeast RNAP II EC with α-amanitin (Brueckner and Cramer, 2008). In this structure, the RNA:DNA hybrid is in the post-translocated position while the downstream DNA adopts the intermediary between the pre- and post-translocated states. The next template DNA base (+1 base; colored magenta; in Figure 1-9) is trapped above the bridge helix by stacking against the bridge helix residues (Ala 832 and Thr

831; yeast RNAP II numbering). The +1 template base is located at the +2 position in the pre-translocated complex before the translocation process. The central part of the bridge helix is shifted upward and occupies the portion of the *i+1* site, thereby inhibiting the +1 base from entering to the *i+1* site. The trigger loop is in a conformation called "a wedge" that stabilizes the shifted bridge helix. Based on these observations, Brueckner and Cramer proposed a two-step translocation model (Brueckner and Cramer, 2008). During the first step, the RNA:DNA hybrid translocates from the pre- to the post-translocated position, and the downstream DNA moves until the next DNA template base (+1 template base) interacts with bridge helix residues leading to the formation of the translocation intermediate. During the second step, the wedge trigger loop relaxes and the bridge helix straightens causing the +1 DNA base to rotate 90° to reach the *i+1* site. This twisting of the +1 DNA base induces the flipping of phosphate between the +1 and +2 DNA base, allowing the movement of the downstream DNA to the post-translocated position (Figure 1-9). Although structural studies are useful, their inability to recapitulate changes occurring during the translocation event in real time has limited our understanding of the transcription mechanism at the molecular level. Currently, work with molecular dynamics simulations is attempting to provide insights into the dynamic nature of the translocation event at atomic resolution and at millisecond timescales.

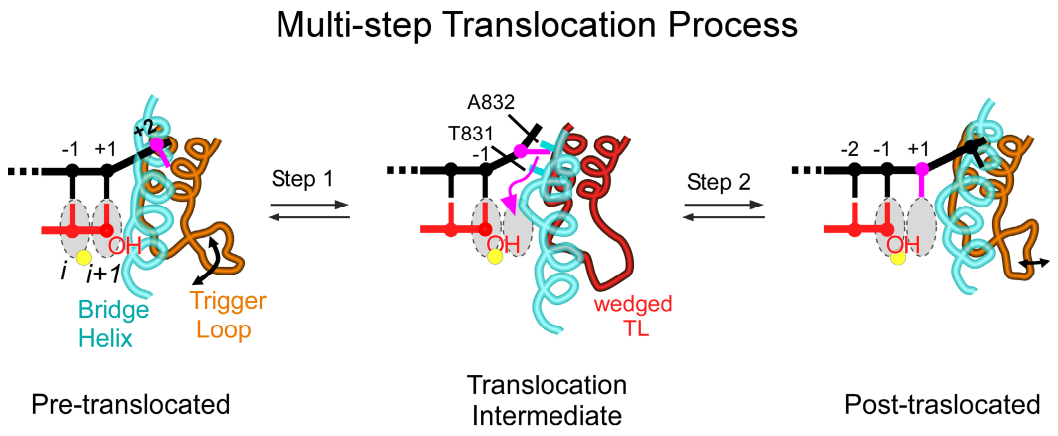


Figure 1-9. A two-step mechanism of translocation model

In this model, forward translocation of RNAP during the nucleotide addition cycle is proposed to occur in two steps. During the first step, the RNA:DNA hybrid translocates from pre- to the post-translocated position along with the downstream DNA until the next incoming DNA template base (+2 base in the pre-translocated state and +1 in the post-translocated state; magenta) contacts two bridge helix (cyan) residues (T831 and A832, yeast RNAP II numbering) through base stacking interactions leading to the formation of the translocation intermediate (Brueckner and Cramer, 2008). In the translocation intermediate state, the central part of bridge helix shifts upwards occluding the portion of *i+1* site trigger loop, restricting the NTP entry into the *i+1* site. The trigger loop (orange) is in a conformation called "a wedge" (red) that stabilizes the shifted bridge helix. During the second step, the wedge trigger loop (red) relaxes and the bridge helix straightens resulting in the 90° rotation of the +1 DNA base to reach the *i+1* site (magenta arrow), which is followed by the movement of the downstream DNA to the post-translocated position.

1-3-6. Effects of transcription factors on RNAP translocation

When RNAP enters the elongation phase of transcription, general or gene-specific transcription elongation factors are recruited to the EC. These factors bind to RNAP and may trigger conformational changes in RNAP that could potentially affect transcription elongation by either increasing or decreasing the processivity, elongation rate, termination, or the enzyme's propensity to fall into the paused state (Roberts et al., 2008).

The multidomain NusA protein is highly conserved among bacteria and also found in archaea (Roberts et al., 2008). Crystallographic studies of the *Mycobacterium tuberculosis* NusA reveal the domain structure: (i) an N-terminal domain that interacts with RNAP; (ii) three global domains, S1, KH1, and KH2; and (iii) a C-teriminal autoinhibitory domain (Beuth et al., 2005) (Gopal et al., 2001). NusA regulates transcription by prolonging transcriptional pausing, increasing intrinsic (RNA hairpin-dependent) termination, and modulating some Rho-dependent termination. Interestingly, the N-terminal domain of NusA (NusA-NTD) is sufficient for the effect on pausing, but the NTD together with the other domains is required for effects on termination (Ha et al., 2010). Tethered chemical protease and protein-RNA crosslinking experiments and 3D reconstruction electronmicroscopy analysis suggest that NusA contacts RNAP near the RNA exit channel around the β-flap domain (Ha et al., 2010; Yang et al., 2009). By using single-molecule force-clamp transcription assays, Zhou *et al.,* reported that NusA reduces the rate of nucleotide addition while increasing RNAP propensity to pause significantly (Zhou et al., 2011). Notably, NusA's inhibitory effect on transcript elongation can be recapitulated by applying opposing force of ~19 pN on transcribing RNAPs in the absence of NusA. Based on these observations, NusA was proposed to delay translocation of RNAP during nucleotide addition, suggesting NusA causes EC bias towards the pre-translocated state.

NusG is the only known universally conserved bacterial transcription elongation factor, homologous to archaeal Spt5 and eukaryotic SPT5 proteins (Roberts et al., 2008; Werner and Grohmann, 2011). NusG associates with RNAP via interactions between an N-terminal NusG (NGN) domain and the RNAP clamp coiled-coil domain. Once bound, NusG stimulates a closure of the clamp domain, thereby increasing processivity and the elongation rate (Hirtreiter et al., 2010; Mooney et al., 2009). The clamp coiled-coil domain is also the target for the NusG paralogue RfaH, a transcription factor that inhibits transcriptional pausing and termination. RfaH is recruited to EC at the operon polarity suppressor (*ops*) site via interactions with RNAP and the *ops* sequence on the nontemplate strand (Artsimovitch and Landick, 2002; Belogurov et al., 2009; Belogurov et al., 2007). Upon interaction of RfaH with the *ops* site, a dramatic all-α to an all-β refolding of the RfaH-CTD occurs, which unmasks the RNAP binding site in RfaH's NTD, allowing RfaH-NTD to bind to RNAP (Burmann et al., 2012). Once recruited to ECs, RfaH suppresses pauses, decreases transcription termination at a subset of intrinsic terminators, and inhibits rho-dependent termination by excluding NusG from the EC, and facilitates translation via RfaH-CTD/ribosomal protein S10 interaction (Sevostyanova et al., 2011).

It has been shown that NusG increases transcription elongation by promoting forward translocation of RNAP, therefore favoring the post-translocated state in single-molecule studies (Herbert et al., 2010). Similarly, RfaH has been proposed to shift the translocation equilibrium towards the post-translocated state (Svetlov et al., 2007). Although these studies have yielded insights into the mechanistic contributions of various transcription elongation factors to the translocation, a direct measurement of rapid translocation in real time is essential to understanding the translocation process. In chapter four, I will provide a mechanistic model of

the translocation process and its regulation by transcription elongation factors using recently described fluorescence-based assays (Malinen et al., 2012).

1-4. Regulation of Transcript Elongation by Pauses

Transcriptional pausing plays key regulatory roles in gene regulation by controlling RNA synthesis. In bacteria, pausing coordinates the coupling of transcription and translation and prevents premature termination by allowing a translating ribosome to catch up to a transcribing RNAP (Landick et al., 1985). Pausing is an initial step in both Rho-dependent and intrinsic (Rho-independent) termination. Pauses also facilitate the proper folding of nascent RNA (Pan et al., 1999; Pan and Sosnick, 2006) and allow the recruitment of regulators to ECs at specific sites. For instance, pausing halts RNAP to ensure binding of RfaH to a paused RNAP at *ops* (operon polarity suppressor) sites in the leader regions of RfaH-regulated operons (Artsimovitch and Landick, 2002) and of λQ to promoter-proximally paused RNAP (Roberts et al., 1998). In eukaryotes, pausing by human RNAP II modulates activator protein Tat-mediated HIV-1 transcription (Palangat et al., 1998). Additionally, specific pause sites are important for polyadenylation (Park et al., 2004; Yonaha and Proudfoot, 1999) and splice site selection in alternatively spliced mRNAs (de la Mata et al., 2003; Robson-Dixon and Garcia-Blanco, 2004).

1-4-1. The mechanism of transcription pausing

A detailed mechanistic understanding of transcriptional pausing is incomplete. Multiple lines of evidence suggest that transcriptional pausing is a branched mechanism and involves the initial rearrangement of RNAP into an off-pathway state called a short-lived or elemental pause that delays nucleotide addition (Landick, 2006). The probability that RNAP will isomerize into

the pause conformation strongly depends upon the kinetic competition between the on-line elongation pathway and the off-line pausing. At a pause site, some fraction of RNAPs enters the pause state and escapes back to the elongation pathway at a slow rate (Figure 1-10). Thus, pausing is characterized by the pause efficiency, the fraction of RNAPs that recognize the pause signal and enter the pause state, and the dwell time ($1/k_e$ where k_e is the rate of escape) or the pause half-life ($\ln 2/k_e$) of the paused EC. Single molecule and solution biochemical studies revealed that the elemental pauses derive from an alteration in the active site rather than from backtracking or hypertranslocation of RNAP (Kireeva and Kashlev, 2009; Neuman et al., 2003; Toulokhonov et al., 2007). In these studies, bacterial RNAP pauses approximately once every 100-200 bp for durations of 1-6 s and these pauses are not affected by application of force to either assist or resist translocation that would either prevent or promote backtracking (Landick, 2006; Neuman et al., 2003). A recent study characterized one pause site, the T7 A1 D111 C37 pause, at which *E. coli* RNAP halts nucleotide addition without affecting translocation or NTP binding. The C37-paused RNAP is not backtracked and appears to result from the isomerization of the enzyme active site, consistent with an elementary pause (Kireeva and Kashlev, 2009; Landick, 2009). These studies support the model that the altered RNAP active site found in the elemental pause may serve as a common precursor to all pauses.

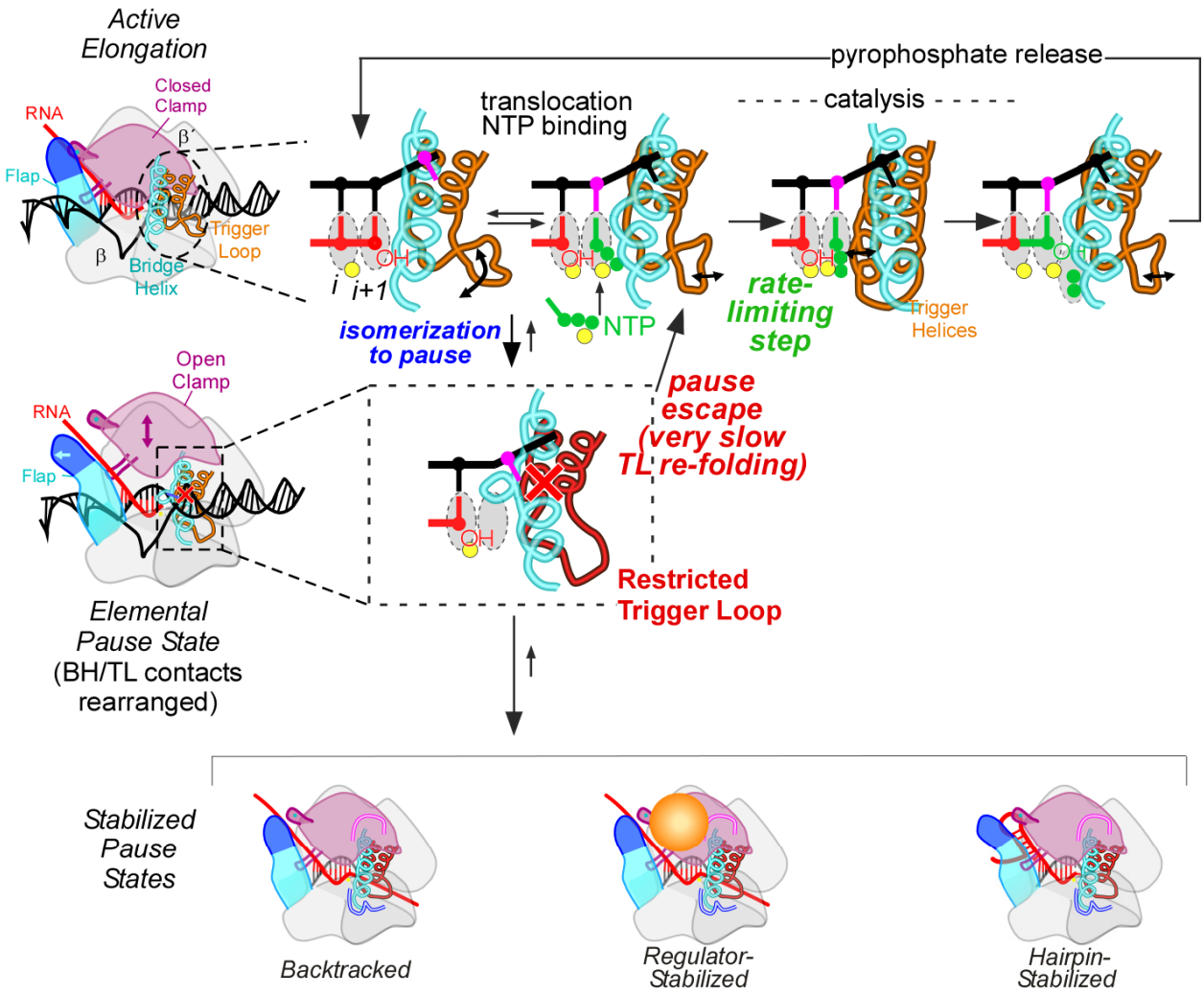


Figure 1-10. Mechanism of transcriptional pausing

A cartoon model of a transcription elongation complex with closed-clamp (purple) conformation (top left). The blow-up shows the active site of an active EC and folding and unfolding of the trigger loop (orange) into trigger helices during the nucleotide addition cycle (top right). The elemental pause elongation complex (ePEC) has an open-clamp conformation (middle left) and transiently inhibits nucleotide addition (middle right). Subsequent arrangements of the ePEC such as backtracking of RNAP (bottom left), interactions with regulators (bottom middle) or

RNA secondary structure formation within the nascent transcript (bottom right) stabilizes the ePEC leading to long-lived pauses.

Once RNAP enters the elemental pause state, subsequent arrangements or interactions such as backtracking of RNAP or the secondary RNA structure formation in the transcript occur in ECs leading to long-lived pauses that increase pause duration (Landick, 2006; Toulokhonov et al., 2001) (Figure 1-11). Three classes of stabilized pauses have been characterized to date: (i) RNA hairpin-stabilized pause, (ii) backtrack pauses, and (iii) non-backtrack pauses (Kireeva and Kashlev, 2009). The first class of stabilized pauses is the hairpin-stabilized pause which has been observed only in prokaryotes and requires formation of a hairpin structure in the nascent RNA transcript (Artsimovitch and Landick, 2000; Chan and Landick, 1993; Toulokhonov et al., 2001). The hairpin-stabilized pause is well characterized in *E. coli*, and therefore makes a good model system for studying transcriptional regulation by pausing. The structural mechanism underlying the hairpin-stabilized pause will be discussed in detail in the next section. A second example of a stabilized pause, backtrack pauses, has been found in both prokaryotes and eukaryotes, and occurs when a weak RNA:DNA hybrid causes backtracking of RNAP in order to reestablish a more stable transcription bubble and RNA:DNA hybrid (Komissarova and Kashlev, 1997; Palangat et al., 1998). One example of a backtrack pause is the pausing of human RNAP II at +62 position on the HIV-1 DNA template. This pause allows formation of an alternative RNA structure (TAR) required for recruitment of the Tat activator protein (Palangat et al., 1998). A third class of pauses arise from the active site rearrangement leading to misalignment of the incoming NTP to the 1+ template DNA base, as described after addition of C37 on the T7 A1 D111 template by *E. coli* RNAP in the previous section (Kireeva and Kashlev, 2009).

1-4-2. RNA hairpin-stabilized transcriptional pause

Hairpin-stabilized pausing occurs in the leader regions of enterobacterial amino acid biosynthetic operons to synchronize the transcription and translation during transcriptional attenuation (Landick, 2006). Transcriptional pausing at these regions allows proper ribosome loading and provides an elegant feedback control system to ensure that cells need to make that particular amino acid at that moment. Once translating ribosomes catch up to RNAP, they disrupt the paused RNA hairpin and cause RNAPs to escape from the pause. One of the best studied and understood hairpin-stabilized pauses is the *his* pause, found at the beginning of the histidine biosynthetic operon in *E. coli*. At the *his* pause, a secondary RNA hairpin structure forms in the nascent RNA and interacts with RNAP (Figure 1-10). This interaction inhibits nucleotide addition and may require (or induce) conformational changes in RNAP (Toulokhonov et al., 2007). Like other pauses characterized to date, kinetic competition of RNAPs between active elongation and pausing occurs so that only a subset of RNAPs to recognize the pause. The fraction of transcribing RNAPs that recognizes the pause enters the off-line elementary pause state through alterations in active site conformation. The paused fraction escapes back to the active elongation pathway at a slow rate. The *his* pause signal is multipartite: the RNA: DNA hybrid, the 2-3 nt spacer between the hybrid and the hairpin, the bases in the active site, and downstream DNA all contribute to the active site rearrangement in the elemental pause (Chan and Landick, 1993; Chan et al., 1997). Together with these elements, the pause RNA hairpin further stabilizes the pause state resulting in a long-lived pause (Chan et al., 1997).

Structural basis of hairpin-stabilized transcriptional pausing

A recent structure of a bacterial elementary pause elongation complex (ePEC) reveals rearrangements of the protein and nucleic acid components in the active site and a movement of the clamp domain (Weixlbaumer et al., 2013). In the crystal structure of ePEC, the clamp domain adopts an open conformation similar to open-clamp structures observed for nucleic acid free core RNAP (Zhang et al., 1999) and the EC bound to a transcription factor (*T. thermophilus* Gfh1) (Tagami et al., 2010), in contrast to the closed-clamp conformation observed in ECs (Figure 1-11; Vassylyev et al., 2007a). These structural observations, consistent with fluorescent resonance energy transfer studies of bacterial RNAP in solution, suggest that the open-clamp state is predominant for the core enzyme, whereas the active EC exists in the closed-clamp state (Chakraborty et al., 2012). Based on these observations, a model for transcriptional pausing that involves clamp movements was proposed (Figure 1-11A). In this model, opening of the clamp is required for ePEC formation, and results in a loss of contacts to the nucleic acid scaffold, especially the DNA template and kinking of the bridge helix (Weixlbaumer et al., 2013). The kinked BH sterically blocks the +1 template DNA base and the substrate from entering the *i+1* site, leading to inhibition of nucleotide addition (Figure 1-10). As a result of the steric clash, the +1 DNA base sits in a position near the side of the BH. The RNA:DNA hybrid in the ePEC closely resembles the hybrid of the translocation intermediate structure obtained with the yeast RNAP II EC in the presence of the α-amanitin (Brueckner and Cramer, 2008).

The ePEC is post-translocated, in which the 3′ end of RNA occupies the *i* site. However, the access to *i+1* site is sterically blocked by the kinked bridge helix as a result of the clamp

opening (Figure 1-10). The clamp opening also causes the widening of the RNA exit channel, which may provide space for the formation of the pause RNA hairpin. Thus, formation of a hairpin in the RNA exit channel is proposed to stabilize the widened exit channel and the open-clamp conformation (Figure 1-11B). Additionally, the formation of a pause RNA hairpin may shift the RNA:DNA hybrid towards pre-translocated state, but the kinked BH in the PEC that sterically blocks the *i+1* site forces the RNA 3′ end to occupy an alternative position outside of the *i+1* site by fraying out of the active site. Biochemical studies support the idea that the ePEC is likely pre-translocated and that RNA pause hairpin formation and active site rearrangements fray the RNA 3′ end out of the active site (Toulokhonov et al., 2007). Consistently, the pause RNA hairpin prolongs the pause dwell time without affecting the pause efficiency or the formation of the elementary pause (Chan et al., 1997).

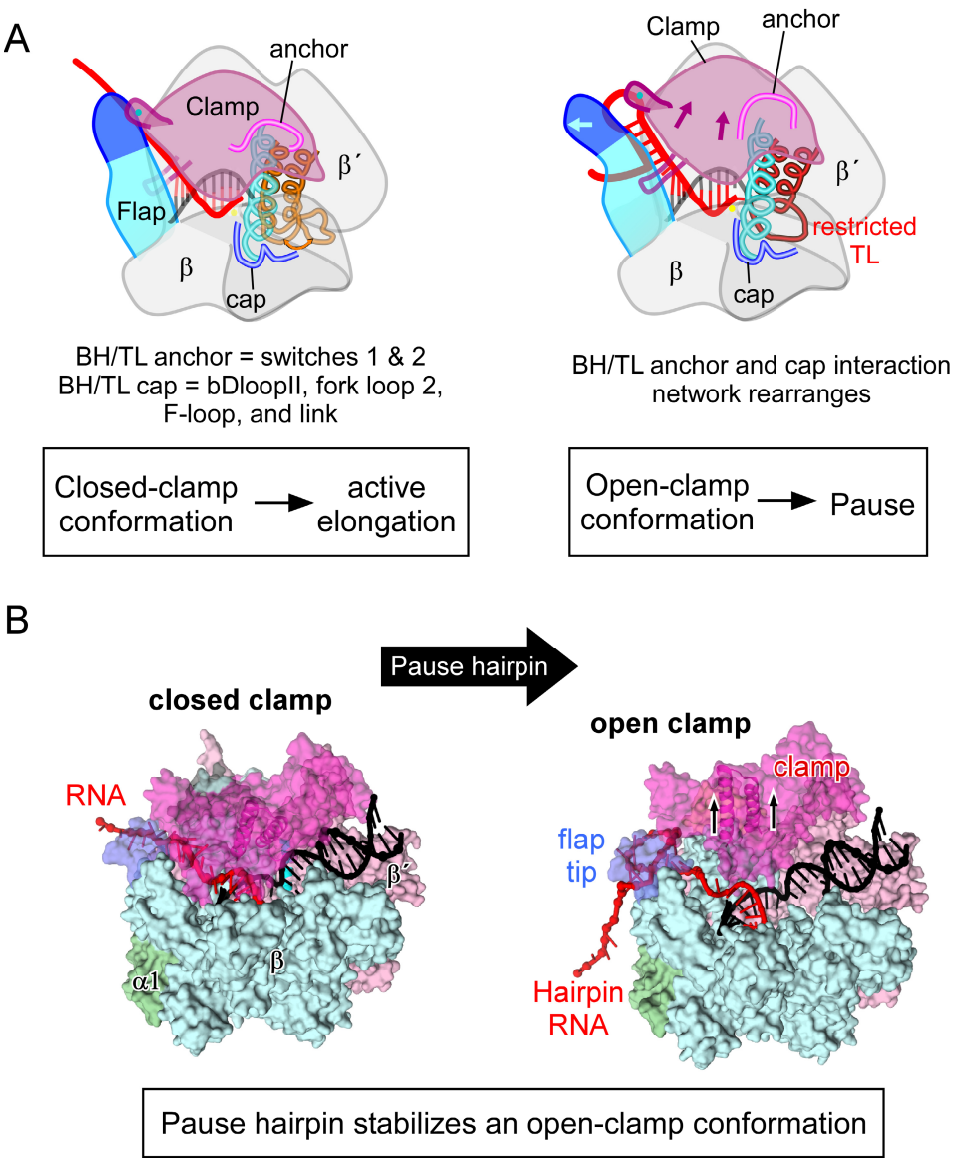


Figure 1-11. Transcriptional pausing mediates RNAP clamp movements

(A) Two models depicting closed (left) and open (right) clamp ECs. The clamp domain is highlighted in magenta. Interactions of the clamp, BH, TL, and BH/TL anchor and cap facilitate TL-TH oscillation and the nucleotide addition cycle.

Figure 1-11 legend. (cont.)

(B) Bacterial EC crystal structures representing RNAP clamp opening. The structural models of a closed-clamp elongation complex (EC) and an open-clamp EC are built based on Vassylyev et al., 2007 (PDB: 2o5j) (Vassylyev et al., 2007b) and (PDB: 4gzy and 3aoi) (Tagami et al., 2010; Weixlbaumer et al., 2013).

Coordinated movements of the trigger-loop, the bridge helix, and the clamp domain during transcription pausing

In multisubunit RNAPs, folding of the highly conserved TL into two anti-parallel α-helices (trigger helices, TH) packed against the bridge helix (BH) in a three-helix bundle is crucial for catalysis (Zhang, 2009). The TL also plays a key role in transcriptional pausing by orchestrating the formation of PECs when its mobility is dramatically restricted. RNA-protein crosslinking experiments provided biochemical evidence that the TL is trapped near the RNA 3′ nucleotide upon formation of the PEC (Toulokhonov et al., 2007). Furthermore, the strong effects on pausing of various substitutions of TL residues (Bar-Nahum et al., 2005; Kireeva et al., 2008; Toulokhonov et al., 2007), the observation of multiple TL conformations in X-ray crystal structures (Kettenberger et al., 2004; Temiakov et al., 2005; Tuske et al., 2005; Vassylyev et al., 2007b; Wang et al., 2009; Wang et al., 2006), and detection of changes in TL conformations using Cys-pair reporters (Nayak et al., 2013) are all consistent with the view that TL folding is inhibited in a PEC even after NTP-binding. Trigger loop conformation may be coupled to clamp movements through the bridge helix anchor and cap regions, so RNA hairpin effects on clamp position could affect TL folding and thus pausing (Figure 1-11). However, a direct demonstration of the regulatory communication between formation of the RNA hairpin and the active site of RNAP through clamp movement is lacking. The evidence supporting the energetic linkage between the RNA hairpin formation and the active site of RNAP via the clamp movement will be discussed in detail in chapters three and four.

Potential contributions of slow translocation to pausing

One possible explanation for the slow escape rate in the active elongation pathway is the inhibition of translocation at the pause site. Importantly, pause ECs with the hairpin RNA at the *his* pause site are pretranslocated so the RNA 3′ end flips out of the active site, contacting the abnormal site and hindering the enzyme's ability to translocate on the DNA (Toulokhonov et al., 2007). Although the recently described crystal structure of elemental pause elongation complex (ePEC) captures a post-translocated state with a kinked bridge helix, it was suggested that pause RNA hairpin formation could shift the complex to the pre-translocated state, which would promote pausing (Weixlbaumer et al., 2013). However, the relative contributions of slow translocation to transcriptional pausing are unknown and have never been measured directly. In chapter four, I will further investigate the effect of pause hairpin on forward translocation of RNAP at the hairpin-stabilized pause site.

The current view of the hairpin-stabilized pause is that upon encountering the pause signal, some fraction of RNAP enters the off-line elemental pause state due to weakened clamp-nucleic acid contacts during translocation. This loss of contacts causes active site rearrangements including the bridge helix kinking and opening of the clamp, which in turn favor the widened RNA exit channel. The enlarged exit channel provides the opportunity for the formation of the RNA hairpin. The pause RNA hairpin stabilizes the open clamp conformation which changes the anchor- BH/TL contacts, resulting in the inhibition of TL folding (Nayak et al., 2013). The formation of the pause hairpin also inhibits forward translocation of RNAP by stabilizing the PECs into the pre-translocated state. The general pausing mechanism by which changes in the conformational state of the clamp control TL folding and translocation may be conserved throughout all domains of life (Werner, 2012).

1-4-3. Mutations in RNAP that affect pausing

When the pause RNA hairpin forms, it makes direct contact with the flap domain of RNAP β subunit, far from the active site of RNAP and appears to alter the activity of the RNAP catalytic center allosterically by shifting the position of the clamp domain of the enzyme (Toulokhonov and Landick, 2003; Toulokhonov et al., 2007). Modeling suggests that a pause hairpin formed in the RNA exit channel will generate steric clash with the β flap and β′ clamp, leading to the movement of these two domains (Figure 1-12A). Two types of changes to RNAP prevent regulatory communication between hairpin formation and the active site of RNAP: (i) deletion of a portion of the exit channel called the flap tip (FT) and (ii) amino-acid substitutions in the βDloop II or in the bridge helix of the enzyme (Toulokhonov and Landick, 2003; Toulokhonov et al., 2007) (Figures 1-12B and 1-12C). A third type of change, deletion of or substitutions in a key active site element called the trigger loop are proposed to influence the communication between the active site, the RNA exit channel (Epshtein et al., 2007; Toulokhonov et al., 2007). Previous studies have established that mutant RNAPs such as RpoB5101 (TI563, PS560 substitutions in the β Dloop II near the active site), F773V RNAP (substitution in the N-terminal part of the bridge helix), ΔFTH RNAP (βΔ(900-909) deletion in the flap tip helix), and ΔFT RNAP (βΔ(890-914) deletion in the flap tip region) are defective in pausing at the *his* pause and in NusA response (Artsimovitch et al., 2003; Pan et al., 1999; Toulokhonov et al., 2001; Toulokhonov and Landick, 2003; Toulokhonov et al., 2007) (Figures 1-12B & C). These pause-suppressing substitutions could affect the clamp position by favoring the closed conformation even at the pause site, which may in turn inhibit hairpin formation if pause hairpin formation requires the clamp opening. I will investigate the effects of these

changes in RNAP (flap-tip deletions, βDloop II substitutions, and stabilization of either open or closed clamp domain) on the formation of nascent RNA structure in chapter four.

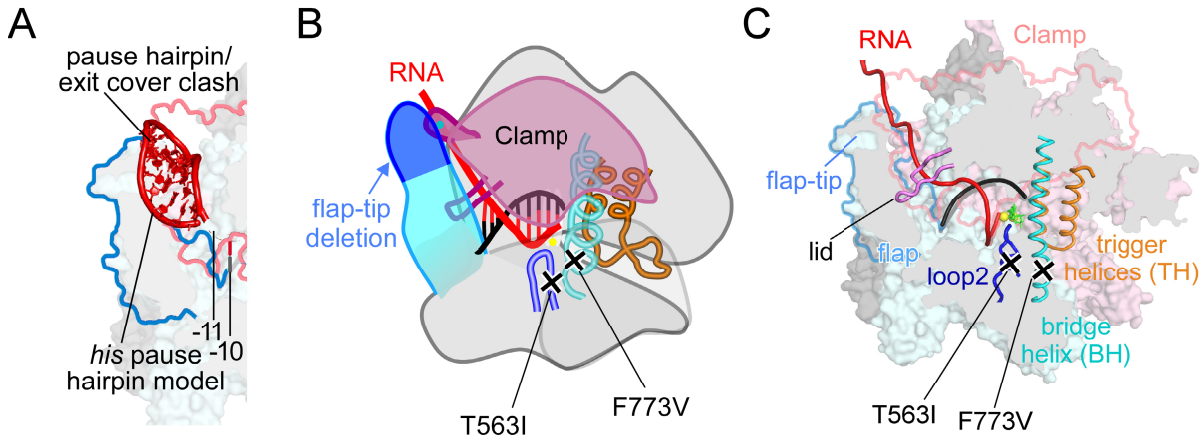


Figure 1-12. Steric clashes between a pause hairpin RNA and the exit channel cover in a closed-clamp EC and mutations in RNAP that suppress pausing at the *his* pause site

(A) *his* pause hairpin model showing the clash with clamp and flap tip.

(B) A model of an elongation complex (EC) with amino-acid substitutions in the βDloop II or in the bridge helix of the enzyme that lies directly adjacent to the active site are indicated. The blue arrow indicates the deletion of a portion of the exit channel called the β flap tip (FT; dark blue).

(C) Cutaway view of a structure of bacterial EC with flap (blue) and clamp (salmon) outlines showing amino-acid substitutions (T563I in the βDloop II and F773V in the N-terminal bridge helix) that decrease pausing at the *his* pause site.

1-4-4. Modulation of transcriptional pausing by regulatory factors

Pause elongation complexes formed during transcriptional pausing serve as a regulatory target for the control of RNA synthesis by elongation factors. As described above, NusA increases hairpin dependent pausing by interacting with both the nascent RNA secondary structure and the RNAP RNA exit channel (Kyzer et al., 2007; Toulokhonov et al., 2001). Gusarov and Nudler propose that NusA stimulates the pause RNA hairpin formation indirectly by binding to and displacing nascent RNA from a ssRNA-specific binding site near the exit channel (Gusarov and Nudler, 1999). Alternatively, a direct model of NusA interaction posits that N-terminal domain of NusA makes direct contact to the hairpin RNA and this interaction either inhibits translocation or stabilizes open-clamp position (Ha et al., 2010; Kolb et al., 2013).

Both eukaryotic and prokaryotic elongation factors can associate with RNAP's clamp and and suppress clamp opening and pausing (Werner, 2012). For example, *E. coli* elongation factor RfaH inhibits PEC formation at the hairpin-stabilized *his* pause site by stabilizing the closed-clamp conformation (Artsimovitch and Landick, 2002; Sevostyanova et al., 2011). As predicted, binding of pause-suppressing elongation factor RfaH to PECs promotes trigger helices formation and facilitates nucleotide addition. RfaH may inhibit pause hairpin formation by favoring the closed-clamp conformation, but this has not been tested, nor are there high-resolution crystal structures of PEC-RfaH.

1-5. Key research questions addressed by the thesis work

Understanding RNAP translocation and its control is crucial to the study of mechanisms underlying transcript elongation. However, it is unclear what controls the translocation register of RNAP, the position of the transcript 3′ end relative to the enzyme's catalytic site. Previous

studies concluded that RNA:DNA hybrid length controls the translocation register (Kashkina et al., 2006). However, these results might be explained by differences in the sequence of the 3′ end RNA. Therefore, I determined what components of the elongation complex (EC) control the translocation register of RNAP (chapter two).

At the hairpin-stabilized *his* pause, a pause RNA hairpin structure makes direct contact with the flap domain of RNAP β subunit and alters the activity of the RNAP catalytic center allosterically by shifting the position of the clamp domain of the enzyme (Toulokhonov et al., 2001; Toulokhonov et al., 2007). This process is further regulated by transcription factors that suppress or increase pausing (*e.g.*, bacterial RfaH, and NusA, respectively). NusA enhances pausing through direct interaction with both the nascent RNA secondary structure and the RNAP RNA exit channel (Ha et al., 2010), whereas pause-suppressing transcription factor RfaH is proposed to stabilize the closed-clamp conformation of RNAP through contacts to the β′clamp and β gate loop of RNAP (Sevostyanova et al., 2011). However, the mechanistic contributions of these factors to pausing remain poorly defined. Interestingly, the effect of nascent transcript secondary structures can be mimicked by annealing short RNA oligonucleotides (8-10 nucleotides long) to the nascent RNA of preformed ECs (Ha et al., 2010) (chapter three). In chapter three, I present work in which we determined the optimal length of the duplex is required for enhancement of oligo-mediated pausing, the differences between RNA:RNA and RNA:DNA duplexes, and determine how pausing is affected by regulators like NusA and RfaH.

The current model of the hairpin-stabilized transcription pause suggests that the pause hairpin stabilizes the open-clamp conformation. This change in clamp position causes the active site rearrangements, which result in inhibition of nucleotide addition (Toulokhonov et al., 2007; Weixlbaumer et al., 2013). I tested if the active site and the secondary structure formation are

energetically linked. Additionally, I asked if this regulatory communication between formation of hairpin structure and the active site of RNAP through the movements of the clamp can be disrupted by either the deletion the flap tip or by amino-acid substitutions in the βD loop II of the enzyme by testing if hairpin formation is different for wild-type and these mutant RNAPs described above (chapter four).

Recent crystal structure of the elemental pause elongation complex reveals the shifting of the clamp position by favoring the open state at the pause site (Toulokhonov et al., 2007; Weixlbaumer et al., 2013). Furthermore, modeling suggests that hairpin RNA clashes in the narrow RNA exit channel observed in ECs with the closed-clamp conformation. Therefore, the opening of the clamp widens the RNA exit channel, which is proposed to be important for formation of the pause hairpin RNA. Important questions raised by these studies are: (i) Is hairpin formation or pausing affected by restricting the clamp opening? (ii) Does RfaH inhibit hairpin formation by favoring the closed-clamp conformation? (iii) Can the pause hairpin effect be mimicked by stabilizing the open-clamp conformation in the absence of the hairpin? In other words, can we switch ECs into PECs by shifting the clamp positions without hairpin? These questions are addressed in chapter four.

One possible explanation for the slow escape rate from the *his* pause site is the inhibition of translocation at the pause site. Indeed, pausing favors the pretranslocated state. Notably, the RNA 3′ end flips out of the active site, hindering the enzyme's ability to translocate on the DNA while maintaining the pretranslocated state in PECs at the *his* pause site (Toulokhonov et al., 2007). These observations led to proposals that the hairpin RNA may influence RNAP translocation and thus pausing. In chapter four, I describe studies that address the hairpin effect on translocation at the *his* pause site by employing a fluorescent based translocation assay, while

annealing RNA oligos to the nascent RNA transcript of the TECs to mimic the hairpin stem. This approach allowed us to determine the precise translocation state of RNAP in real-time without perturbing the original translocation register.

Transcription factors in the cell, which increase or suppress pausing, also affect RNAP translocation. The transcription elongation factor, NusA, slows elongation rate, enhances hairpin dependent pausing, and is proposed to hinder forward translocation of RNAP (Toulokhonov et al., 2001; Zhou et al., 2011). It is unclear if pause stimulation by NusA is a result of slowing RNA translocation at the pause site. Another transcription elongation factor, RfaH, can reduce the pause-enhancing effects of the RNA hairpin perhaps by promoting forward translocation (Artsimovitch and Landick, 2002; Svetlov et al., 2007). I describe studies that test the modulation of translocation process by elongation factors, NusA and RfaH at the *his* pause site in chapter four.

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.

Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell *42*, 599-610.

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

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

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

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell *109*, 193-203.

Bai, L., Shundrovsky, A., and Wang, M.D. (2004). Sequence-dependent kinetic model for transcription elongation by RNA polymerase. Journal of molecular biology *344*, 335-349.

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.

Belogurov, G.A., Mooney, R.A., Svetlov, V., Landick, R., and Artsimovitch, I. (2009). Functional specialization of transcription elongation factors. EMBO J *28*, 112-122.

Belogurov, G.A., Vassylyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassylyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. Mol Cell *26*, 117-129.

Beuth, B., Pennell, S., Arnvig, K.B., Martin, S.R., and Taylor, I.A. (2005). Structure of a Mycobacterium tuberculosis NusA-RNA complex. EMBO J *24*, 3576-3587.

Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992). GreA protein: a transcription elongation factor from Escherichia coli. Proc Natl Acad Sci U S A *89*, 8899-8902.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from E. coli. Cell *72*, 459-466.

Browning, D.F., and Busby, S.J. (2004). The regulation of bacterial transcription initiation. Nat Rev Microbiol *2*, 57-65.

Brueckner, F., and Cramer, P. (2008). Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. Nat Struct Mol Biol *15*, 811-818.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. (1969). Factor stimulating transcription by RNA polymerase. Nature *221*, 43-46.

Burmann, B.M., Knauer, S.H., Sevostyanova, A., Schweimer, K., Mooney, R.A., Landick, R., Artsimovitch, I., and Rosch, P. (2012). An alpha helix to beta barrel domain switch transforms the transcription factor RfaH into a translation factor. Cell *150*, 291-303.

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.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001). Structural mechanism for rifampicin inhibition of bacterial rna polymerase. Cell *104*, 901-912.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R.*, et al.* (2012). Opening and closing of the bacterial RNA polymerase clamp. Science *337*, 591-595.

Chan, C.L., and Landick, R. (1993). Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. Journal of molecular biology *233*, 25-42.

Chan, C.L., Wang, D., and Landick, R. (1997). Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. Journal of molecular biology *268*, 54-68.

Cheung, A.C., and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature *471*, 249-253.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J.*, et al.* (2008). Structure of eukaryotic RNA polymerases. Annu Rev Biophys *37*, 337-352.

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

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.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell *12*, 525-532.

Ebright, R.H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. Journal of molecular biology *304*, 687-698.

Ederth, J., Mooney, R.A., Isaksson, L.A., and Landick, R. (2006). Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. Journal of molecular biology *356*, 1163-1179.

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). An allosteric path to transcription termination. Mol Cell *28*, 991-1001.

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

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

Geszvain, K., and Landick, R. (2005). The structure of bacterial RNA polymerase. In The bacterial chromosome, N.P. Higgins, ed. (Washington, D. C.: ASM Press), pp. 283-296.

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.

Gopal, B., Haire, L.F., Gamblin, S.J., Dodson, E.J., Lane, A.N., Papavinasasundaram, K.G., Colston, M.J., and Dodson, G. (2001). Crystal structure of the transcription elongation/anti-termination factor NusA from Mycobacterium tuberculosis at 1.7 A resolution. Journal of molecular biology *314*, 1087-1095.

Guajardo, R., and Sousa, R. (1997). A model for the mechanism of polymerase translocation. Journal of molecular biology *265*, 8-19.

Guo, Q., and Sousa, R. (2006). Translocation by T7 RNA polymerase: a sensitively poised Brownian ratchet. Journal of molecular biology *358*, 241-254.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Mol Cell *3*, 495-504.

Ha, K.S., Toulokhonov, I., Vassylyev, D.G., and Landick, R. (2010). The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. Journal of molecular biology *401*, 708-725.

Haugen, S.P., Ross, W., and Gourse, R.L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol *6*, 507-519.

Hein, P.P., Palangat, M., and Landick, R. (2011). RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. Biochemistry *50*, 7002-7014.

Herbert, K.M., Zhou, J., Mooney, R.A., Porta, A.L., Landick, R., and Block, S.M. (2010). E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. Journal of molecular biology *399*, 17-30.

Hirata, A., Klein, B.J., and Murakami, K.S. (2008). The X-ray crystal structure of RNA polymerase from Archaea. Nature *451*, 851-854.

Hirtreiter, A., Damsma, G.E., Cheung, A.C., Klose, D., Grohmann, D., Vojnic, E., Martin, A.C., Cramer, P., and Werner, F. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. Nucleic Acids Res *38*, 4040-4051.

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.

Izban, M.G., and Luse, D.S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. Genes Dev *6*, 1342-1356.

Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., and Ebright, R.H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science *314*, 1144-1147.

Kashkina, E., Anikin, M., Tahirov, T.H., Kochetkov, S.N., Vassylyev, D.G., and Temiakov, D. (2006). Elongation complexes of Thermus thermophilus RNA polymerase that possess distinct translocation conformations. Nucleic Acids Res *34*, 4036-4045.

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., Kashlev, M., and Burton, Z.F. (2010). Translocation by multi-subunit RNA polymerases. Biochim Biophys Acta.

Kireeva, M.L., and Kashlev, M. (2009). Mechanism of sequence-specific pausing of bacterial RNA polymerase. Proc Natl Acad Sci U S A *106*, 8900-8905.

Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purtov, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N., and Kashlev, M. (2008). Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol Cell *30*, 557-566.

Kolb, K., Hein, P., and Landick, R. (2013). Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit-channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. J Biol Chem *in press*.

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

Krummel, B., and Chamberlin, M.J. (1992). Structural analysis of ternary complexes of Escherichia coli RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. Journal of molecular biology *225*, 239-250.

Kyzer, S., Ha, K.S., Landick, R., and Palangat, M. (2007). Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. J Biol Chem *282*, 19020-19028.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.

Landick, R. (2009). Transcriptional pausing without backtracking. Proc Natl Acad Sci U S A *106*, 8797-8798.

Landick, R., Carey, J., and Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the trp operon leader region. Proc Natl Acad Sci U S A *82*, 4663-4667.

Landick, R., Stewart, J., and Lee, D.N. (1990). Amino acid changes in conserved regions of the beta-subunit of Escherichia coli RNA polymerase alter transcription pausing and termination. Genes Dev *4*, 1623-1636.

Laptenko, O., Lee, J., Lomakin, I., and Borukhov, S. (2003). Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. EMBO J *22*, 6322-6334.

Larson, M.H., Zhou, J., Kaplan, C.D., Palangat, M., Kornberg, R.D., Landick, R., and Block, S.M. (2012). Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. Proc Natl Acad Sci U S A *109*, 6555-6560.

Lee, D.N., Phung, L., Stewart, J., and Landick, R. (1990). Transcription pausing by Escherichia coli RNA polymerase is modulated by downstream DNA sequences. J Biol Chem *265*, 15145-15153.

Lowery-Goldhammer, C., and Richardson, J.P. (1974). An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor. Proc Natl Acad Sci U S A *71*, 2003-2007.

Malinen, A.M., Turtola, M., Parthiban, M., Vainonen, L., Johnson, M.S., and Belogurov, G.A. (2012). Active site opening and closure control translocation of multisubunit RNA polymerase. Nucleic Acids Res *40*, 7442-7451.

Marchand, B., and Gotte, M. (2003). Site-specific footprinting reveals differences in the translocation status of HIV-1 reverse transcriptase. Implications for polymerase translocation and drug resistance. J Biol Chem *278*, 35362-35372.

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H., and Severinov, K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci U S A *98*, 892-897.

Miropolskaya, N., Artsimovitch, I., Klimasauskas, S., Nikiforov, V., and Kulbachinskiy, A. (2009). Allosteric control of catalysis by the F loop of RNA polymerase. Proc Natl Acad Sci U S A *106*, 18942-18947.

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

Mooney, R.A., Schweimer, K., Rosch, P., Gottesman, M., and Landick, R. (2009). Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. Journal of molecular biology *391*, 341-358.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. Mol Cell *50*, 882-893.

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.

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

Ninio, J. (1991). Connections between translation, transcription and replication error-rates. Biochimie *73*, 1517-1523.

Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: protein-DNA interactions holding together the elongation complex. Science *273*, 211-217.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., and Darst, S.A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. Cell *114*, 335-345.

Palangat, M., Meier, T.I., Keene, R.G., and Landick, R. (1998). Transcriptional pausing at +62 of the HIV-1 nascent RNA modulates formation of the TAR RNA structure. Mol Cell *1*, 1033-1042.

Pan, T., Artsimovitch, I., Fang, X.W., Landick, R., and Sosnick, T.R. (1999). Folding of a large ribozyme during transcription and the effect of the elongation factor NusA. Proc Natl Acad Sci U S A *96*, 9545-9550.

Pan, T., and Sosnick, T. (2006). RNA folding during transcription. Annu Rev Biophys Biomol Struct *35*, 161-175.

Park, N.J., Tsao, D.C., and Martinson, H.G. (2004). The two steps of poly(A)-dependent termination, pausing and release, can be uncoupled by truncation of the RNA polymerase II carboxyl-terminal repeat domain. Mol Cell Biol *24*, 4092-4103.

Peters, J.M., Vangeloff, A.D., and Landick, R. (2011). Bacterial transcription terminators: the RNA 3'-end chronicles. Journal of molecular biology *412*, 793-813.

Rees, W.A., Weitzel, S.E., Das, A., and von Hippel, P.H. (1997). Regulation of the elongation-termination decision at intrinsic terminators by antitermination protein N of phage lambda. Journal of molecular biology *273*, 797-813.

Rees, W.A., Weitzel, S.E., Yager, T.D., Das, A., and von Hippel, P.H. (1996). Bacteriophage lambda N protein alone can induce transcription antitermination in vitro. Proc Natl Acad Sci U S A *93*, 342-346.

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

Reynolds, R., and Chamberlin, M.J. (1992). Parameters affecting transcription termination by Escherichia coli RNA. II. Construction and analysis of hybrid terminators. Journal of molecular biology *224*, 53-63.

Richardson, J.P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim Biophys Acta *1577*, 251-260.

Roberts, J.W., Shankar, S., and Filter, J.J. (2008). RNA polymerase elongation factors. Annu Rev Microbiol *62*, 211-233.

Roberts, J.W., Yarnell, W., Bartlett, E., Guo, J., Marr, M., Ko, D.C., Sun, H., and Roberts, C.W. (1998). Antitermination by bacteriophage lambda Q protein. Cold Spring Harb Symp Quant Biol *63*, 319-325.

Robson-Dixon, N.D., and Garcia-Blanco, M.A. (2004). MAZ elements alter transcription elongation and silencing of the fibroblast growth factor receptor 2 exon IIIb. J Biol Chem *279*, 29075-29084.

Saecker, R.M., Tsodikov, O.V., McQuade, K.L., Schlax, P.E., Jr., Capp, M.W., and Record, M.T., Jr. (2002). Kinetic studies and structural models of the association of E. coli sigma(70) RNA polymerase with the lambdaP(R) promoter: large scale conformational changes in forming the kinetically significant intermediates. Journal of molecular biology *319*, 649-671.

Seibold, S.A., Singh, B.N., Zhang, C., Kireeva, M., Domecq, C., Bouchard, A., Nazione, A.M., Feig, M., Cukier, R.I., Coulombe, B.*, et al.* (2010). Conformational coupling, bridge helix dynamics and active site dehydration in catalysis by RNA polymerase. Biochim Biophys Acta *1799*, 575-587.

Sevostyanova, A., Belogurov, G.A., Mooney, R.A., Landick, R., and Artsimovitch, I. (2011). The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Mol Cell *43*, 253-262.

Skordalakes, E., and Berger, J.M. (2006). Structural insights into RNA-dependent ring closure and ATPase activation by the Rho termination factor. Cell *127*, 553-564.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J *22*, 2234-2244.

Sosunov, V., Zorov, S., Sosunova, E., Nikolaev, A., Zakeyeva, I., Bass, I., Goldfarb, A., Nikiforov, V., Severinov, K., and Mustaev, A. (2005). The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. Nucleic Acids Res *33*, 4202-4211.

Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A., and Mustaev, A. (2003). Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. Proc Natl Acad Sci U S A *100*, 15469-15474.

Steitz, T.A. (1998). A mechanism for all polymerases. Nature *391*, 231-232.

Svetlov, V., Belogurov, G.A., Shabrova, E., Vassylyev, D.G., and Artsimovitch, I. (2007). Allosteric control of the RNA polymerase by the elongation factor RfaH. Nucleic Acids Res *35*, 5694-5705.

Sweetser, D., Nonet, M., and Young, R.A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc Natl Acad Sci U S A *84*, 1192-1196.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., and Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. Nature *468*, 978-982.

Telesnitsky, A., and Chamberlin, M.J. (1989). Terminator-distal sequences determine the in vitro efficiency of the early terminators of bacteriophages T3 and T7. Biochemistry *28*, 5210-5218.

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.

Thomen, P., Lopez, P.J., and Heslot, F. (2005). Unravelling the mechanism of RNA-polymerase forward motion by using mechanical force. Phys Rev Lett *94*, 128102.

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

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

Toulokhonov, I., and Landick, R. (2006). The role of the lid element in transcription by E. coli RNA polymerase. Journal of molecular biology *361*, 644-658.

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.

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

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. (2007a). 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. (2007b). Structural basis for substrate loading in bacterial RNA polymerase. Nature *448*, 163-168.

Vogel, U., and Jensen, K.F. (1994). The RNA chain elongation rate in Escherichia coli depends on the growth rate. J Bacteriol *176*, 2807-2813.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., and Kornberg, R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science *324*, 1203-1206.

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

Weinzierl, R.O. (2010). The nucleotide addition cycle of RNA polymerase is controlled by two molecular hinges in the Bridge Helix domain. BMC Biol *8*, 134.

Weixlbaumer, A., Leon, K., Landick, R., and Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. Cell *152*, 431-441.

Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. Journal of molecular biology *417*, 13-27.

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol *9*, 85-98.

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.

Whalen, W., Ghosh, B., and Das, A. (1988). NusA protein is necessary and sufficient in vitro for phage lambda N gene product to suppress a rho-independent terminator placed downstream of nutL. Proc Natl Acad Sci U S A *85*, 2494-2498.

Yang, X., Molimau, S., Doherty, G.P., Johnston, E.B., Marles-Wright, J., Rothnagel, R., Hankamer, B., Lewis, R.J., and Lewis, P.J. (2009). The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. EMBO Rep *10*, 997-1002.

Yarnell, W.S., and Roberts, J.W. (1992). The phage lambda gene Q transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. Cell *69*, 1181-1189.

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

Yonaha, M., and Proudfoot, N.J. (1999). Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. Mol Cell *3*, 593-600.

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, 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.

Zhang, J., Palangat, M., and Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. Nat Struct Mol Biol *17*, 99-104.

Zhang, J.a.L., R. (2009). Substrate Loading, Nucleotide Addition, and Translocation by RNA Polymerase. In RNA Polymerase as Molecular Motors, H.a.S. Buc, T., ed. (Cambridge, UK: Royal Society of Chemistry), pp. 206-235.

Zhou, J., Ha, K.S., La Porta, A., Landick, R., and Block, S.M. (2011). Applied force provides insight into transcriptional pausing and its modulation by transcription factor NusA. Mol Cell *44*, 635-646.

Chapter Two

RNA Transcript 3′-Proximal Sequence Affects Translocation Bias of RNA Polymerase

This chapter is adapted with permission from:

"RNA Transcript 3′-Proximal Sequence Affects Translocation Bias of RNA Polymerase" Pyae Hein, Murali Palangat, and Robert Landick. 2011. Biochemistry. 50 (32): 7002-7014.

Abstract

Translocation of RNA polymerase on DNA is thought to involve oscillations between pretranslocated and posttranslocated states that are rectified by nucleotide addition or pyrophosphorolysis. The pretranslocated register is also a precursor to transcriptional pause states that mediate regulation of transcript elongation. However, the determinants of bias between the pretranslocated and posttranslocated states are incompletely understood. To investigate translocation bias in multisubunit RNA polymerases, we measured rates of pyrophosphorolysis, which occurs in the pretranslocated register, in minimal elongation complexes containing *T. thermophilus* or *E. coli* RNA polymerase. Our results suggest that the identity of RNA:DNA nucleotides in the active site are strong determinants of susceptibility to pyrophosphorolysis, and thus translocation bias, with the 3′ RNA nucleotide favoring the pretranslocated state in the order U > C > A > G. The preference of 3′ U vs G for the pretranslocated register appeared to be universal among both bacterial and eukaryotic RNA polymerases and was confirmed by exonuclease III footprinting of defined elongation complexes. However, the relationship of pyrophosphate concentration to the rate of pyrophosphorolysis of 3′ U- versus 3′ G-containing elongation complexes did not match predictions of a simple mechanism in which 3′-RNA seqeunce affects only translocation bias and pyrophosphate (PPi) binds only to the pretranslocated state.

Introduction

Cellular gene expression in all free-living organisms relies on evolutionarily conserved, highly regulated, multi-subunit RNA polymerases (RNAPs). During transcription, synthesis of the RNA transcript requires stepping the DNA template strand through the RNAP active site one nucleotide (nt) at a time in a process called translocation (recently reviewed in refs Kireeva et al., 2010; Zhang, 2009). Translocation is one of four steps in a nucleotide addition cycle (NAC; Figure 2-1) that also includes nucleoside triphosphate (NTP) binding, catalysis, and pyrophosphate (PPi) release. At the beginning of the NAC, the RNA 3′ nt is in the *i+1* subsite of RNAP, corresponding to the pretranslocated register of the elongating complex (EC). Translocation of DNA through RNAP generates the posttranslocated register, positions the RNA 3′ nt into the product (P) or *i* subsite, and opens the *i+1* (A) subsite to bind the incoming cognate NTP complexed with Mg^{2+}II. Catalysis occurs by alignment of the phosphate of the bound NTP with the RNA 3′ OH upon formation of a three-helix bundle in the active site consisting of the trigger helices (TH; which form from the trigger loop; TL) and the bridge helix (BH). TH-NTP contacts facilitate an S_N2 nucleophilic reaction in which a trigonal bipyramidyl transition state is stabilized by two catalytic Mg^{2+} ions (Sosunov et al., 2003). This reaction extends the RNA transcript by one nt and generates PPi, the release of which leaves the EC in the pretranslocated register ready for the next round of nucleotide addition. The NAC is reversible at high PPi concentration. Although the basic features of the NAC are established, the order in which translocation occurs relative to NTP binding and PPi release, which step is rate-limiting, and the path of NTP loading remains uncertain and under study (Kireeva et al., 2010; Zhang, 2009).

Translocation is thought to bear features of a thermal ratchet mechanism in which RNAP can oscillate between the pre and posttranslocated registers, with the bias between registers

determined by their relative stabilities and the net motion of DNA through RNAP resulting from nucleotide addition or pyrophosphorolysis (Kireeva et al., 2010). However, the elemental rates of forward and reverse translocation, the effects of scaffold structure (*e.g.*, RNA:DNA hybrid length) on translocation bias and the sequence-specific effects of interactions between nucleic acid bases and RNAP side chains on translocation bias are unknown (Kireeva et al., 2010). In the simplest form of this mechanism (Figure 2-1), translocation bias should affect the apparent binding affinities of NTP and PPi by dictating the fraction of time their binding sites are available. It also should determine susceptibility of the EC to transcriptional pausing, which plays many roles in the regulation of transcript elongation by multisubunit RNAPs. Although the mechanism of pausing remains under study, current proposals all identify the pretranslocated EC as the starting point for either (*i*) a structural isomerization in the EC that can be followed by reverse translocation (backtracking) or by other events that stabilize the paused EC (Landick, 2006, 2009), (*ii*) direct conversion to backtracked (paused) states (Depken et al., 2009; Galburt et al., 2007; Mejia et al., 2008), or (*iii*) inhibition of conversion to the posttranslocated state by thermodynamic stability of the pretranslocated state (Bai et al., 2004). The energetics of translocation bias during backtracking are better understood and are governed by the relative stabilities of the RNA:DNA hybrid and the transcription bubble in different translocation registers (Galburt et al., 2007; Komissarova and Kashlev, 1997; Nudler et al., 1997). Translocation bias may also affect susceptibility to termination, which is thought to occur *via* either hypertranslocation without RNA synthesis or hybrid shearing (Larson et al., 2008; Santangelo and Roberts, 2004), both of which should be easier from the posttranslocated state (although alternative, non-translocational termination models also have been proposed; Refs. (Epshtein et al., 2007; Epshtein et al., 2010).

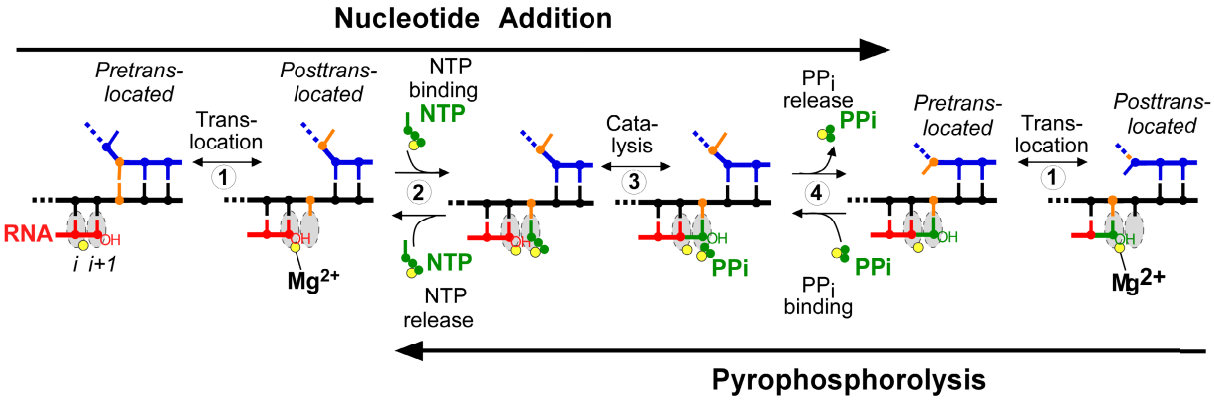


Figure 2-1. Nucleotide addition and pyrophosphorolysis cycle.

Template DNA is shown in black, non-template DNA in blue, RNA in red, incoming nucleotide triphosphate and pyrophosphate in green, and Mg^{2+} in yellow. One template/nontemplate position is colored orange to illustrate translocation. The EC alternates between pre and posttranslocated states with the RNA 3′ nt in the *i* or *i+1* subsites, prior to NTP binding (step 1). NTPs enter the active site when the EC is posttranslocated (step 2). Catalysis (step 3) requires a rate-limiting conformational change in which the trigger loop folds into the trigger helices (not shown). Release of PPi (step 4) completes the nucleotide addition cycle. Pyrophosphorolysis is the reverse reaction of nucleotide addition. Since catalysis itself is reversible, net pyrophosphorolysis requires that the thermodynamic driving force of PPi conversion to $NTP+RNA^{-1}$ exceed the conversion of NTP to $PPi+RNA^{+1}$. Pyrophosphorolysis requires the EC be in the pretranslocated register, but whether PPi binding/release and translocation occur with obligate order is not established.

Thus, understanding the nature of translocation bias is important not only to gain insight in the NAC but also to understand regulatory events like pausing and termination and the mechanisms of regulators like NusA and NusG that are proposed to alter translocational bias (Bar-Nahum et al., 2005), (Pasman and von Hippel, 2000).

We have investigated the contribution of hybrid length and scaffold sequence to translocation bias using sensitivity to pyrophosphorolysis, which can occur only from the pretranslocated register (Rozovskaya et al., 1982). Initially, we eliminated contributions of the transcription bubble, upstream DNA, and upstream RNA using minimal scaffolds containing only the hybrid and downstream DNA (Figure 2-2A). We then extended the experiments to use full scaffolds, test contributions to translocation bias by exonuclease III footprinting, and remove the effect of the reverse reaction (nucleotide addition) by using apyrase (Apy) to destroy NTPs generated by pyrophosphorolysis. Our findings suggest that the 3´ RNA dinucleotide sequence is a primary determinant of translocation bias, and that the PPi concentration-dependence of the pyrophosphorolysis reaction is inconsistent with the simple thermal ratchet model of the reaction.

Results

RNA 3´ sequence, not an 8-bp *vs.* 9-bp hybrid, primarily controls translocation bias

We first investigated the relationship between 8- *vs.* 9 bp hybrids, transcript sequence, and translocation register using minimal scaffolds that eliminate effects of transcription bubble energetics on translocation bias, and sequences that led Kashkina et al. (Kashkina et al., 2006) to suggest that the pre or posttranslocated registers were favored by 8-bp or 9-bp hybrids, respectively. We reconstituted ECs using *Thermus thermophilus* RNAP (*Tth*RNAP) on scaffolds

that differed either by hybrid length or by 3′ trinucleotide sequence, and then assessed translocational bias by measuring the rates of pyrophosphorolysis for each scaffold (Figure 2-2). We observed the same difference reported by Kashkina et al. (Kashkina et al., 2006) between the sensitive 8-bp hybrid scaffold $EC8^{GAU}$ (superscript indicates the 3′-proximal RNA sequence in the EC) and the resistant 9-bp hybrid scaffold $EC9^{AUA}$ (Figure 2-2B). However, when we examined pyrophosphorolysis of ECs in which the 3′ RNA sequences of 8-bp and 9-bp hybrid were swapped, we found that $EC8^{AUA}$ was resistant to pyrophosphorolysis whereas $EC9^{GAU}$ was sensitive to pyrophsophorolysis (Figure 2-2B). In other words, both $EC8^{GAU}$ and $EC9^{GAU}$ appear biased toward the pretranslocated state whereas both $EC8^{AUA}$ and $EC9^{AUA}$ appear biased toward the posttranslocated state. This result strongly suggests that the RNA 3′ sequence in the sensitive 3′-proximal GAU and resistant 3′-proximal AUA is the dominant determinant of translocation bias relative to any effects of 8- *vs.* 9-bp hybrids.

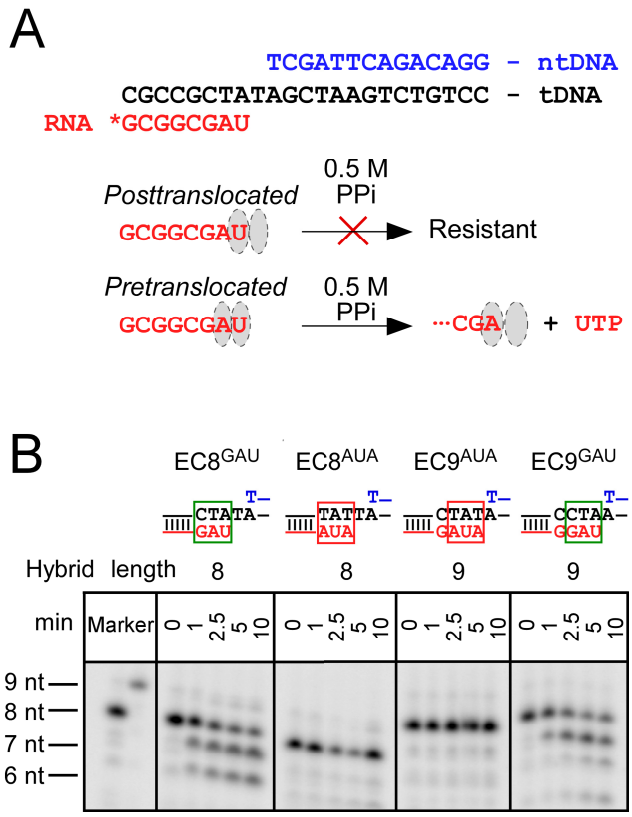


Figure 2-2. Pyrophosphorolysis by *Tth*RNAP on a minimal scaffold.

(A) A representative example of a minimal nucleic acid scaffold used in this study (yields $EC8^{AU}$). *, 5′ ^{32}P. The color scheme for RNA and DNA is the same as in Figure 2-1A. The posttranslocated state is resistant to pyrophosphorolysis; the pretranslocated state is competent for pyrophosphorolysis, yielding UTP in this case.

(B) ECs containing *Tth*RNAP were assembled on four minimal nucleic acid scaffolds that differ only in the sequence shown above the gel panels. The time course of pyrophosphorolysis (0.5 mM PPi) at 60° C is shown, with the sizes of RNAs indicated.

RNA 3′-dinucleotide dictates translocation bias

Given the dominant effect of the 3′-proximal sequence on pyrophosphorolysis, we next sought to determine which RNA 3′-dinucleotides are most and least sensitive to pyrophosphorolysis (presumably corresponding to the most and least biased toward the pretranslocated register). To accomplish this test, we generated 16 different ECs in which the 3′ dinucleotide sequence was varied to create all possible combinations, again using minimal scaffolds to avoid effects of transcription bubble energetics. When we tested pyrophosphorolysis in the 16 different ECs, we noticed that they exhibited differences not only in the rate of pyrophosphorolysis but also in the fraction of ECs in which the RNA that was shortened (*e.g.*, compare $EC9^{CG}$ and $EC9^{CU}$; Figures 2-3A & B). Incomplete RNA shortening was not due to incomplete reconstitution of ECs because addition of cognate NTP resulted in extension of all 9mer RNA (+NTP lane; Figures 2-3A & B). Rather, we suspected (and later confirmed; see below) that incomplete pyrophosphorolysis reflected equilibration of pyrophorphorolysis with the reverse reaction of nucleotide addition as the concentration of NTP produced by pyrophorphorolysis increased. Therefore, we fit the reaction progress curves to a simple reversible mechanism (reaction 1: see Materials and Methods).

This yielded a value for k_1 that approximated the initial rate of pyrophosphorolysis and allowed us to rank the pyrophosphorolysis sensitivities of the 16 different ECs by assigning both an observed rate of pyrophosphorolysis and an equilibrium fraction ($[EC8]_{eq}/[EC9]_{eq}$) of pyrophosphorolysed transcript (Table 2-1 and Figure 2-3C). For any given penultimate RNA nucleotide, both the rate of pyrophosphorolysis and the equilibrium fraction yielded the same order of sensitivity to pyrophosphorolysis for the 3′-terminal nucleotide: U > C > A > G. Interestingly, 3′-penultimate nucleotide gave the inverse order of effects on sensitivity to

pyrophosphorolysis: G > A > C > U. In general, faster pyrophosphorolysis rate correlated with a larger sensitive EC fraction (Figure 2-3C). These data are consistent with preferential interaction of G > A > C > U in the product (*i*) subsite and U > C > A > G in the NTP-binding (*i+1*) subunit (see Discussion).

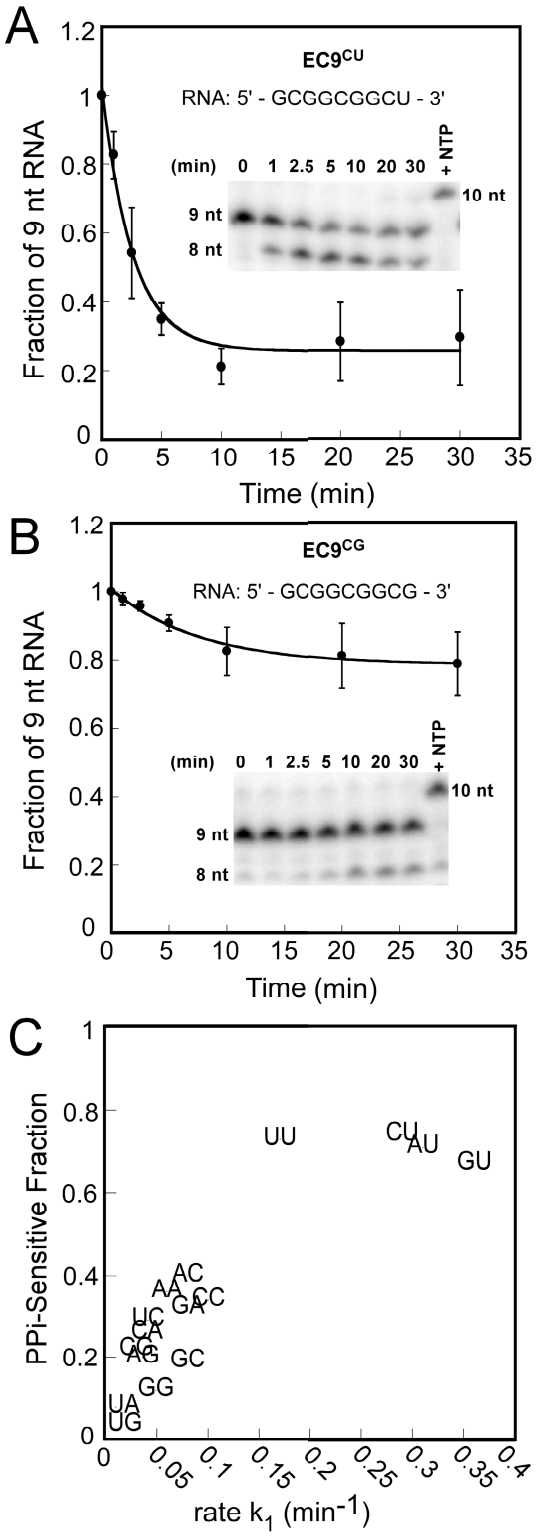


Figure 2-3. Identity of RNA 3´ terminal nucleotide affects translocation register of *Tth*RNAP.

Figure 2-3 legend.

(A) Quantitative analysis of pyrophosphorolysis of $EC9^{CU}$, which favors the pretranslocated state. *Tth*$EC9^{CU}$ was reconstituted from tDNA #6059, ntDNA #5848, and RNA #6042 (Table 2-3). ECs (50 nM) were incubated with 0.5 mM pyrophosphate at 60°C and disappearance of the 9 nt RNA measured at the times indicated in the inset. +NTP, UTP (1mM) was added at the end of the time course to confirm that the unreacted $EC9^{CU}$ remained active and could extend the RNA 1 nt by UMP incorporation. Errors are SD from three independent experiments. Data were fit to a simple reversible mechanism of pyrophosphorolysis (equation 3; see *Materials and Methods*). A value for k_1 that approximated the initial rate of pyrophosphorolysis (0.29 ± 0.04 min^{-1}) and a final fraction of pyrophosphorolysed transcript (PPi-sensitive fraction, remaining 9-nt RNA/total RNA = 0.74 ± 0.03) were determined by nonlinear regression.

(B) Quantitative analysis of pyrophosphorolysis of $EC9^{CG}$, which favors the posttranslcated state. The experiment was performed identically to that shown in panel A except that *Tth*$EC9^{CG}$ was reconstituted from tDNA #6062, ntDNA #5848, and RNA #6045 (Table 2-3).

(C) A plot of fraction of EC susceptible to pyrophosphorolysis (PPi-sensitive fraction) with 0.5 mM PPi versus the rate of pyrophosphorolysis for 16 combinations of 3′-proximal dinucleotide sequence.

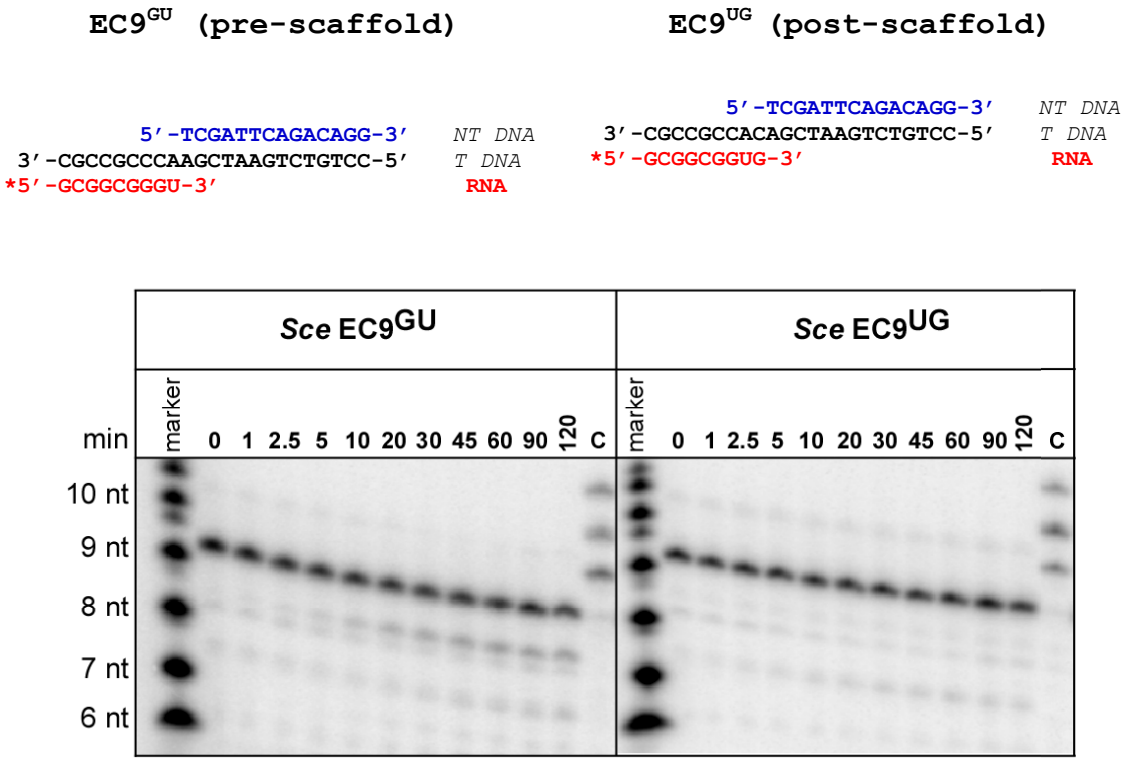
TABLE 2-1. Effect of 3′ diribonucleotide sequence on pyrophosphorolysis of 50 nM EC at 0.5 mM PPi.

EC RNA Sequence	Resistant Fraction	k_1	k_{-1}	K_{eq} (EC8/EC9)
5'-GCGGCGG(xx)		min^{-1}	min^{-1}	
AA	0.64 ± 0.03	0.06 ± 0.01	0.10 ± 0.03	0.6 ± 0.2
AC	0.60 ± 0.02	0.08 ± 0.01	0.11 ± 0.02	0.7 ± 0.2
AG	0.80 ± 0.01	0.04 ± 0.01	0.13 ± 0.03	0.3 ± 0.1
AU	0.29 ± 0.03	0.31 ± 0.05	0.12 ± 0.03	2.6 ± 0.8
CU	0.26 ± 0.03	0.29 ± 0.04	0.10 ± 0.02	2.9 ± 0.7
CA	0.74 ± 0.02	0.04 ± 0.01	0.11 ± 0.03	0.4 ± 0.1
CC	0.66 ± 0.02	0.10 ± 0.02	0.19 ± 0.03	0.5 ± 0.1
CG	0.78 ± 0.01	0.03 ± 0.01	0.10 ± 0.02	0.3 ± 0.1
GU	0.33 ± 0.03	0.36 ± 0.09	0.17 ± 0.05	2.1 ± 0.8
GA	0.68 ± 0.01	0.08 ± 0.01	0.16 ± 0.02	0.5 ± 0.1
GG	0.78 ± 0.01	0.05 ± 0.01	0.16 ± 0.02	0.3 ± 0.1
GC	0.81 ± 0.01	0.08 ± 0.02	0.36 ± 0.07	0.2 ± 0.1
UU	0.27 ± 0.03	0.17 ± 0.02	0.06 ± 0.01	2.8 ± 0.6
UA	0.92 ± 0.01	< 0.01	~ 0.14	≤ 0.09
UC	0.71 ± 0.01	0.04 ± 0.01	0.10 ± 0.02	0.4± 0.1
UG	0.96 ± 0.01	< 0.007	~ 0.14	≤ 0.05

RNA 3´ dinucleotide effects on pyrophosphorolysis are evolutionarily conserved

To investigate whether similar effects of RNA 3´ dinucleotides would be observed in ECs reconstituted with other multisubunit RNAPs, we compared ECs reconstituted on the most and least sensitive scaffolds ($EC9^{GU}$ and $EC9^{UG}$, respectively) with *Thermus thermophilus* RNAP (*Tth*RNAP), *Escherichia coli* RNAP (*Eco*RNAP), *Saccharomyces cerevisiae* RNAPII (*Sce*RNAPII), and calf thymus (*Bos taurus*) RNAPII (*Bta*RNAPII). For all four sources of RNAP, $EC9^{GU}$ was sensitive to pyrophosphorolysis and $EC9^{UG}$ was resistant (Table 2-2; Figure 2-4). We also confirmed that incomplete pyrophosphorolysis was not due to incomplete reconstitution of ECs by the addition of cognate NTP, which resulted in extension of all 9mer RNA (+NTP lane; Figures 2-3A & 2-3B; C lane; Figure 2-4A & 2-4B) Thus, the same RNA 3´ dinucleotide characteristics appear to control translocation bias for evolutionarily diverse multisubunit RNAPs.

A



B

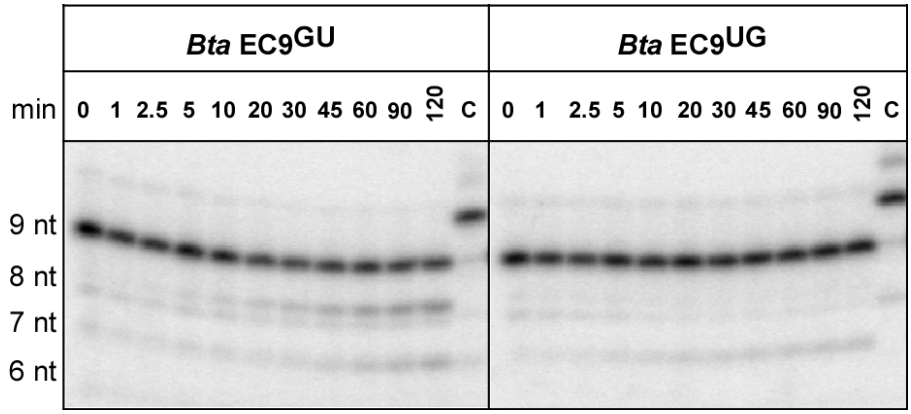


Figure 2-4. RNA 3′ dinucleotide effects on pyrophosphorolysis are evolutionarily conserved.

Figure 2-4 legend.

(A) Pyrophosphorolysis in ECs containing *Saccharomyces cerevisiae* RNAPII. The most and least pyrophosphorolysis sensitive scaffolds are shown at the top (yields $EC9^{GU}$ and $EC9^{UG}$). *, 5′ ^{32}P. *Sce*$EC9^{GU}$ and *Sce*$EC9^{UG}$ were reconstituted and subjected to pyrophosphorolysis as described in Materials and Methods, except the pyrophosphorolysis reactions were conducted at 30 °C (with 0.5 mM PPi). Samples were withdrawn and separated by denaturing electrophoresis at the times indicated. The "C" lane in each panel is a chase sample in which 1 mM UTP (the next cognate NTP) was added at the end of pyrophosphorolysis to confirm that the remaining ECs were active by showing the RNA can be extended by one nt.

(B) Pyrophosphorolysis in ECs containing calf thymus RNAPII. *Bta*$EC9^{GU}$ and *Bta*$EC9^{UG}$ were reconstituted and subjected to pyrophosphorolysis as described for panel A.

TABLE 2-2. Rate of pyrophosphorolysis by 50 nM EC containing different RNAPs at 0.5 mM PPi.

RNAP	Temperature	EC9^{GU}			EC9^{UG}
		k_1	k_{-1}	K_{eq}	k_1 [1]
		min^{-1}			min^{-1}
*Tth*RNAP	60 °C	0.36 ± 0.09	0.17 ± 0.05	2.1 ± 0.8	< 0.007
*Eco*RNAP	37 °C	0.01 ± 0.003	0.024 ± 0.007	0.4 ± 0.2	< 0.001
*Sce*RNAPII	30 °C	0.003 ± 0.0007	0.006 ± 0.004	0.5 ± 0.4	< 0.001
*Bta*RNAPII	30 °C	0.007 ± 0.001	0.009 ± 0.003	0.8 ± 0.3	< 0.001

[1]Pyrophosphorolysis was not detectable. An upper limit on the possible rate was estimated based on experimental error in measurements on the EC9^{GU} scaffold.

NTP accumulation causes incomplete pyrophosphorolysis

We assumed that incomplete pyrophosphorolysis (*e.g.*, Figure 2-3) could be explained by reverse-reaction (nucleotide addition) with NTPs generated during pyrophosphorolysis because the K_{NTP} is low and nucleotide addition is fast (kcat≈103 s-1; Refs. Foster et al., 2001; Vassylyev et al., 2007b). To test this hypothesis, we examined pyrophosphorolysis of $EC9^{GU}$ in the presence of apyrase. Apyrase rapidly hydrolyzes NTPs to NDPs (Molnar and Lorand, 1961). We used *Eco*RNAP instead of *Tth*RNAP for these experiments because the optimal temperature for apyrase is 30 °C, a temperature at which *Tth*RNAP is relatively inactive. *Eco*$EC9^{GU}$ exhibited sensitivity to pyrophosphorolysis similar to *Tth*$EC9^{GU}$, but with a higher unreacted fraction (~65% *vs* ~30%; compare Figure 2-5A to Figure 2-3C). If the 65% unreacted fraction was due to reverse-reaction with NTP, addition of apyrase should shift the reaction toward completion. Consistent with our hypothesis, addition of apyrase at the beginning of the reaction eliminated the plateau at partial completion and gave a pseudo-first-order rate of 0.016 ± 0.001 min^{-1} for the most of reaction (Figure 2-5A). Addition of apyrase after the plateau had been achieved (at 60 min) caused resumption of pyrophosphorolysis, unambiguously establishing that incomplete reaction reflected accumulation of NTP. Since initial concentration of *Eco*$EC9^{GU}$ is 50 nM and 65% of it remains unreactive when the reaction reaches the equilibrium, we infer that the net forward flux of 0.5 mM PPi + ~30 nM *Eco*$EC9^{GU}$ is equal to the reverse flux from 20 nM UTP + 20 nM *Eco*$EC8^{G}$. We noted, however, that at later times in the reaction (after ~2 hours) the equilibrium in the absence of apyrase was replaced by further pyrophorphorolysis and the rate of pyrophosphorolysis in the presence of apyrase slowed, most likely because one or more reaction components degraded upon prolonged incubation.

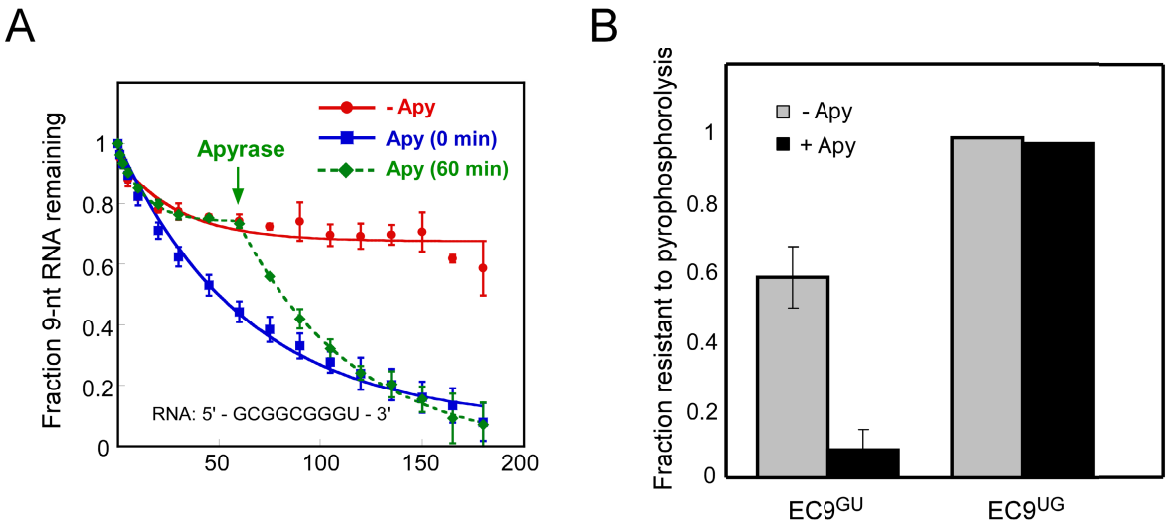


Figure 2-5. Incomplete pyrophosphorolysis by *E.coli* RNAP is caused by NTP accumulation.

(A) Rates of pyrophosphorolysis of *Eco*EC9^{GU}. *Eco*EC9^{GU} reconstituted using *E.coli* RNAP and the pre-translocation favoring scaffold (#6063, #5848, #6046; Table 2-3). *Eco*EC9^{GU} (50 nM) was incubated at 37 °C with 0.5 mM PPi alone (red), in the presence of 0.5 U apyrase/ml (blue), or with addition of apyrase to 0.5 U/ml 60 min after the reaction was initiated (green) for the time indicated. Samples were removed at the times indicated and the fraction EC9^{GU} remaining was plotted as a function of time. The error bars represent standard deviations obtained from 4 different experiments. The data from reaction without apyrase (red) are fit to a simple reversible mechanism of pyrophosphorolysis (equation 3; see Materials and Methods), whereas data from reaction with apyrase (blue) are fit to a single exponential for a pseudo-first-order reaction. (B) The fraction of *Eco*EC9^{GU} and *Eco*EC9^{UG} resistant to pyrophosphorolysis at 0.5 mM PPi in the absence (gray) or presence (black) of apyrase.

We next tested whether the resistance of *Eco*EC9^{UG} to pyrophosphorolysis could possibly be explained by an exceptionally low K_{GTP} for the *Eco*EC8^{U} + GTP reaction. To test this possibility, we performed pyrophosphorolysis of *Eco*EC9^{UG} in the presence of apyrase in parallel to reactions in which apyrase made pyrophosphorolysis of *Eco*EC9^{GU} go to completion. Although we observed the expected shift in *Eco*EC9^{GU} completion, no reaction of *Eco*EC9^{UG} was observed even in the presence of apyrase (Figure 2-5B and Figure 2-6A). We verified that apyrase effectively destroyed NTPs by preincubating 200 nM UTP with apyrase under the same conditions used for pyrophophorolysis, then adding *Eco*EC9^{UG}, and observing that elongation occurred only when apyrase was omitted (Figure 2-6B). We concluded that the resistance of *Eco*EC9^{UG} to pyrophosphorolysis is an inherent property of *Eco*EC9^{UG}, most likely a strong preference for the posttranslocated register, and not to a strong reverse-reaction of *Eco*EC8^{U} with GTP.

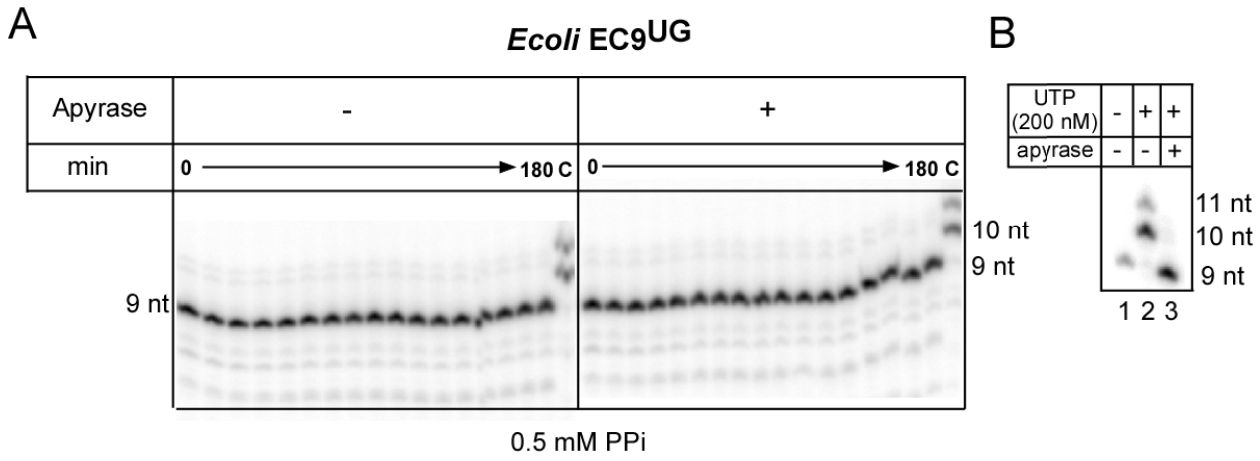


Figure 2-6. Lack of effect of apyrase on pyrophosphorolysis of EC9^{UG}.

(A) Pyrophosphorolysis in the presence or absence of apyrase of EC9^{UG} (pyrophosphorolysis insensitive EC) reconstituted with *Eco*RNAP and the post-favored scaffold (#6070, #5848, #6053; Table 2-3) containing 5′- ^{32}P-labeled 9nt RNA (0 min). EC9^{UG} (50 nM) was incubated with 0.5 mM pyrophosphate only (left), or with 0.5 mM pyrophosphate and 0.5 unit of apyrase/ mL (right) at 37 °C for 1, 2.5, 5, 10, 20, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180 min. The "C" lane represents the chase lane in which 1 mM of the next cognate NTP is added at the end of pyrophosphorolysis to confirm that the remaining complexes are active and RNA can be extended by one nt.

(B) Activity of apyrase under the conditions of pyrophosphorolysis reactions (see Materials and Methods). EC9^{UG} was first reconstituted with *Eco*RNAP as in panel A (lane 1). Addition of 200 nM UTP caused extension of all EC9^{UG} to EC10^{UGU} (lane 2; with some misincorporation to produce EC11). The action of apyrase on low concentration NTP was confirmed by preincubating 200 nM UTP with apyrase for 30 min before mixing with 46 nM EC9^{UG} (lane 3). The absence of extension on lane 3 confirms that apyrase destroyed all UTP.

Effects of RNA 3´-dinucleotide sequence on pyrophosphorolysis were preserved on a complete nucleic acid scaffold

We used minimal nucleic-acid scaffolds in initial experiments to avoid complications from transcription bubble energetics. To test whether the strong effects of RNA 3´ dinucleotide sequence persisted on complete scaffolds containing fully complementary DNA strands, a transcription bubble, and upstream and downstream duplexes, we tested complete-scaffold versions of the PPi-resistant EC9^{UG} and the PPi-sensitive EC9^{GU} reconstituted ECs with *Tth*RNAP or *Eco*RNAP (Figure 2-7A and Table 2-3). Similar to the minimal scaffolds, *Eco*EC9^{UG} (<0.002 min^{-1}) and *Tth*EC9^{UG} (0.12 ± 0.04 min^{-1}) were less sensitive to pyrophosphorolysis than *Eco*EC9^{GU} (0.034 ± 0.004 min^{-1}) and *Tth*EC9^{GU} (1.7 ± 0.1 min^{-1}) (Figure 2-7B). In contrast to the minimal scaffolds, however, both *Eco*EC9^{UG} and *Tth*EC9^{UG} exhibited measurable rates of pyrophosphorolysis. We also tested the effects of apyrase on the complete scaffold version of *Eco*EC9^{GU} and verified that apyrase addition caused the 40% unreactive *Eco*EC9^{GU} to complete pyrophosphorolysis (Figure 2-8). We drew two conclusions from these results. First, pyrophosphorolysis is faster in ECs reconstituted with complete scaffolds than in those with minimal scaffolds. This could reflect a greater bias of complete *vs.* minimal scaffolds toward the pretranslocated register for these specific complexes or a lower activation barrier to formation of the trigger helices in the presence of an intact fork junction (trigger helices formation strongly stimulates the rate of pyrophosphorolysis (Vassylyev et al., 2007b). Second, faster pyrophosphorolysis and presumably greater bias toward the pretranslocated register for the RNA 3´dinucleotide GU *vs*. UG occurs on both minimal and complete scaffolds.

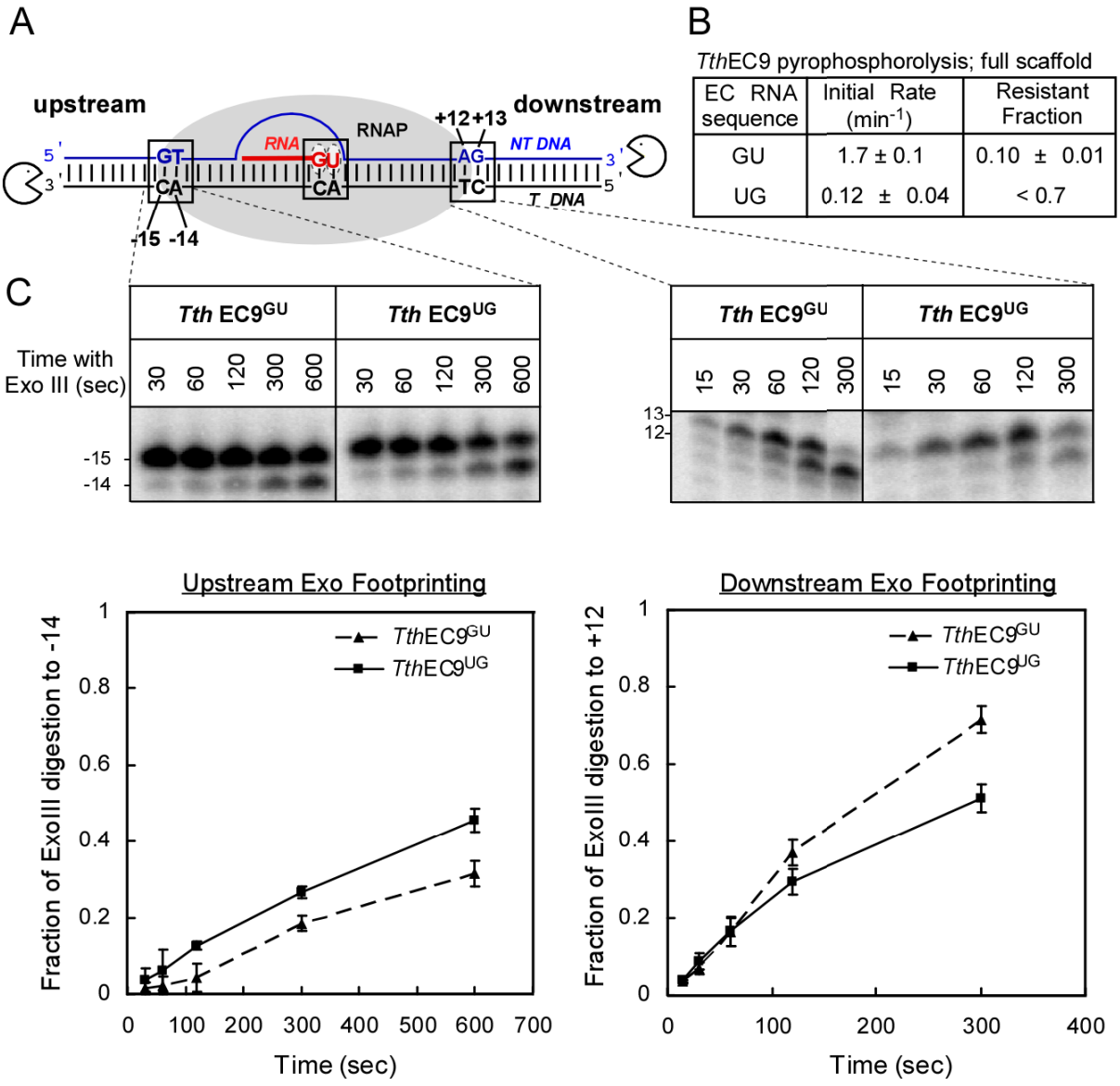


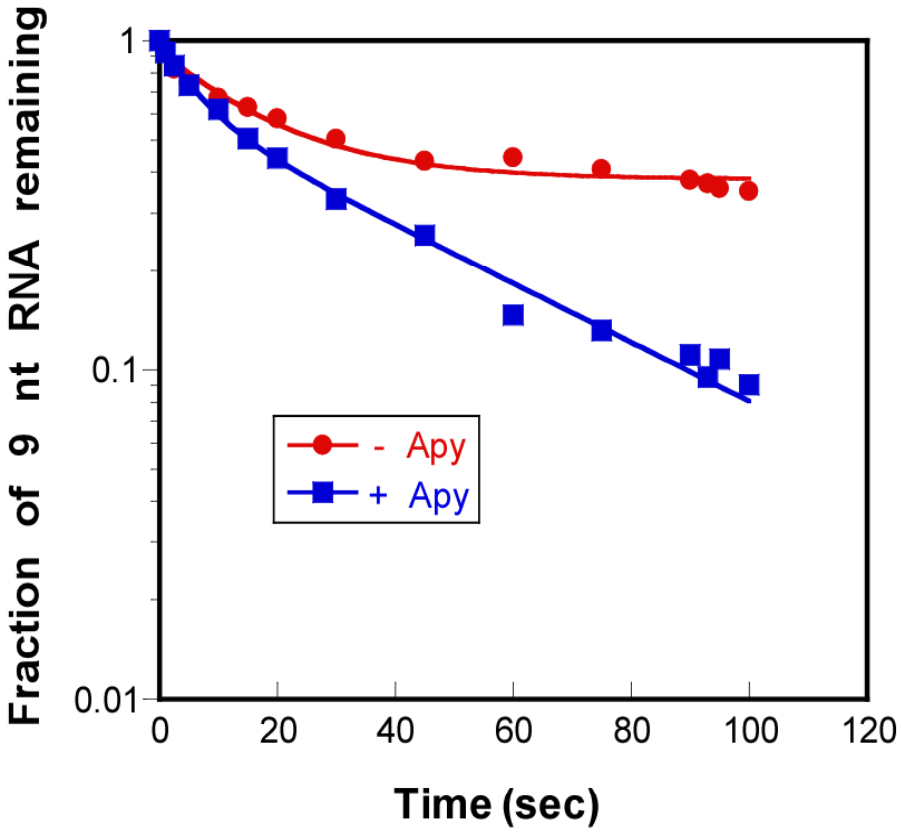
Figure 2-7. ExoIII footprints of *Tth*EC9^{GU} (pretranslocated) and *Tth*EC9^{UG} (posttranslocated).

Figure 2-7 legend.

(A) Schematic of the complete scaffold used for ExoIII footprinting experiments. The sizes of the template DNA (upstream) and the nontemplate DNA (downstream) fragments protected by RNAP from digestion by ExoIII are illustrated. For the nontemplate strand assay, the nontemplate strand contained a 5′ ^{32}P label and the template strand contained a 3′ phosphorothioate bond. For template strand assay, the template strand contained a 5′ ^{32}P label and the nontemplate strand containing a 3′ phosphorothioate bond.

(B) Pyrophosphorolysis results of EC9^{GU} and EC9^{UG} on complete scaffolds used for ExoIII footprinting experiments. The RNA and template DNA (GU: #6046, #6355, and UG: #6053, #6357; Table 2-3) were first annealed, then mixed with *Tth*RNAP, then annealed to the nontemplate strand (GU: #6354 and UG: #6356; Table 2-3; see Materials and Methods). ECs (50 nM) were incubated with 0.5 mM PPi at 60°C. The initial rate of pyrophosphorolysis (k_1) and the fraction resistant to pyrophosphorolysis were determined as described in Materials and Methods.

(C) Upstream ExoIII footprinting (left) and downstream ExoIII footprinting (right) of EC9^{GU} and EC9^{UG}. Plot on the left depicts the appearance of the nontemplate DNA -14 band, which should be faster when the pretranslocated register is favored. Plot on the right depicts the appearance of the template DNA +12 band, which should be faster when the posttranslocated register is favored.




```
                    GCGGCGGGU
5'-GGTCAGTACGTCCTAA          TCGATTCAGACAGG-3'
3'-CCAGTCATGCAGGATTCGCCGCCCAAGCTAAGTCTGTCC-5'
                    GCGGCGGGU
```

Figure 2-8. Incomplete pyrophosphorolysis caused by NTP accumulation using a complete scaffold.

Figure 2-8 legend.

$EC9^{GU}$ was reconstituted on the pre-favored complete scaffold (#6355, #6354, #6046; Table 2-3) with *Eco*RNAP. Pyrophosphorolysis was conducted in the presence (blue) or absence (red) of apyrase to 0.5 unit/mL at 37 °C (see Materials and Methods). The remaining fraction of 5′-end labeled 9 nt RNA of $EC9^{GU}$ was plotted as a function of time. The pre-favored nucleic acid scaffold is shown below the plot with template and nontemplate DNA in black and RNA in red.

Direct ExoIII footprinting confirmed translocation bias due to RNA 3′ dinucleotide sequence

To verify that the differences in pyrophosphorolysis sensitivity were due to differences in translocation register and not to inherent differences in the chemical reactivities of RNA 3′ dinucleotides with PPi or to sequence-specific effects of the RNAP catalytic center on pyrophosphorolysis, we assayed translocation register using the independent criterion of exonuclease III (ExoIII) footprinting (Kashkina et al., 2006; Kireeva et al., 2008; Landick and Yanofsky, 1987). ExoIII is a double-strand-specific deoxyribonuclease that processively digests a single-strand of DNA from the 3′ end and detects the boundary of RNAP on DNA when bound RNAP inhibits access of the ExoIII active site to the DNA phosphodiester backbone. Because RNAP may be in rapid oscillation between pre and posttranslocated registers and because ExoIII exhibits some sequence specificity in cleavage rates, assessment of translocation register with ExoIII requires measuring the rates of ExoIII digestion rather than arbitrary reaction endpoints (Kireeva et al., 2008).

We determined the upstream and downstream ExoIII boundaries for *Tth*EC9^{GU} and *Tth*EC9^{UG} reconstituted on complete scaffolds (Figure 2-7A). On this scaffold, an EC with greater pretranlocation bias will exhibit a greater barrier to ExoIII digestion of -15 to -14 on the template strand and a weaker barrier to ExoIII digestion of +13 to +12 on the nontemplate strand (Figure 2-7A). The upstream footprint slowly shifts one bp downstream in both *Tth*EC9^{GU} and *Tth*EC9^{UG} but the *Tth*EC9^{GU} presents a stronger barrier that results in an initial lag in the rate of the shift and a difference in the final distribution of DNA fragments (Figure 2-7C). This is consistent with the hypothesis that *Tth*EC9^{GU} was more biased in the pretranslocated direction than *Tth*EC9^{UG}. A different effect was observed during ExoIII digestion of downstream DNA,

with the initial rates of digestion being similar, but *Tth*EC9^{UG} posing a stronger barrier as the reaction progressed. This result also is consistent with *Tth*EC9^{GU} being more pretranslocated than *Tth*EC9^{UG}. These ExoIII footprinting assays are less clear-cut than one might hope, but the consistent pattern of effects strongly supports the interpretation that RNA 3′ GU more favors the pretranslocated register than RNA 3′ UG.

EC9^{UG} resistance to pyrophosphorolysis was not caused by backtracking

Although our results strongly favored the idea that EC9^{UG} is biased toward the posttranslocated register, we wished to rule out the alternative possibility that EC9^{UG} is backtracked and therefore resistant to pyrophosphorolysis. To this end, we performed hydrolytic transcript cleavage in *Tth*EC9^{UG} reconstituted on a minimal scaffold. The transcript hydrolysis (*i.e.,* intrinsic cleavage) reaction cleaves nascent RNA at the phosphodiester bond located in the RNAP active site so that a pretranslocated EC yields a one nt 3′ cleavage product and a backtracked EC yield a larger 3′ cleavage product (Figure 2-9A); the hydrolysis reaction is faster at elevated pH and high Mg^{2+} concentration (Orlova et al., 1995; Sosunov et al., 2003; Surratt et al., 1991; Zhang et al., 2010). We first generated *Tth*EC9^{UG} and *Tth*EC9^{GU} in transcription buffer at pH 9 lacking Mg^{2+} and then initiated transcript cleavage by addition of Mg^{2+} to 20 mM. No cleavage products were observed for *Tth*EC9^{UG}, whereas *Tth*EC9^{GU} generated the one-nt cleavage product expected for a pretranslocated EC (Figures 2-9B and C). We confirmed that these ECs were fully reconstituted and active by extending the transcripts from 9 to 10 nt by incubation with UTP (Figure 2-9D). Further, cleavage products were not observed when RNAP was omitted (no enzyme control; Figure 2-10). The absence of cleavage products >1 nt is consistent with strong preference of the RNA 3′ GU dinucleotide for the pretranslocated register

relative to the GG dinucleotide that would occupy the active site if the EC were to backtrack by 1 bp. We concluded that *Tth*EC9^{UG} is not backtracked and that the simplest explanation for its resistance to pyrophosphorolysis is bias toward the posttranslocated register.

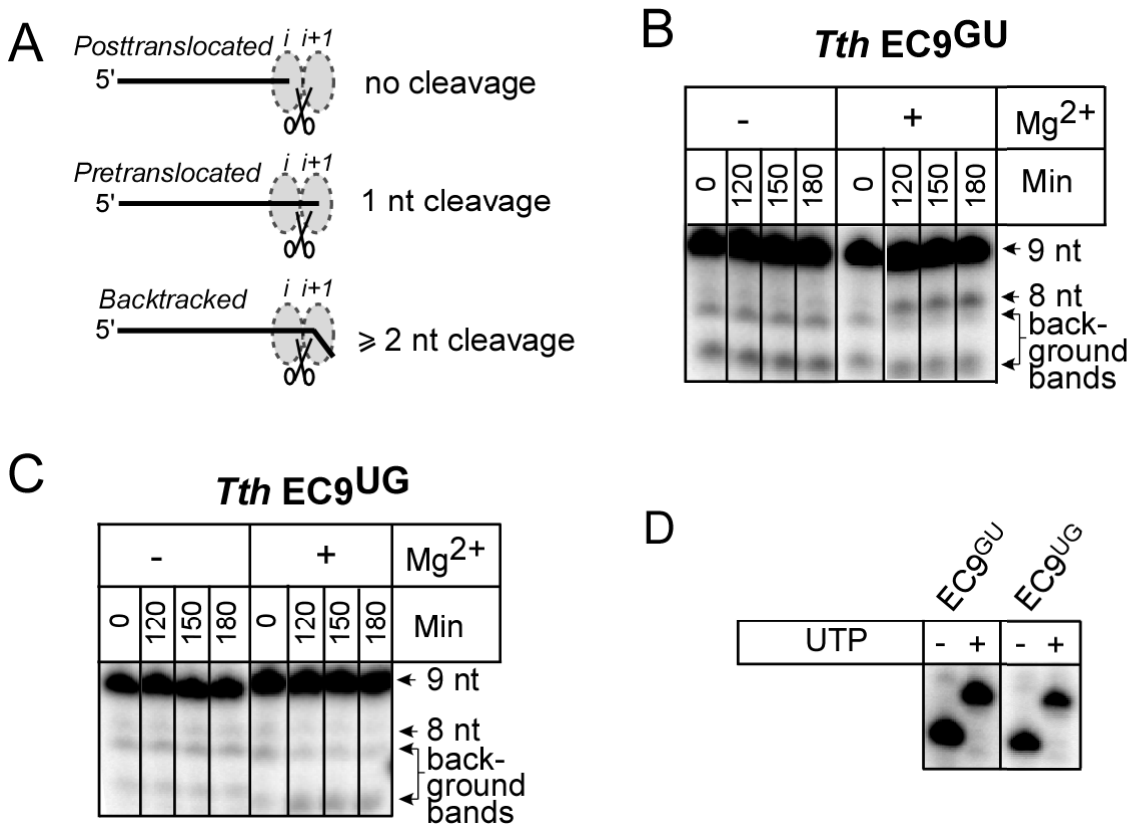


Figure 2-9. The pyrophosphorolysis-resistant *Tth*EC9^{UG} is not backtracked.

(A) Predicted outcomes for intrinsic cleavage reactions of pretranslocated, posttranslocated, and backtracked ECs. Posttranslocated ECs does not generate cleavage products, whereas pretranslocated and backtracked complexes produce 1-nt cleavage products, and 2 or more cleavage products, respectively.

(B) Intrinsic cleavage reaction of *Tth*EC9^{GU}. *Tth*EC9^{GU} was reconstituted using *Tth*RNAP and the pretranslocation-favoring scaffold (#6063, #5848, #6046; Table 2-3) containing 5′ end ^{32}P-labeled RNA at pH 9 without Mg^{2+}. Intrinsic cleavage reaction was initiated by the addition of 20 mM Mg^{2+} and samples were removed and separated by electrophoresis at the times indicated (see Materials and Methods).

Figure 2-9 legend (cont.)

(C) Intrinsic cleavage reaction of *Tth*EC9^{UG}. *Tth*EC9^{UG} was reconstituted and assayed as in panel A except with the postranslocation-favoring scaffold (#6070, #5848, #6053: Table 2-3).

(D) EC9^{GU} and EC9^{UG} were incubated with 100 μM UTP to extend their respective RNAs by one nucleotide, thus showing the ECs were active and not arrested.

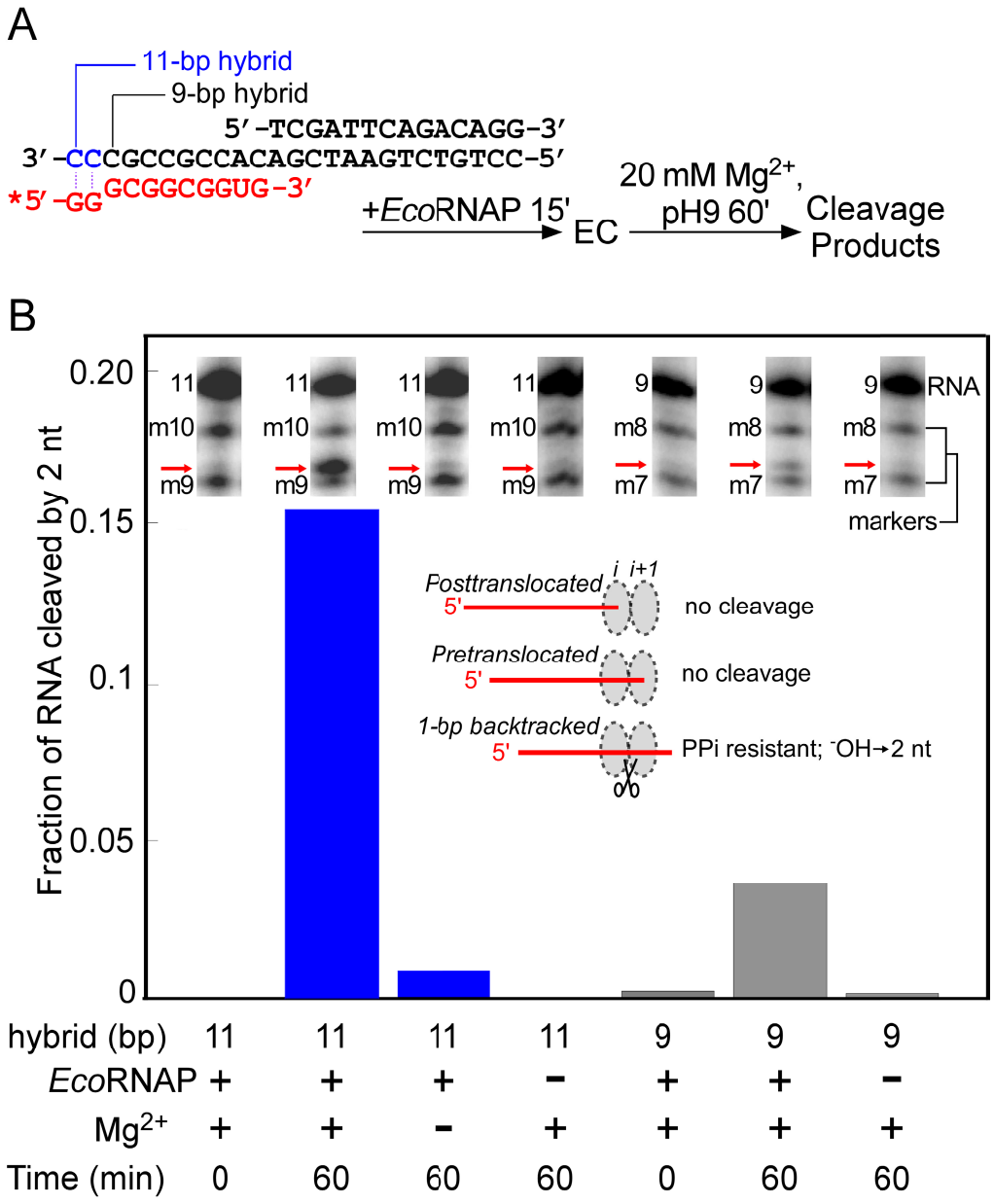


Figure 2-10. Extension of the RNA:DNA hybrid past 9-bp shift translocation bias.

Figure 2-10 legend.

Intrinsic cleavage reaction of *Eco*EC9^{UG} (9-bp hybrid) and *Eco*EC11^{UG} (11-bp hybrid).

(A) Nucleic acid scaffolds with different potential RNA:DNA hybrid length are shown above the plot. RNA is shown in red, template and nontemplate DNA strands are in black. The 9-bp and 11-bp potential RNA:DNA hybrid are created by varying the length of template DNA without changing the length of the 11-nt RNA (the template DNAs in 9-bp and 11-bp hybrids end at the positions marked with a black line, or a blue line, respectively). RNA was labeled with ^{32}P at its 5′ end (*).Experimental set up of intrinsic cleavage reaction is illustrated. Nucleic acid scaffolds with 9-bp hybrid or 11-bp hybrid (50 nM) were incubated with *Eco*RNAP at 37°C for 15 min to reconstitute ECs *in vitro*. ECs were then incubated with 20 mM Mg^{2+} or without Mg^{2+} at 37°C for 60 minutes to initiate intrinsic RNA cleavage reaction (see Materials and Methods).

(B) Histograms indicate the fraction of 2-nt cleaved RNA by *Eco*EC11^{UG} (blue) and *Eco*EC9^{UG} (gray). The reaction conditions are indicated below the graph. Representative gel panels for each corresponding histogram and outcomes for intrinsic cleavage reactions of posttranslocated, pretranslocated, and 1bp-backtracked ECs are shown in the inset.

Inhibiting forward translocation with extended hybrids makes EC9^{UG} PPi-sensitive

We reasoned that if bias toward the posttranslocated register in EC9^{UG} rather than an inherently slow catalytic reaction explained its pyrophosphorolysis resistance (Figure 2-11A), then it might be possible to increase the rate of pyrophosphorolysis by increasing the length of the RNA:DNA past 9 bp. A longer hybrid, especially one with GC-bp at the upstream end, should be harder to melt and thus be shifted toward pretranslocated register to fit the longer hybrid within the RNAP main channel. To test this prediction, we used *Eco*RNAP to allow use of apyrase (to reduce potential interference from reverse nucleotide addition) and measured pyrophosphorolysis of *Eco*EC9^{UG}, *Eco*EC10^{UG}, and *Eco*EC11^{UG} reconstituted on minimal scaffolds (Figure 2-11B). These scaffolds use the same 11-nt, 3′-UG RNA and differ only in potential hybrid length due to changes in the 3′ portion of the template strands, which are 3′-GCC··· for EC9^{UG}, 3′-CGCC··· for EC10^{UG}, and 3′-CCGCC··· for EC11^{UG}. In agreement with the idea that pyrophosphorolysis resistance of EC9^{UG} reflects translocation bias, EC10^{UG} exhibited increased sensitivity to PPi and EC11^{UG} exhibited a greater increase in sensitivity to PPi (Figure 2-11C). To confirm that longer hybrids inhibit forward translocation, we also performed hydrolytic transcript cleavage in *Eco*EC9^{UG} and *Eco*EC11^{UG}. As expected, *Eco*EC11^{UG} generated a two-nt 3′ cleavage product, which indicated that forward translocation is hindered in this EC, whereas *Eco*EC9^{UG} gave little cleavage product (Figure 2-10). These results are inconsistent with inherence resistance of the 3′-UG dinucleotide to pyrophosphorolysis but consistent with the view that posttranslocation bias confers PPi-resistance on EC9^{UG}. Pyrophosphorolysis of EC11^{UG} was noticeably biphasic. This could either reflect structural heterogeneity in the elongation complexes or an inability of apyrase to hydrolyze NTPs rapidly at the concentrations generated in this experiment.

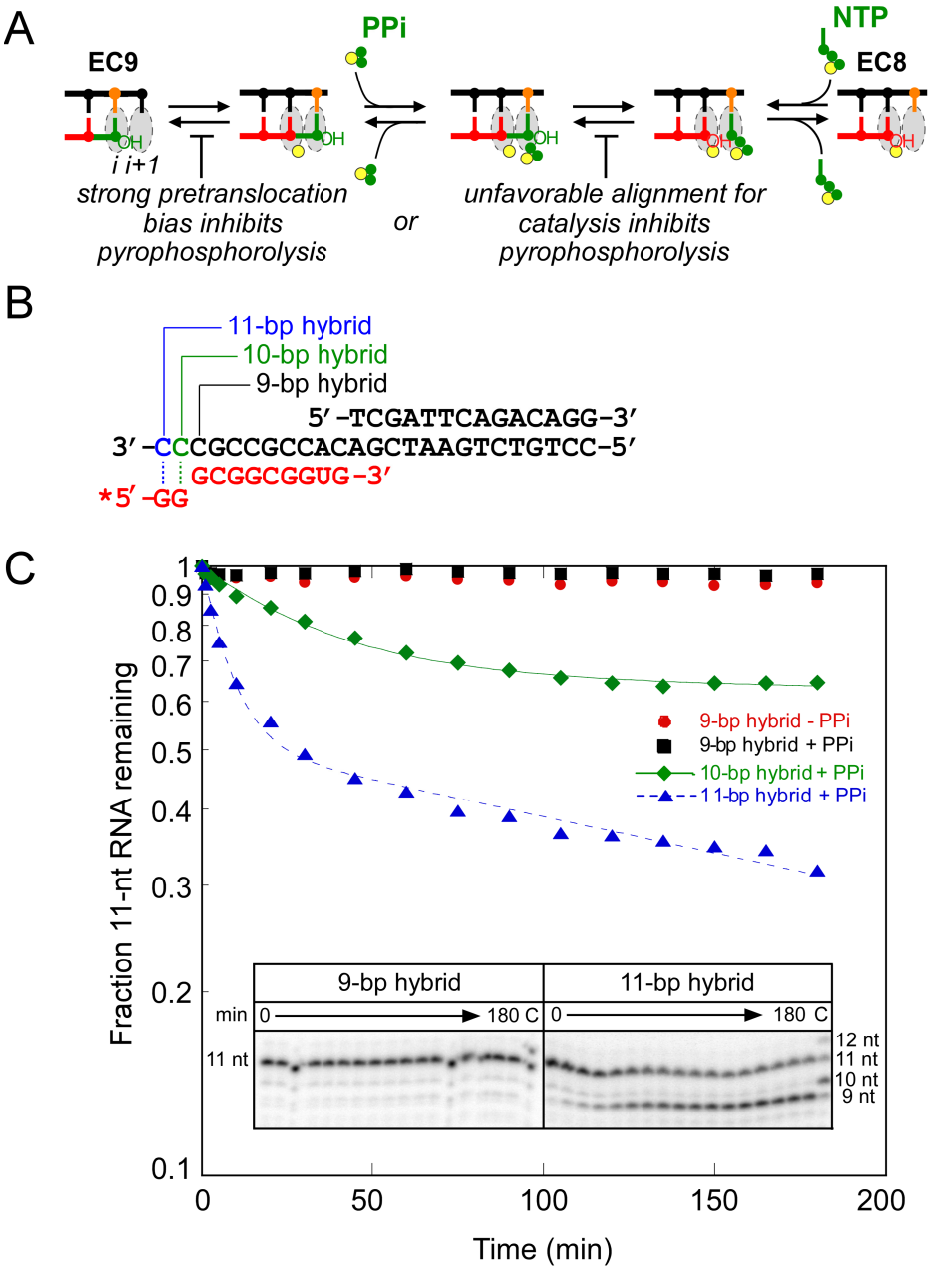


Figure 2-11. The RNA 3′ dinucleotide UG is not intrinsically resistant to pyrophosphorolysis by *Eco*RNAP.

Figure 2-11 legend.

(A) translocation bias towards posttranslocated state or a slow catalysis could account for the insensitivity of $EC9^{UG}$.

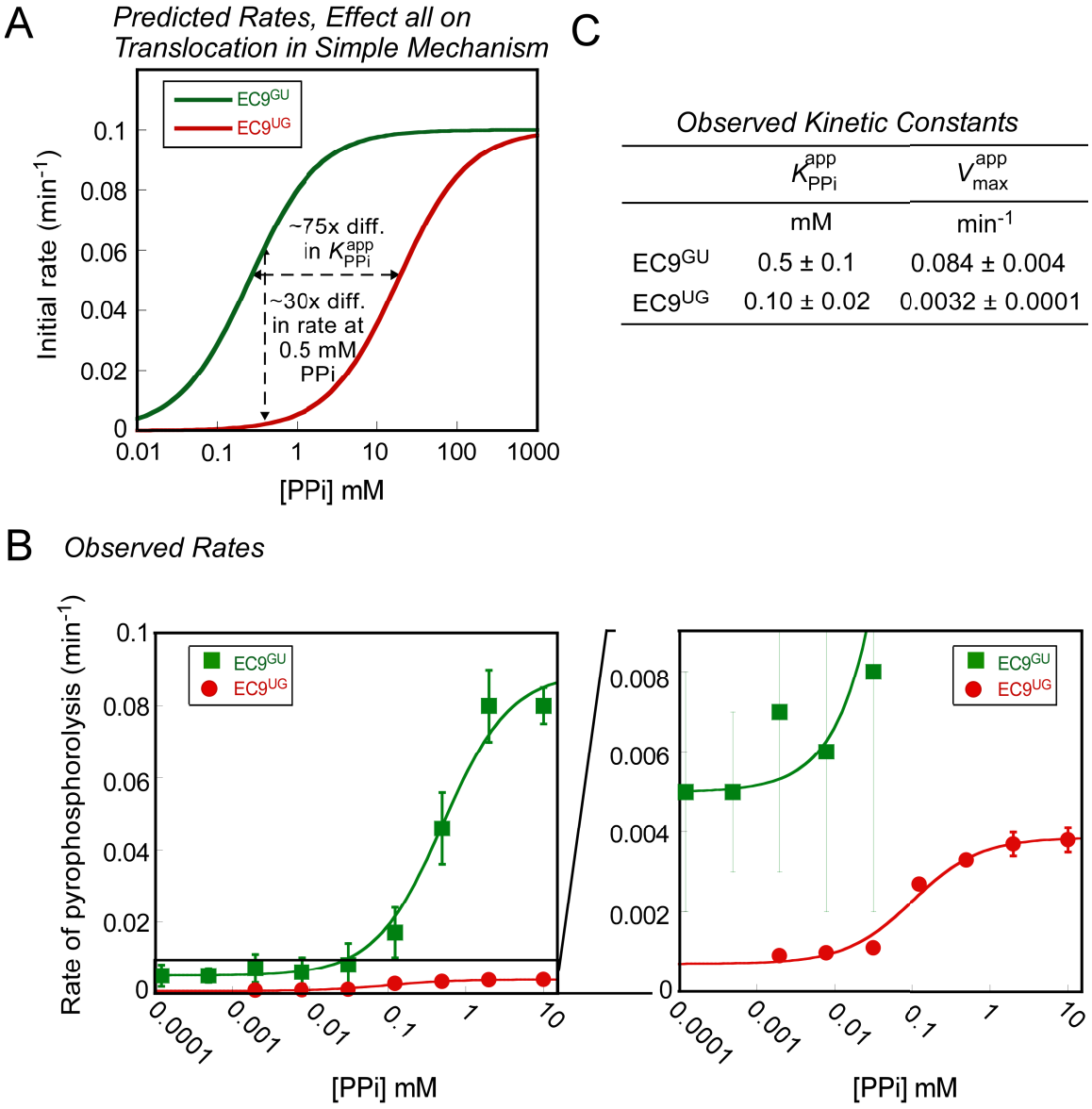
(B) Nucleic acid scaffolds with different potential RNA:DNA hybrid length. RNA is shown in red, template and nontemplate DNA strands are in black. The 9-bp, 10-bp, and 11-bp potential RNA:DNA hybrid are created by varying the length of template DNA without changing the length of the 11-nt RNA (the template DNAs in 9-bp, 10-bp, and 11-bp hybrids end at the positions marked with a black line, a green, or a blue line, respectively. RNA was labeled with ^{32}P at its 5′ end (*).

(C) Plot of pyrophosphorolysis of ECs reconstituted with *Eco*RNAP and different scaffolds shown in Figure 2-9B. ECs were incubated with 0.5 mM PPi, 0.5 U apyrase/ml at 37 °C for the times indicated. Representative gel panels for $EC9^{UG}$ (9-bp hybrid) and EC11UG (11-bp hybrid) are shown in the inset. “C”, chase lane in which ECs were chased with 1 mM UTP at the end of pyrophosphorolysis time course (after 3 hours).

We note that the 2-nt cleavage interval observed for *Eco*EC9^{UG} and *Eco*EC11^{UG} (Figure 2-10) differs from the 1-nt hydrolytic cleavage interval observed for *Tth*EC9^{GU} (Figure 2-9). This could reflect, at least in part, differences in *Eco* and *Tth* RNAPs. However, it is notable that our analysis of RNA 3′ dinucleotide effects on pyrophosphorolysis (Figure 2-3C and Table 2-1) suggested that the UG dinucleotide disfavors the pretranslocated register. Upon a 1 nt backtrack, the active site of *Eco*EC9^{UG} and *Eco*EC11^{UG} would be occupied by the GU dinucleotide that appears to favor the pretranslocated register. Thus, the observed difference in hydrolytic cleavage intervals is consistent with favorable interactions of the GU dinucleotide and unfavorable interactions of the UG dinucleotide in the pretranslocated active site.

PPi concentration-dependence of pyrophosphorolysis is inconsistent with ordered translocation and PPi binding and effects of 3′ dinucleotide only on translocation bias

A simple model in which rapid translocation equilibrium is linked to PPi binding to the pretranslocated EC (Figure 2-11A) predicts that increasing PPi concentration should reduce occupancy of the posttranslocated state. Indeed, studies of the single-subunit polymerases HIV-1 reverse transcriptase and T7 RNAP show that binding of PPi or the PPi-analog foscarnet does shift the translocation bias towards the pretranslocated state, suggesting these single-subunit polymerases adhere to the predictions of this simple model (Guo and Sousa, 2006),(Marchand et al., 2007). If such a model operates for multisubunit RNAPs and if all the effects of RNA 3′ dinucleotide sequence are on translocation bias, then a greater posttranslocation bias should increase apparent K_{PPi} by reducing the fraction of time the pretranslocated state is available to bind PPi. Thus, in this scenario, EC9^{UG} should undergo pyrophosphorolysis if incubated at a high [PPi] and it should exhibit a higher apparent K_{PPi} than EC9^{GU} (*e.g.*, Figure 2-12A).



	K_{PPi}^{app}	V_{max}^{app}
	mM	min^{-1}
$EC9^{GU}$	0.5 ± 0.1	0.084 ± 0.004
$EC9^{UG}$	0.10 ± 0.02	0.0032 ± 0.0001

Figure 2-12. PPi-concentration dependence of pyrophosphorolysis in *Eco*EC9^{GU} and *Eco*EC9^{UG}.

Figure 2-12 legend.

(A) A predicted relationship for the rate of pyrophosphorolysis *vs.* [PPi] for the ordered translocation/PPi binding mechanism shown in Figure 2-11A and Figure 2-1. The difference in apparent K_{PPi} (~75x) between $EC9^{GU}$ and $EC9^{UG}$ and in rate at 0.5 mM PPi concentration are indicated on the plot.

(B) Kinetic analysis of [PPi]-dependence of pyrophosphorolysis in $EC9^{GU}$ and $EC9^{UG}$. Rate of pyrophosphorolysis for *Eco*$EC9^{GU}$ and *Eco*$EC9^{UG}$ over a wide range of PPi concentrations.

(C) Approximate, apparent kinetic constants of $EC9^{GU}$ and $EC9^{UG}$ are calculated from the plot shown in Figure 2-12B (see Materials and Methods).

To investigate this prediction, we sought to measure the PPi-concentration dependence of pyrorophosphorolysis for $EC9^{UG}$ and $EC9^{GU}$ using *Eco*RNAP and complete nucleic acid scaffolds (Figure 2-7A and Figure 2-13). Although we included apyrase in these reactions, we found it was not possible to obtain accurate kinetic data at low or high PPi concentrations. At low PPi concentration, the background rate of intrinsic transcript hydrolysis prevented detection of slow pyrophosphorolysis (Figure 2-13). At high PPi concentation, the reactions of $EC9^{GU}$ became biphasic (Figure 2-13). The biphasic kinetics could be explained by structural heterogeneity of the reconstituted ECs or by an inability of apyrase to degrade NTPs rapidly at very low NTP concentration (or the combination of both; Figure 2-13). We favor the latter view, but were unable to establish it conclusively.

Since complete kinetic profiling of pyrophosphorolysis proved difficult, we instead estimated initial rates of pyrophosphorolysis by fitting the relative [EC9] at different [PPi] using the mechanism 4 in the program KinTek Global Kinetic Explorer (See Materials and Methods). These calculations yielded estimates of initial rates that saturated at ~0.08 min^{-1} for $EC9^{GU}$ and ~0.003 min^{-1} for $EC9^{UG}$ (Figures 2-12B and C). Although these data could suggest that the intrinsic rate of $EC9^{UG}$ is inherently slow even at saturated PPi, such a conclusion is inconsistent with the observation that $EC11^{UG}$ reacts rapidly with even 0.5 mM PPi (>0.04 min^{-1}, Figure 2-11C, *vs.* <0.002 min^{-1} for $EC9^{UG}$, Table 2-2). Further, the apparent half-maximal concentration of PPi (K_{PPi}^{app}) was actually lower for $EC9^{UG}$ (0.1 ± 0.02 mM) than for $EC9^{GU}$ (0.5 ± 0.1 mM) (Figures 2-12B and C). Taken together, these results are inconsistent with a simple model in which a rapid equilibrium of pre and posttranslocated states is coupled to PPi binding to the pretranslocated EC and in which posttranslocation bias completely explains the pyrophosphorolysis of $EC9^{UG}$. Such a model (*e.g.*, Figure 2-11A) predicts a roughly 75-fold

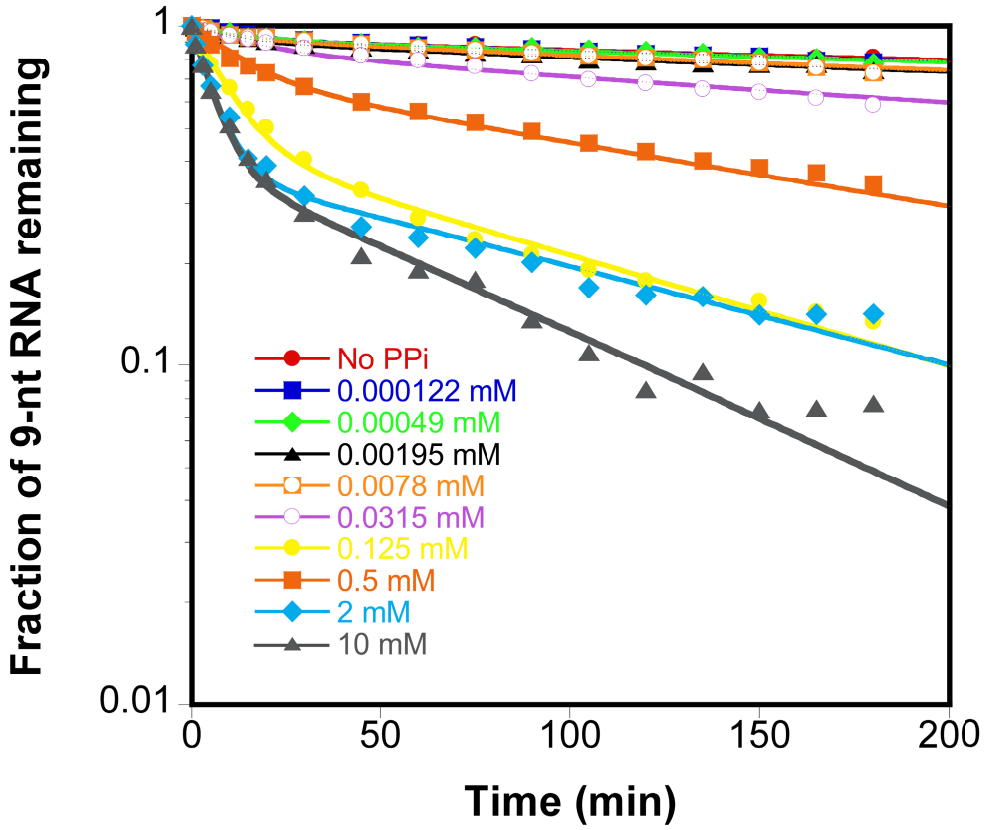


Figure 2-13. PPi-concentration dependence of pyrophosphorolysis of *Eco*EC9^{GU} reconstituted on a complete nucleic acid scaffold.

*Eco*EC9^{GU} was reconstituted as described in Materials and Methods using 5′-end labeled RNA and oligos #6355, #6354, and #6046 (Table 2-3). Pyrophosphorolysis was conducted as described in Materials and Methods in the presence apyrase (0.5 unit/mL) and increasing concentrations of PPi. The concentration of Mg^{2+} was adjusted for 10 mM PPi by addition of $MgCl_2$ to the final concentration of 10 mM. EC9^{GU} (50 nM) were incubated with indicated concentrations of pyrophosphate at 37 °C for the time indicated. To measure the background rate of intrinsic transcript hydrolysis, EC9^{GU} is incubated with the reaction buffer alone (EB buffer; Materials and Methods) at 37 °C for the time indicated (*red*; No PPi). The fraction remaining EC9^{GU} was plotted as a function of time.

greater K_{PPi}^{app} for $EC9^{UG}$ *vs.* $EC9^{GU}$ to account for the 30-fold faster observed rate of pyrophosphorolysis at 0.5 mM PPi (Figure 2-12A).

These results led us to ask if it is possible to explain the observed rates of pyrophosphorolysis for $EC9^{UG}$ and $EC9^{GU}$ if translocation and PPi binding were not ordered events; that is, if PPi could bind to either the pre or posttranslocated EC. Although PPi release has been postulated to cause translocation at least in single-subunit RNAPs (Yin and Steitz, 2004), we are unaware of data establishing the order of PPi release and translocation for multisubunit RNAPs. To evaluate the consequences of a random order for translocation and PPi binding/release, we considered the hypothetical reaction scheme in which PPi can bind to pre- and posttranslocated ECs (Figure 2-14A). We considered only the simplest version of such a mechanism in which pre- and posttranslocated ECs bind PPi with equal affinity and in which the translocation equilibrium is unaffected by PPi binding. Even with these constraints, we found that the random-order mechanism yielded predictions that were remarkably similar to our observed data by assuming a 50x pretranslocation bias for $EC9^{GU}$ and a 50x posttranslocation bias for $EC9^{UG}$ (Figures 2-14B and C). We draw no conclusion about the actual translocation equilibria in $EC9^{UG}$ and $EC9^{GU}$ from this prediction, given the arbitrary constraints introduced to simplify modeling and the lack of comprehensive kinetic data. Rather, we conclude that a mechanism involving random order of PPi binding and translocation can, in principle, generate rates of pyrophosphorolysis that match those we observed.

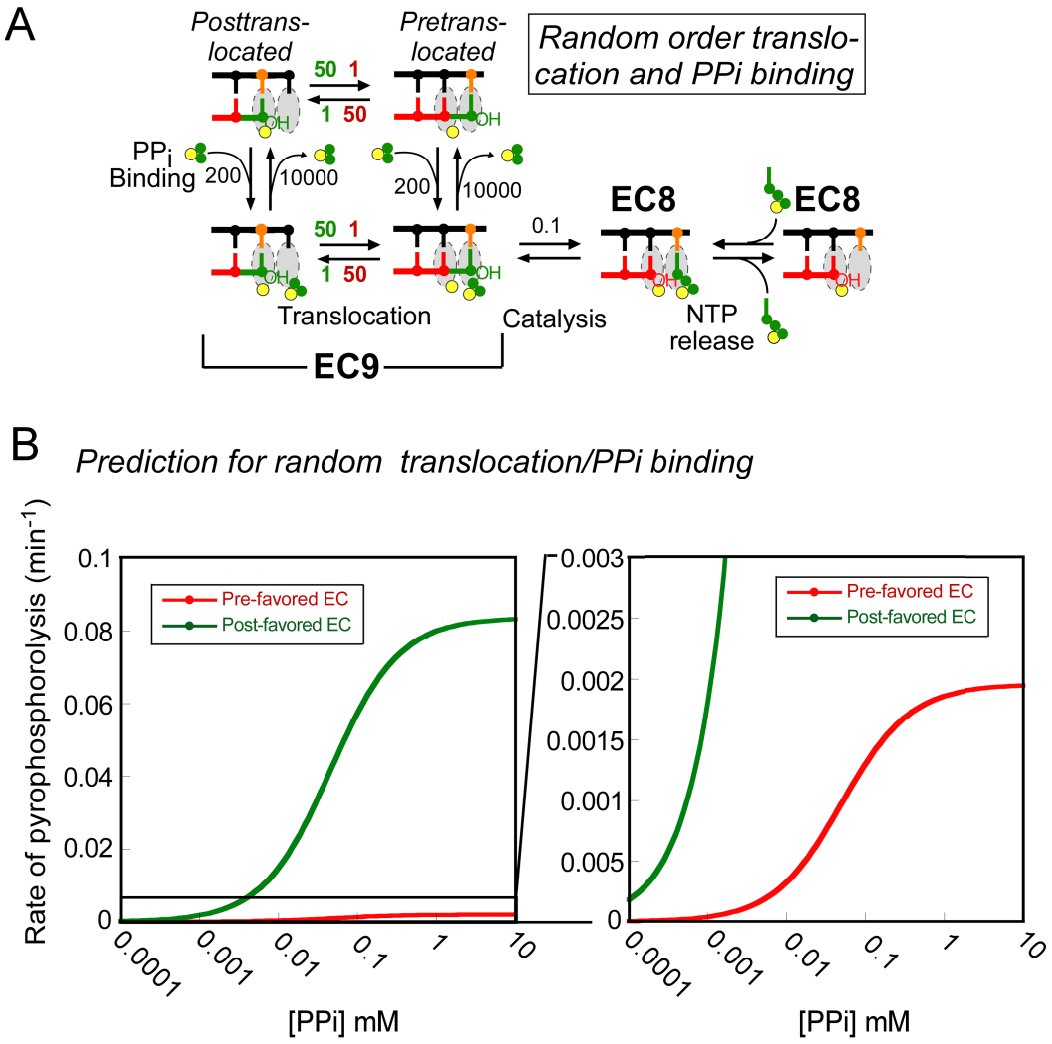


Figure 2-14. A random-order of translocation of PPi binding/release can explain differences in pyrophosphorolysis of EC9^{UG} and EC9^{GU}.

(A) A random-order translocation and PPi binding/release mechanism. With the arbitrarily chosen rate constants shown in the figure, the mechanism can account for a Vmax difference with little effect on apparent K_{PPi}. Translocation rates were assigned based on translocation bias of a particular EC. Forward and backward translocation rate constants for pre-favored ECs are shown in green (50 and 1 s^{-1}, respectively), whereas those for post-favored ECs are in red (1 s^{-1}; forward rate constant and 50 s^{-1}; reverse rate constant).

Figure 2-14 legend (cont.)

(B) A kinetic simulation graph of pretranslocation-favoring $EC9^{GU}$ (green) and posttranslocation favoring $EC9^{UG}$ (red) ECs was generated using the reaction scheme shown in Figure 2-14A using the program KinTek Global Kinetic Explorer (see Materials and Methods).

Discussion

Our investigation of the mechanistic basis of translocation bias uncovered basic effects of RNA:DNA hybrid sequence that appear universal among multisubunit RNAPs. As such, they provide insight into how RNAP interacts with the nucleic acid scaffold in an EC that will be generally applicable to understanding the regulation of transcript elongation. Additionally, our findings uncovered preliminary evidence that highlights the need for experiments that establish the order of translocation and PPi release/binding in the nucleotide addition cycle.

RNA 3′ dinucleotide sequence strongly influences translocation bias

Our central findings are that the RNA 3′ dinucleotide sequence has a major effect on bias of ECs between the pre and posttranslocated registers, and that this effect is conserved from bacterial to mammalian RNAPs. This translocation bias was evident in both the initial rate of pyrophosphorolysis and the point at which product inhibition (build up of NTP) caused the EC^{9} and EC^{8} species to equilibrate. The rate-limiting step in pyrophosphorolysis appears to be folding of the trigger loop into the trigger helices, as proline substitutions that block folding decrease the rate of pyrophosphorolysis by a factor of ~200 (Vassylyev et al., 2007b). Thus, we also needed to consider the possibility that different 3′ dinucleotides could affect trigger loop folding and give rise to the differences in the rates pyrophosphorolysis rather than being attributable to translocation bias. This concern was heightened by our finding that the rates of pyrophosphorolysis differed between $EC9^{GU}$ and $EC9^{UG}$ even at saturating PPi (Figure 2-12).

Although we cannot exclude differences in k_{cat} as contributing factors in the observed differences in pyrophosphorolysis rate, three arguments favor a dominant contribution of translocation bias. First, the strong inhibition of pyrophosphorolysis observed in $EC9^{UG}$ was

dramatically lessened when the RNA:DNA hybrid was overextended to 11 bp, presumably because the longer hybrid favored the pretranslocated register through steric clash between the upstream end on the hybrid and the wall of the hybrid-binding cleft (Figure 2-11). Second, an independent measure of translocation bias, exoIII footprinting, also detected a pretranslocation bias of $EC9^{GU}$ *vs* $EC9^{UG}$ (Figure 2-7). Third, an effect of RNA sequence on pyrophosphorolysis rate was also observed for the 3′-penultimate position of the RNA (G > A > C > U). It is not obvious how the 3′-penultimate base would directly affect folding of the trigger loop during catalysis (unlike for the 3′ base, which contacts the trigger helices), whereas these effects are easily rationalized by differences in contacts to active-site side chains that could favor or disfavor positioning nucleotide in the *i* site. It is notable in this regard that the order of effects of the penultimate nucleotide is the reverse of the order of effects of the 3′-nucleotide on pyrophosphorolysis rate (U > C > A > G). This inverse relationship is consistent with a preference for binding of G > A > C > U in the *i* site and a preference for binding of U > C > A > G in the *i+1* site.

Although our results favor a dominant contribution of 3′-dinucleotide sequence to pyrophosphorolysis, we do not mean to suggest that hybrid length makes no contribution. Indeed, we observed a strong effect of increasing pretranslocation bias of lengthening the hybrid past 9 bp. Our results disfavor the idea that an 8-bp hybrid necessarily favors the pretranslocated register (Kashkina et al., 2006), but we note that a 8-bp hybrid may not be a state ordinarily observed in ECs. Crystal structures of both yeast and bacterial RNAP in a posttranslocated EC detect 9-bp hybrids (Kettenberger et al., 2004; Vassylyev et al., 2007a; Westover et al., 2004), whereas paused ECs thought to be in the pretranslocated register appear to contain a 10-bp hybrid (Kyzer et al., 2007). A pretranslocated yeast RNAPII EC resolved only 9 nt of nascent

RNA, leaving the hybrid length uncertain but consistent with 10-bp in the pretranslocated state (Gnatt, 2002).

We note that assays of translocation bias employed in our study are suboptimal in that they are all indirect. More robust assays based on cleavage of DNA from locally generated free radicals have been used to detect translocation states in T7 RNAP and in reverse transcriptase (Guo and Sousa, 2006; Marchand and Gotte, 2003), but have not yet been applied to multisubunit RNAPs. A direct assay of sequence length on translocation bias, for instance using fluorescence quenching (Kashkina et al., 2007; Liu and Martin, 2001), would also be desirable. In general, the paucity of good methods to assay translocation register/bias limits our understanding. Although pre- and posttranlocated registers are thought to equilibrate in halted ECs, we lack measurements of the rates of interconversion. Development of such methods for multisubunit RNAPs would significantly advance study of the regulation of transcript elongation.

PPi binding may not be tightly coupled to translocation register

The relationship between PPi release (or PPi binding during pyrophosphorolysis) and translocation is one of the least well understood aspects of the nucleotide addition cycle catalyzed by multisubunit RNAPs. No structure of a multisubunit RNAP with bound PPi has been reported. Our attempts to reconcile pretranslocation bias of RNA 3′-GU relative to 3′UG with the apparently greater K_{PPi} of 3′-GU for initial pyrophosphorolysis rate suggests that another possibility should be considered, specifically that PPi release (or binding) might occur in either a pre or posttranslocated EC. There is no obvious structural impediment to PPi binding to a posttranslocated EC, unlike the obvious steric clash that would result from NTP binding to a pretranslocated EC. Even a highly constrained version of a model that allows random order of

translocation *vs*. PPi binding/release appears able to explain a lack of effect of 3′-dinucleotide sequence on apparent K_{PPi} even while the sequence has a large effect on translocation bias. We suggest that this idea merits further investigation.

Translocation bias may contribute to transcriptional pausing

Transcriptional pausing is thought to arise initially by a structural rearrangement in the active site of a pretranslocated EC (Landick, 2006, 2009). Although the sequence and structure contributions to transcriptional pausing are complex and include structures in the exiting RNA, the sequence of the RNA:DNA hybrid, the bases in the active site, and the sequence of the downstream DNA duplex (Chan et al., 1997; Chan and Landick, 1993; Kireeva and Kashlev, 2009), the RNA 3′ nucleotide has among the strongest effects. Pausing is favored by an RNA 3′ U or C (Aivazashvili et al., 1981; Chan and Landick, 1993). Thus, our finding that an RNA 3′ U or C exhibits a pretranslocation bias relative to an RNA 3′ A or G suggests a possible mechanistic basis for at least one component of the contribution of the 3′ nucleotide to pause proclivity. By increasing the fraction of time that ECs spend in the pretranslocated register prior to NTP binding, an RNA 3′ U or C may increase the probability that the EC isomerizes to a paused state and conversely decrease the probability of NTP binding. An increased probability of pausing at templates positions that favor the pretranslocated register is distinct from the proposed existence of "pretranslocated pauses" (Bai et al., 2004). In the mechanism we propose, the pretranslocated register increases the probability of an isomerization to an off-line paused state in a branched mechanism such that a fraction of RNAPs pause. The proposed "pretranslocated pause" is an on-line state in which the energetic barrier to formation of the posttraslocated state is so high that all RNAPs are delayed at the pause position.

The effect of RNA 3′ dinucleotide sequence on translocation bias may synergize with a G at -10 that also could favor the pretranslocated state by inhibiting hybrid melting at the upstream end (Herbert et al., 2008; Kyzer et al., 2007). Thus, an important component of pausing may be a hybrid sequence that favors the pretranslocated state, thereby allowing isomerization into the paused state, and one way that regulators may influence pausing could be to affect translocation bias.

Materials and Methods

Materials

All DNA and RNA oligonucleotides (Table 2-3) were obtained from IDT (Corvalville, IA) and purified by denaturing PAGE before use. [γ-^{32}P]ATP was from PerkinElmer Life Sciences and NTPs were from GE Healthcare (Piscataway, NJ). Exonuclease III (100,000 U/mL), T4 polynucleotide kinase, and apyrase (50 units/mL) were obtained from New England Biolabs (Ipswich, MA).

Proteins

Core *E. coli* and *T. thermophilus* RNAPs were purified as described previously (Toulokhonov et al., 2007; Vassylyeva et al., 2002). Calf thymus RNAP was purified as described previously (Hu et al., 2006). Yeast (*Saccharomyces cerevisiae*) RNA polymerase II was a generous gift from Dr. Yuichiro Takagi (Indiana University, Indianapolis).

In Vitro EC Reconstitution

Nucleic acid scaffolds for reconstituting ECs were assembled in reconstitution buffer (RB; 10 mM Tris·HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$) by heating 5′-^{32}P-labeled RNAs (500 nM; Table 2-3), tDNAs (1 μM; Table 2-3), and ntDNAs (1 μM; Table 2-3) to 95 °C for 2 min, rapidly cooling to 45 °C, and then cooling to room temperature in 2 °C/2 min steps as described previously (Kyzer et al., 2007). Reconstitution of ECs was performed by incubating core RNAPs with the nucleic acid scaffold (2:1 RNAP:nucleic acid scaffold) in elongation buffer (EB; 25 mM HEPES-KOH, pH 8.0, 130 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.15 mM EDTA, 5% glycerol, and 25 μg of acetylated bovine serum albumin/ml) for 15 min at 37 °C as described previously (Kyzer et al., 2007). When fully complementary DNA strands were used, RNA and

tDNA were preannealed and incubated with RNAP at 37 °C followed by incubation with ntDNA for an additional 10 min at 37 °C.

Pyrophosphorolysis

Pyrophosphorolysis was performed by incubating reconstituted ECs (~50 nM) with different concentrations of PPi in EB at 37 °C for *Eco*RNAP, 60 °C for *Thermus thermophilus* RNAP, and 30 °C for *Saccharomyces cerevisiae* RNAPII and *Bos taurus* RNAPII. Aliquots were removed at the indicated times and quenched with an equal volume of 2X stop buffer (8 M urea, 50 mM EDTA, 90 mM Tris-borate buffer, pH 8.3, 0.02% bromphenol blue, and 0.02% xylene cyanol) and analyzed by denaturing (8 M urea) 25% polyacrylamide (19:1) gel electrophoresis (45 mM Tris, pH 8.3, 1.25 mM EDTA, 45 mM Boric acid). To remove the NTP accumulated from pyrophosphorolysis, 1 μL of apyrase (50 milliunits) was added to 100 μL of reaction mixture containing 50 nM EC.

Exonuclease III Footprinting

Exonuclease III digestion was performed as described previously (Toulokhonov et al., 2007). ECs were assembled with *Tth*RNAP by annealing appropriate oligonucleotides as described above. Either the template or nontemplate DNA strand was 5′-end labeled with ^{32}P. The non-labeled DNA strand in the scaffold contained phosphothioate linkage at the 3′ penultimate position to inhibit cleavage of that strand by ExoIII. ExoIII digestion was initiated by addition of 100 U of ExoIII to ECs (44 μL) at 37 °C. Aliquots were removed at the indicated times and stopped with the addition of an equal volume of 2X stop buffer. Products of the reaction and G+A nucleotide-sequencing reaction (Maxam and Gilbert, 1980) of the same DNA fragment were analyzed by denaturing 15% polyacrylamide gel electrophoresis.

Intrinsic transcript cleavage assay

ECs were reconstituted in 20 mM Tris·HCl, pH 9.0, 20 mM NaCl, 0.1mM EDTA by incubating *Tth*RNAP with nucleic acid scaffolds at 37 °C for 15 min as described above. The cleavage reaction was initiated by the addition of $MgCl_2$ to 20 mM and incubated at 37 °C. At the indicated times, aliquots were removed, and samples were processed as for the pyrophosphorolysis assay.

Data Quantitation and Analysis

Gels were exposed to phosphorimager screens, scanned using a Typhoon PhosphorImager, and quantitated using the ImageQuant Software (GE Healthcare). The 9 nt RNA present in each lane was quantitated as a fraction of the total RNA in each lane (Figure 2-3) and corrected for the fraction remaining in the chase lane. The rate of pyrophosphorolysis was then determined as described below.

Estimation of Rates and Kinetic Modeling

To calculate the apparent rate of pyrophosphorolysis of 16 different variants in table 2-1, we determined rapid decay of 9 nt RNA by fitting the fraction of 9 nt (Figure 2-3) using the equation (3). A simple reversible mechanism of pyrophosphorolysis was used in the fitting for simplicity.

$$EC9 \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} EC8 \qquad (1)$$

The overall rate of pyrophosphorolysis (equation 1) is given by

$$\text{Rate} = k_1[EC9] - k_{-1}[EC8] \qquad (2)$$

Integration between t = 0 and t = t at concentrations $[EC9]_o$ and [EC9] gives

$$[EC9] = \frac{k_1[EC9]_o e^{-(k_1+k_{-1})t} + k_{-1}[EC9]_o}{k_1 + k_{-1}} \qquad (3)$$

The values of k_1 and k_{-1} for 16 ECs are summarized in Table 2-1 and the rate constant (k_1) of 16 ECs was used to generate a plot in Figure 2-3C.

To characterize the [PPi]-dependence of pyrophosphorolysis in ECs with 3′ U or 3′ G transcripts, we assayed pyrophosphorolysis of *Eco*EC9^{GU} and *Eco*EC9^{UG} at different [PPi]. The fraction remaining EC9 was plotted as a function of time (Figure 2-13). Since we included apyrase in these reactions, we modified reaction 1 and included the rate of apyrase action (reaction 4).

$$\mathrm{EC9} \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} \mathrm{EC8.NTP} \underset{k_{-2}}{\overset{k_2}{\longleftrightarrow}} \mathrm{EC8 + NTP} \overset{\mathrm{Apy}}{\underset{k_{-3}}{\overset{k_3}{\longleftrightarrow}}} \mathrm{Apy.NTP + EC8} \xrightarrow{k_4} \mathrm{Apy + NDP + Pi + EC8} \qquad (4)$$

We then estimated initial rates of pyrophosphorolysis by fitting the [EC9] versus time at different [PPi] to mechanism 4 using the program KinTek Global Kinetic Explorer (Johnson et al., 2009a, b), with diffusion limited NTP binding (k_{-2} = 6000 μM^{-1} min^{-1}) a dissociation rate to give a K_{NTP} of ~55 μM (k_2 =330,000 min^{-1}) (Foster et al., 2001). At the resulting NTP concentrations, reported kinetic values for apyrase were consistent with NTP degradation becoming rate limiting for overall pyrophosphorolysis (diffusion limited k_3 = 6000 μM^{-1} min^{-1}; k_{-3} = 132,000 μM^{-1} min^{-1}; k_4 = 190 min^{-1} ; (Molnar and Lorand, 1961). The concentration of apyrase, which varied somewhat due to the error in small-volume pipetting, was allowed to float. The initial rate (k_1) of pyrophosphorolysis of EC9^{GU} was obtained using these simulations and plotted versus [PPi] (Figure 2-12B). Because it was too slow for the apyrase reaction to become rate limiting, the pyrophosphorolysis rate of EC9^{UG} was determined by nonlinear regression assuming a pseudo-first order reaction and plotted versus [PPi] (Figure 2-12B).

To obtain the kinetic simulation graph depicted in Figure 2-14B, the arbitrary kinetic parameters shown in Figure 2-14A were used. The synthetic rate of pyrophosphorolysis for pre-favored EC and that of post-favored EC over a wide range of PPi concentrations were generated using KinTek Global Kinetic Explorer. The predicted values were used to construct complete curves by nonlinear regression to a hyperbola (Figures. 2-14B).

TABLE 2-3. Oligonucleotides used in this study

Stock number	Description	Sequence (5′ -> 3′)
DNA		
5847	Template DNA	CCTGTCTGAATCGATATCGCCGC
5926	Template DNA	CCTGTCTGAATCGATTATGCCGC
6056	Template DNA	CCTGTCTGAATCGA**TT**CCGCCGC
6057	Template DNA	CCTGTCTGAATCGA**GT**CCGCCGC
6058	Template DNA	CCTGTCTGAATCGA**CT**CCGCCGC
5927	Template DNA	CCTGTCTGAATCGAA**T**CCGCCGC
6059	Template DNA	CCTGTCTGAATCGAA**G**CCGCCGC
6060	Template DNA	CCTGTCTGAATCGA**TG**CCGCCGC
6061	Template DNA	CCTGTCTGAATCGA**GG**CCGCCGC
6062	Template DNA	CCTGTCTGAATCGA**CG**CCGCCGC
6063	Template DNA	CCTGTCTGAATCGAA**C**CCGCCGC
6064	Template DNA	CCTGTCTGAATCGA**TC**CCGCCGC
6065	Template DNA	CCTGTCTGAATCGA**CC**CCGCCGC
6066	Template DNA	CCTGTCTGAATCGA**GC**CCGCCGC
6067	Template DNA	CCTGTCTGAATCGAA**A**CCGCCGC
6068	Template DNA	CCTGTCTGAATCGA**TA**CCGCCGC
6069	Template DNA	CCTGTCTGAATCGA**GA**CCGCCGC
6070	Template DNA	CCTGTCTGAATCGA**CA**CCGCCGC
5848	Nontemplate DNA	TCGATTCAGACAGG
6355	Template DNA (Complete)	CCTGTCTGAATCGAA**C**CCGCCGCTTAGGACGTACTGACC
6354	Non-template DNA (Complete)	GGTCAGTACGTCCTAACGACGCTCATCGATTCAGACAG*G
6357	Template DNA (Complete)	CCTGTCTGAATCGA**CA**CCGCCGCTTAGGACGTACTGACC
6356	Non-template DNA (Complete)	GGTCAGTACGTCCTAACGACGCTTGTCGATTCAGACAG*G
6435	Non-template DNA (Downstream Exo)	GGTCAGTACGTCCTAAGCGGCGGTGGGAAGAGATTCAGAGCCATCCCAATGGACACC
6436	Template DNA (Downstream Exo)	GGTGTCCATTGGGATGGCTCTGAATCTCTTCC**C**A**C**CGCCGCTTAGGACGTACTGAC*C
6437	Nontemplate DNA (Downstream Exo)	GGTCAGTACGTCCTAAGCGGCGGGTGGAAGAGATTCAGAGCCATCCCAATGGACACC
6438	Template DNA (Downstream Exo)	GGTGTCCATTGGGATGGCTCTGAATCTCTTCC**A**C**C**CCGCCGCTTAGGACGTACTGAC*C
RNA		
5855	RNA	GCGGC**GAU**
5928	RNA	GCGGC**AUA**

Stock number	Description	Sequence (5′ -> 3′)
6039	RNA	GCGGCGG**AA**
6040	RNA	GCGGCGG**AC**
6041	RNA	GCGGCGG**AG**
5929	RNA	GCGGCGG**AU**
6042	RNA	GCGGCGG**CU**
6043	RNA	GCGGCGG**CA**
6044	RNA	GCGGCGG**CC**
6045	RNA	GCGGCGG**CG**
6046	RNA	GCGGCGG**GU**
6047	RNA	GCGGCGG**GA**
6048	RNA	GCGGCGG**GG**
6049	RNA	GCGGCGG**GC**
6050	RNA	GCGGCGG**UU**
6051	RNA	GCGGCGG**UA**
6052	RNA	GCGGCGG**UC**
6053	RNA	GCGGCGG**UG**

* indicates position of phosphorothioate

Acknowledgements

We are grateful to Dr. Yuichiro Takagi for providing Yeast (*Saccharomyces cerevisiae*) RNA polymerase II. We thank the members of the Landick laboratory for critical reading and comments on the manuscript.

References

Aivazashvili, V.A., Bibilashvili, R., Vartikian, R.M., and Kutateladze, T.A. (1981). [Effect of the primary structure of RNA on the pulse character of RNA elongation in vitro by Escherichia coli RNA polymerase: a model]. Mol Biol (Mosk) *15*, 915-929.

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

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

Chan, C., Wang, D., and Landick, R. (1997). Spacing from the transcript 3' end determines whether a nascent RNA hairpin interacts with RNA polymerase to prolong pausing or triggers termination. J Mol Biol *268*, 54-68.

Chan, C.L., and Landick, R. (1993). Dissection of the *his* leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. J Mol Biol *233*, 25-42.

Depken, M., Galburt, E.A., and Grill, S.W. (2009). The origin of short transcriptional pauses. Biophys J *96*, 2189-2193.

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). An allosteric path to transcription termination. Mol Cell *28*, 991-1001.

Epshtein, V., Dutta, D., Wade, J., and Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. Nature *463*, 245-249.

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.

Galburt, E.A., Grill, S.W., Wiedmann, A., Lubkowska, L., Choy, J., Nogales, E., Kashlev, M., and Bustamante, C. (2007). Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. Nature *446*, 820-823.

Gnatt, A. (2002). Elongation by RNA polymerase II: structure-function relationship. Biochim Biophys Acta *1577*, 175-190.

Guo, Q., and Sousa, R. (2006). Translocation by T7 RNA polymerase: a sensitively poised Brownian ratchet. J Mol Biol *358*, 241-254.

Herbert, K.M., Greenleaf, W.J., and Block, S.M. (2008). Single-molecule studies of RNA polymerase: motoring along. Annu Rev Biochem *77*, 149-176.

Hu, X., Malik, S., Negroiu, C.C., Hubbard, K., Velalar, C.N., Hampton, B., Grosu, D., Catalano, J., Roeder, R.G., and Gnatt, A. (2006). A Mediator-responsive form of metazoan RNA polymerase II. Proc Natl Acad Sci U S A *103*, 9506-9511.

Johnson, K.A., Simpson, Z.B., and Blom, T. (2009a). FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. Anal Biochem *387*, 30-41.

Johnson, K.A., Simpson, Z.B., and Blom, T. (2009b). Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. Anal Biochem *387*, 20-29.

Kashkina, E., Anikin, M., Brueckner, F., Lehmann, E., Kochetkov, S.N., McAllister, W.T., Cramer, P., and Temiakov, D. (2007). Multisubunit RNA polymerases melt only a single DNA base pair downstream of the active site. J Biol Chem *282*, 21578-21582.

Kashkina, E., Anikin, M., Tahirov, T.H., Kochetkov, S.N., Vassylyev, D.G., and Temiakov, D. (2006). Elongation complexes of Thermus thermophilus RNA polymerase that possess distinct translocation conformations. Nucleic Acids Res *34*, 4036-4045.

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., Kashlev, M., and Burton, Z.F. (2010). Translocation by multi-subunit RNA polymerases. Biochim Biophys Acta.

Kireeva, M.L., and Kashlev, M. (2009). Mechanism of sequence-specific pausing of bacterial RNA polymerase. Proc Natl Acad Sci U S A *106*, 8900-8905.

Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purtov, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N., and Kashlev, M. (2008). Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol Cell *30*, 557-566.

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

Kyzer, S., Ha, K.S., Landick, R., and Palangat, M. (2007). Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. J Biol Chem *282*, 19020-19028.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.

Landick, R. (2009). Transcriptional pausing without backtracking. Proc Natl Acad Sci U S A *106*, 8797-8798.

Landick, R., and Yanofsky, C. (1987). Isolation and structural analysis of the Escherichia coli trp leader paused transcription complex. J Mol Biol *196*, 363-377.

Larson, M.H., Greenleaf, W.J., Landick, R., and Block, S.M. (2008). Applied force reveals mechanistic and energetic details of transcription termination. Cell *132*, 971-982.

Liu, C., and Martin, C.T. (2001). Fluorescence characterization of the transcription bubble in elongation complexes of T7 RNA polymerase. J Mol Biol *308*, 465-475.

Marchand, B., and Gotte, M. (2003). Site-specific footprinting reveals differences in the translocation status of HIV-1 reverse transcriptase. Implications for polymerase translocation and drug resistance. J Biol Chem *278*, 35362-35372.

Marchand, B., Tchesnokov, E.P., and Gotte, M. (2007). The pyrophosphate analogue foscarnet traps the pre-translocational state of HIV-1 reverse transcriptase in a Brownian ratchet model of polymerase translocation. J Biol Chem *282*, 3337-3346.

Maxam, A.M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol *65*, 499-560.

Mejia, Y.X., Mao, H., Forde, N.R., and Bustamante, C. (2008). Thermal probing of E. coli RNA polymerase off-pathway mechanisms. J Mol Biol *382*, 628-637.

Molnar, J., and Lorand, L. (1961). Studies on apyrases. Arch Biochem Biophys *93*, 353-363.

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

Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. Proc Natl Acad Sci USA *92*, 4596-4600.

Pasman, Z., and von Hippel, P.H. (2000). Regulation of rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex. Biochemistry *39*, 5573-5585.

Rozovskaya, T.A., Chenchik, A.A., and Beabealashvilli, R. (1982). Processive pyrophosphorolysis of RNA by Escherichia coli RNA polymerase. FEBS Lett *137*, 100-104.

Santangelo, T.J., and Roberts, J.W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. Mol Cell *14*, 117-126.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. Embo J *22*, 2234-2244.

Surratt, C.K., Milan, S.C., and Chamberlin, M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription. Proc Natl Acad Sci USA *88*, 7983-7987.

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.

Vassylyev, D., Vassylyeva, M., Perederina, A., Tahirov, T., and Artsimovitch, I. (2007a). Structural basis for transcription elongation by bacterial RNA polymerase. Nature *448*, 157-162.

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

Vassylyeva, M.N., Lee, J., Sekine, S.I., Laptenko, O., Kuramitsu, S., Shibata, T., Inoue, Y., Borukhov, S., Vassylyev, D.G., and Yokoyama, S. (2002). Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from Thermus thermophilus. Acta Crystallogr D Biol Crystallogr *58*, 1497-1500.

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.

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

Zhang, J., Palangat, M., and Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. Nat Struct Mol Biol *17*, 99-104.

Zhang, J.a.L., R. (2009). Substrate loading, nucleotide addition, and translocation by RNA Polymerase. In RNA Polymerase as Molecular Motors, H.a.S. Buc, T., ed. (Cambridge, UK, Royal Society of Chemistry), pp. 206-235.

Chapter Three

Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit-channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH

This chapter is adapted from a published paper:

"Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit-channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH"
Kellie E. Kolb*, Pyae P. Hein*, and Robert Landick. 2013. Journal of Biological Chemistry *in press* (*co-first authors)

Abstract

Transcript elongation by bacterial RNA polymerase (RNAP) is thought to be regulated at pause sites by open *versus* closed positions of the RNAP clamp domain, pause-suppressing regulators like NusG and RfaH that stabilize the closed-clamp RNAP conformation, and pause-enhancing regulators like NusA and exit-channel nascent RNA structures that stabilize the open-clamp RNAP conformation. However, the mutual effects of these protein and RNA regulators on RNAP conformation are incompletely understood. For example, it is unknown whether NusA directly interacts with exit-channel duplexes and whether formation of exit-channel duplexes and RfaH binding compete by favoring the open and closed RNAP conformations. We report new insights into these mechanisms using antisense oligonucleotide mimics of a pause RNA hairpin from the leader region of the *his* biosynthetic operon of enteric bacteria like *E. coli*. By systematically varying the structure and length of the oligonucleotide mimic, we determined that full pause stabilization requires an RNA:RNA duplex of at least 8-bp or a DNA:RNA duplex of at least 11-bp; RNA:RNA duplexes were more effective than DNA:RNA. NusA stimulation of pausing was optimal with 10-bp RNA:RNA duplexes and was aided by single-stranded RNA upstream of the duplex, but was significantly reduced with DNA:RNA duplexes. Our results favor direct NusA stabilization of exit-channel duplexes, which consequently affect RNAP clamp conformation. Effects of RfaH, which suppresses oligo-stabilization of pausing, were competitive with antisense oligo concentration, suggesting that RfaH and exit-channel duplexes compete *via* opposing effects on RNAP clamp conformation.

Introduction

Transcriptional pausing plays multiple roles in regulating RNA synthesis and gene expression in all domains of life (Landick, 2006). In bacteria, pauses help maintain coupling of transcription and translation, aid proper folding of nascent RNA, provide time for recruitment of regulators, and are the first step in both Rho-dependent and intrinsic (Rho-independent) termination of transcription (Artsimovitch and Landick, 2002; Gusarov and Nudler, 1999; Landick, 2006; Landick et al., 1985; Pan et al., 1999; Pan and Sosnick, 2006; Peters et al.; Proshkin et al., 2010; Wickiser et al., 2005).

Multiple lines of evidence suggest that bacterial RNA polymerase (RNAP) initially enters pause states by a sequence-induced structural isomerization that alters the active site and disrupts the nucleotide addition cycle. This initially formed pause state, termed the elemental pause, forms in competition with nucleotide addition; thus, some elongating RNAPs may bypass entry into the elemental paused state with a probability that varies among pauses (Kassavetis and Chamberlin, 1981; Neuman et al., 2003). At a subset of elemental pauses, additional rearrangements of the elongating transcription complex (EC) or interactions of regulators prolong the pause (Artsimovitch and Landick, 2000; Landick, 2006). These rearrangements include (*i*) reverse translocation (backtracking) of the RNA and DNA chains through RNAP, which removes the RNA 3´ OH from the active site; and (*ii*) formation of RNA secondary structures ("pause hairpins") in the RNA-exit channel of RNAP, which is thought to inhibit nucleotide addition indirectly by stabilizing a conformation of RNAP in which the clamp domain is opened (Figure 3-1).

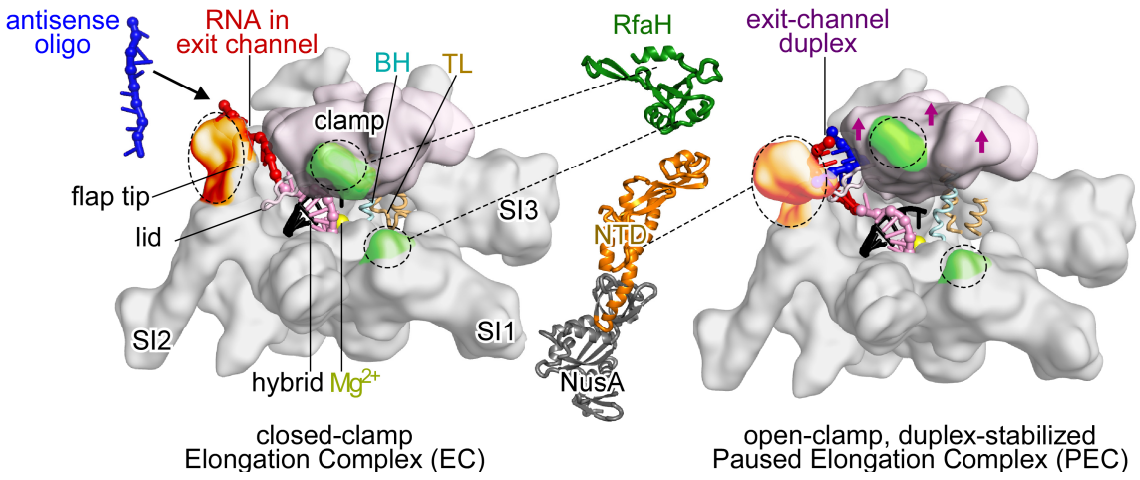


Figure 3-1. Models of *E. coli* RNAP EC and hairpin-stabilized paused EC.

Left, closed-clamp, active EC model (Opalka et al., 2010) showing clamp (light pink), bridge helix (BH), trigger loop (TL) and locations of major sequence insertions (SI1, SI2, SI3; Ref. (Artsimovitch et al., 2003), flap tip, lid, active-site Mg^{2+} (yellow) and binding sites for NusA-NTD (orange) and RfaH-NTD (green) (Belogurov et al., 2007; Ha et al., 2010; Sevostyanova et al., 2011). Most EC nucleic acids are omitted for clarity. The RNA:DNA hybrid (pink and black), exiting RNA (red), and 8-nt antisense RNA (blue) are shown. *Middle*, cartoon representations of RfaH-NTD (PDB 2oug; Ref. (Belogurov et al., 2007; Worbs et al., 2001) and NusA-NTD, S1, KH1, and KH2 (PDB 1hh2; (Belogurov et al., 2007; Worbs et al., 2001). Dotted lines indicate contacts to RNAP. *Right*, open-clamp paused EC (PDB 4gzy; (Weixlbaumer et al., 2013). Arrows indicate clamp movement and an 8-bp exit channel duplex (blue and red) is modeled into the expanded RNA exit channel.

A model pause signal from the *his* operon leader region uses a 5-bp stem, 8-nt loop pause hairpin that forms 12 nt from the RNA 3′ end to synchronize transcription of the *his* transcriptional attenuator with translation of the leader peptide coding region; the translating ribosome disrupts the hairpin to release the paused EC (Landick, 2006). Stabilization of the paused EC by the *his* pause hairpin appears to involve multiple inhibitory effects on steps in the nucleotide addition cycle in the RNAP active site (Toulokhonov et al., 2007). Non-paused nucleotide addition (50-100 s^{-1} for *E. coli* RNAP) requires four steps: (*i*) translocation of RNA and DNA through RNAP to position the RNA 3′ OH and template base in the so-called *i* and *i+1* subsites, (*ii*) NTP binding in the *i+1* subsite, (*iii*) catalysis, and (*iv*) pyrophosphate release (Zhang and Landick, 2009). Translocation shifts the DNA bubble downstream by shortening the RNA:DNA hybrid from 10- to 9-bp, shifting ~5-nt of ssRNA that lies in the RNA exit channel upstream, melting one bp of the downstream DNA duplex directly in front of the RNA 3′ end, and reannealing one bp of upstream DNA duplex (Kireeva et al., 2010; Malinen et al., 2012). Rapid catalysis requires folding of the trigger loop (TL) into an α-helical hairpin (trigger helices; TH) that contacts the NTP substrate and forms a three-helix bundle with the bridge helix (BH). The BH, which spans the active-site cleft, is deformable and may occlude the *i+1* subsite when NTP is absent and the TL is unfolded. The elemental pause appears to form when contacts between the clamp domain, the RNA:DNA hybrid and the downstream DNA duplex are loosened in a partially translocated EC, leading to a rearrangement in which a deformed BH prevents template base entry into the *i+1* subsite (Weixlbaumer et al., 2013). Subsequent formation of the pause RNA hairpin then appears to stabilize the paused EC in the pretranslocated register with the RNA 3′ nt frayed off the template base and the clamp domain opened, inhibiting both translocation and formation of the TH even after translocation and NTP

binding (Artsimovitch and Landick, 2000; Nayak et al., 2013; Zhang et al., 2010). Recent structural, single-molecule, and biochemical studies have provided evidence for the flexibility of the RNAP clamp domain (Chakraborty et al., 2012; Tagami et al., 2010), its movement associated with transcriptional pausing (Weixlbaumer et al., 2013), and the inhibition of TL folding in the hairpin-stabilized paused elongation complex (PEC; Nayak et al., 2013).

Regulators of pausing influence the RNAP active site through contacts to the pause hairpin, exit channel, or RNAP clamp domain (Figure 3-1). NusA increases the duration of hairpin-stabilized pausing through contacts to the hairpin and the tip of the flap domain, which forms the wall of the RNA exit channel opposite the clamp (Chan and Landick, 1993; Ha et al., 2010; Kyzer et al., 2007; Toulokhonov et al., 2001; Toulokhonov and Landick, 2003; Wang et al., 1997; Yang et al., 2009). An alternative model in which NusA acts indirectly to stabilize the hairpins by displacing ssRNA from a different NusA-binding site outside the RNA exit channel has also been proposed (Gusarov and Nudler, 2001). In contrast, the NusG paralog RfaH suppresses hairpin effects on pausing through contacts to RNAP thought to inhibit clamp opening (Belogurov et al., 2007; Sevostyanova et al.; Svetlov et al., 2007).

Both enhancement of pausing by the *his* pause hairpin and the hairpin contribution to NusA enhancement of pausing can be mimicked by formation of an 8-bp RNA duplex using an RNA oligo complementary to the exiting RNA (Ha et al., 2010). Although there is some evidence that changes in hairpin structure can influence pause or NusA effects (Toulokhonov et al., 2001), it remains unclear (*i*) if any duplex that favors clamp opening is sufficient for these effects, (*ii*) if specific features of the duplex are needed to enhance pausing, or (*iii*) if the essential features might differ among exit-channel duplexes for the basic hairpin effect, NusA enhancement of pausing, and RfaH suppression of pausing. To investigate these questions, we systematically

varied RNA exit-channel duplexes using complementary RNA and DNA oligos of varying lengths and concentrations and assayed pause prolongation, NusA enhancement, and RfaH suppression. Our results yield new insights into the mechanisms by which nascent hairpins affect pausing, support a model for direct NusA-hairpin interaction, and establish that duplexes and RfaH compete in a concentration-dependent manner for effect on EC.

Results

The length of RNAP exit-channel RNA:RNA duplexes affects PEC lifetime

An 8-bp RNA:RNA duplex formed by pairing an RNA oligo to the nascent RNA at the location of the *his* pause hairpin (transcript 3′-proximal end of duplex at -12) mimics the pause lifetime-increasing effect of the hairpin (Ha et al., 2010). To test whether the length of the RNA duplex influences this effect, we first asked if shortening the duplex to 7-bp without changing its location altered the duplex effect on pause lifetime. We performed this experiment using reconstituted ECs formed with *E. coli* RNAP on partially complementary nucleic-acid scaffolds known to recapitulate the properties of the *his* paused EC, as described previously (Ref. Kyzer et al., 2007; Figure 3-2 A and B; see Methods and Materials). C18 complexes were labeled with ^{32}P by incubating G17 ECs formed 2 nt upstream from the pause site with [α-^{32}P]CTP. After annealing an 8mer RNA oligo to the exiting RNA, the lifetimes of paused ECs formed at U19 were measured by tracking the ^{32}P-U19 RNA as a function of time after addition of GTP and UTP (to 10 and 100 μM, respectively) using denaturing polyacrylamide gel electrophoresis (Figures 3-2C and D). Consistent with previous findings, formation of the 8-bp exit channel duplex strongly stimulated pausing, and the magnitude of this effect (~10-fold on pause duration) was equivalent to the difference in pause duration between an EC lacking an exit channel duplex and one containing the natural *his* pause RNA hairpin (Figure 3-2F). However, shortening the exit-channel duplex to 7- bp decreased pause duration by a factor of ~2 even at saturating concentrations of complementary 7mer oligo (Figures. 3-2E and F; see below). (*note*: the nascent RNA used to test the 7mer antisense oligo is 1 nt shorter than the RNA used to test the 8mer oligo; thus, C17 and U18 ECs for the 7mer are equivalent to C18 and U19 ECs for the 8mer). Thus, an 8-bp but not a 7-bp RNA duplex is sufficient to recapitulate the full-hairpin effect.

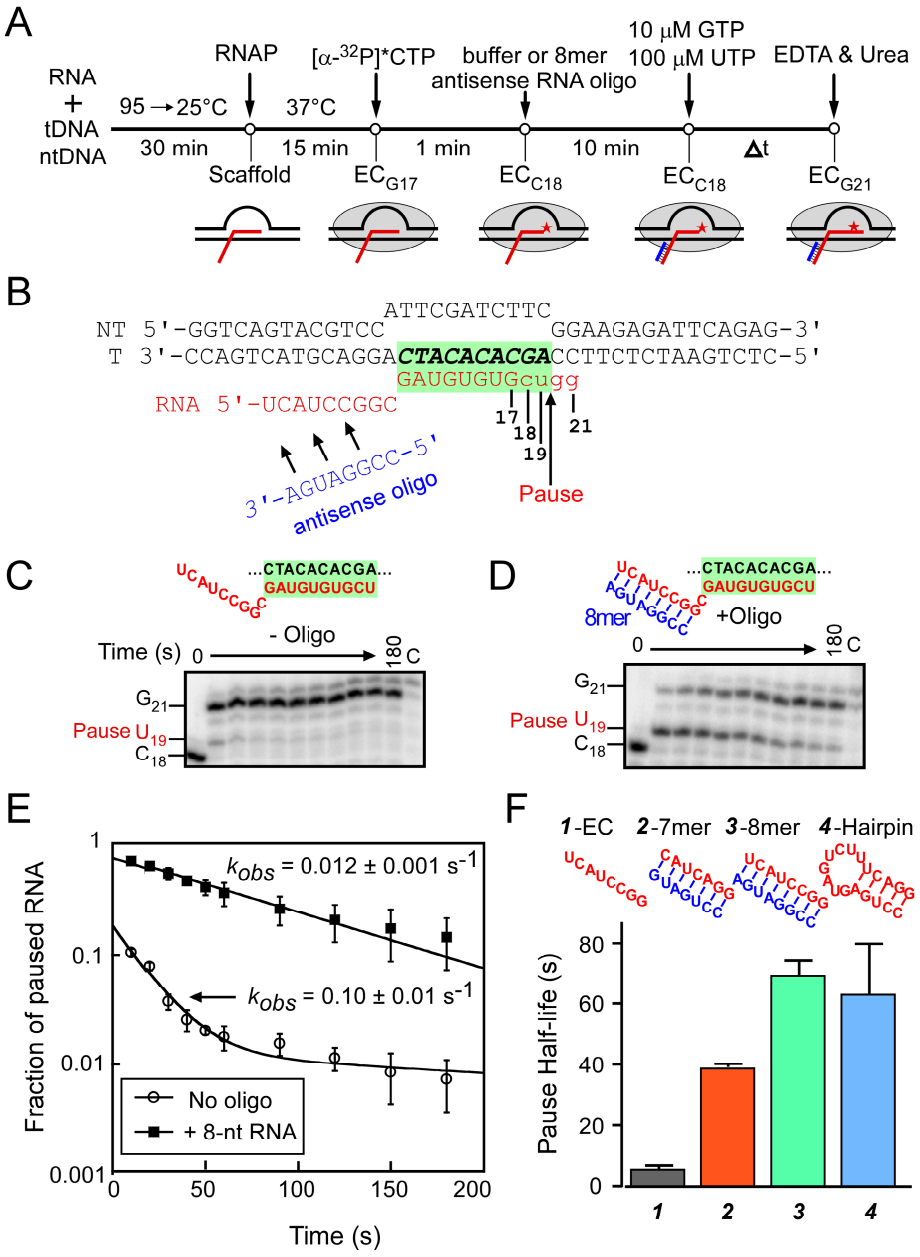


Figure 3-2. An antisense RNA oligo anneals to the nascent RNA and mimics the *his* pause hairpin.

Figure 3-2 legend.

(A) A schematic of the transcriptional pause assay. ECs were first preformed by assembling *E.coli* RNAP on the nucleic acid scaffold (RNA #6593, tDNA #5420, and ntDNA #5069; Table 3-4) shown in B (see Materials and Methods). Preformed ECs (G17) were elongated to the C18 position by addition of [α-^{32}P]CTP. The complexes were then incubated with or without an 8-nt antisense RNA oligonucleotide (oligo). The pause assay was initiated by addition of 10 μM GTP and 100 μM UTP; aliquots were removed at the indicated times and mixed with 2X loading dye, and RNA products were separated on 20% denaturing acrylamide gel.

(B) A representative example of a nucleic acid scaffold used in this study (for 8mer antisense oligo; yields EC G17). The RNA:DNA hybrid region is highlighted in green and the antisense RNA oligonucleotide (blue) anneals to C18 RNA upstream from the RNA:DNA hybrid (a 1-nt spacer separates the hybrid and the exit-channel duplex). Elongation of RNA17 by RNAP is shown in *lowercase type*.

(C,D) Two representative gels of transcription pause assay in the absence (C) or in the presence (D) of 1 μM 8-nt antisense RNA oligo. The corresponding nucleic acid scaffolds are depicted above the gels with the RNA:DNA hybrid highlighted in green. The lane marked "*C"* is a chase lane containing reaction products after ECs were incubated for 3 min with 0.5 mM GTP and UTP after the 3-min time course.

Figure 3-2 legend. (cont.)

(E) The 19mer RNA (U19) present in each lane was quantitated as a fraction of the total RNA in each lane, normalized for the fraction remaining in the chase lane and plotted as a function of reaction time. The rate of escape, $k_{obs,}$ is obtained by fitting the disappearance of U19 RNA to a single exponential (for 8-nt antisense RNA) or to a double exponential equation (for no oligo control). The rates of escape in the presence or absence of 8-nt RNA oligo are indicated on the plot.

(F) Pause half-lives of ECs containing no oligo (gray), 32 μM 7-nt RNA oligo (red), 32 μM 8-nt RNA oligo (green), or hairpin RNA (blue).

This result confirms our previous report (Ha et al., 2010) that the hairpin stem, but not the hairpin loop, generates the hairpin-stabilizing effect and suggests that the length of the stem determines the magnitude of the effect.

Although we allowed ample time for equilibration of antisense oligo binding to C17 ECs used to test the 7mer oligo before assaying pausing at position U18, we were concerned that oligos might bind the U18 ECs more tightly and that our results could be compromised by incomplete equilibration with the transient U18 ECs In principle, the posttranslocated C17 EC, which is thought to predominate at equilibrium (Malinen et al., 2012), should present the same nascent RNA target as the pretranslocated U18 ECs that would form upon nucleotide addition. Nonetheless, to test for possible complications, we compared effects of the 7mer oligo when bound to ECs halted at C17 or at U18 (Figure 3-3A). We found that 7mer oligo binding to C17 ECs actually had stronger effects on pausing, especially at saturating 32 μM concentrations (Figure 3-3B). We suspect that this result reflects backtracking of ECs halted at U18, which is known to occur (Kyzer et al., 2007) and which would shorten the RNA in the exit channel available for hybridization. These results validated our use of C17 complexes as preferred substrates for the oligo-simulated pause assay.

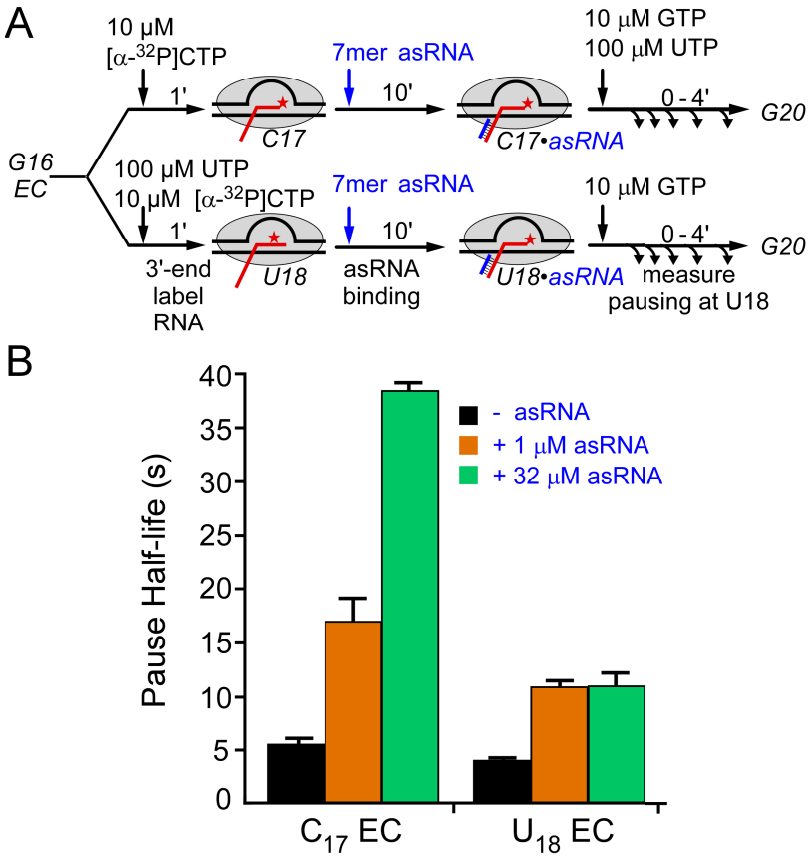


Figure 3-3. Comparison of antisense oligo annealing 1 nt before pause site or at pause site.

(A) Reaction scheme to test effects of EC position on oligo-stimulated pausing. ECs were halted at either C17 (1 nt before the pause) or U18 (at the pause) and then incubated with 1 μM or 32 μM antisense oligo (7mer asRNA) prior to assay of pausing. The nascent RNA is 1 nt shorter than that shown in Figure 3-2 to allow formation of a 7-bp exit channel duplex.

(B) Pause half lives for assays in which no oligo (black), 1 μM 7mer antisense oligo (orange), or 32 μM 7mer antisense oligo (green) was annealed to ECs halted at C17 or U18.

The maximal increase in PEC lifetime requires an exit-channel RNA:RNA duplex ≥8-bp

In principle, differences in either the strength of oligo binding to the nascent RNA in halted ECs or a requirement for a longer duplex for maximum effect on RNAP could explain the reduced effect of the 7-bp exit-channel RNA duplex on pause duration. Therefore, we next tested a range of possible exit-channel RNA duplexes from 5- to 22-bp using 1 μM oligo during the annealing step (Figure 3-4A). RNA oligos generating duplexes longer than 8-bp did not increase enhancement of pause duration significantly, giving ~14-fold longer pause durations than the scaffold lacking an exit channel RNA duplex (Figure 3-4B). Although 1 μM and 32 μM 8mer RNA oligos gave the same pause durations, showing that oligo binding saturated at 1 μM, much less stimulation of pause duration was detected at 1 μM 7mer relative to 32 μM (Figure. 3-4B). This result established that the 7mer oligo binds the G17 EC more weakly than the 8mer oligo. By testing various concentrations of 7mer oligo, we found that the ~7-fold stimulation of pause duration at 32 μM was saturating and that half-maximal stimulation occurred near 1μM 7mer (Figure 3-4C). Consistent with observations from Ha and colleagues that at least 7 complementary nucleotides (nt) are necessary to seed oligo binding (Cisse et al., 2012), we detected little if any increase in pause duration with 5mer and 6mer oligos at 1 μM (1.75-fold maximum; Figure 3-4B) and increasing 5mer or 6mer concentrations did not change this result (data not shown). These results establish (*i*) that an 8-bp or longer RNA duplex maximally stimulates pause duration; (*ii*) that a 7-bp RNA duplex is both more difficult to form and, once formed, stimulates pause duration less than an 8-bp duplex, and (*iii*) that exit-channel RNA duplexes less than 7-bp either are unable to form or have little to no effect on PEC lifetime.

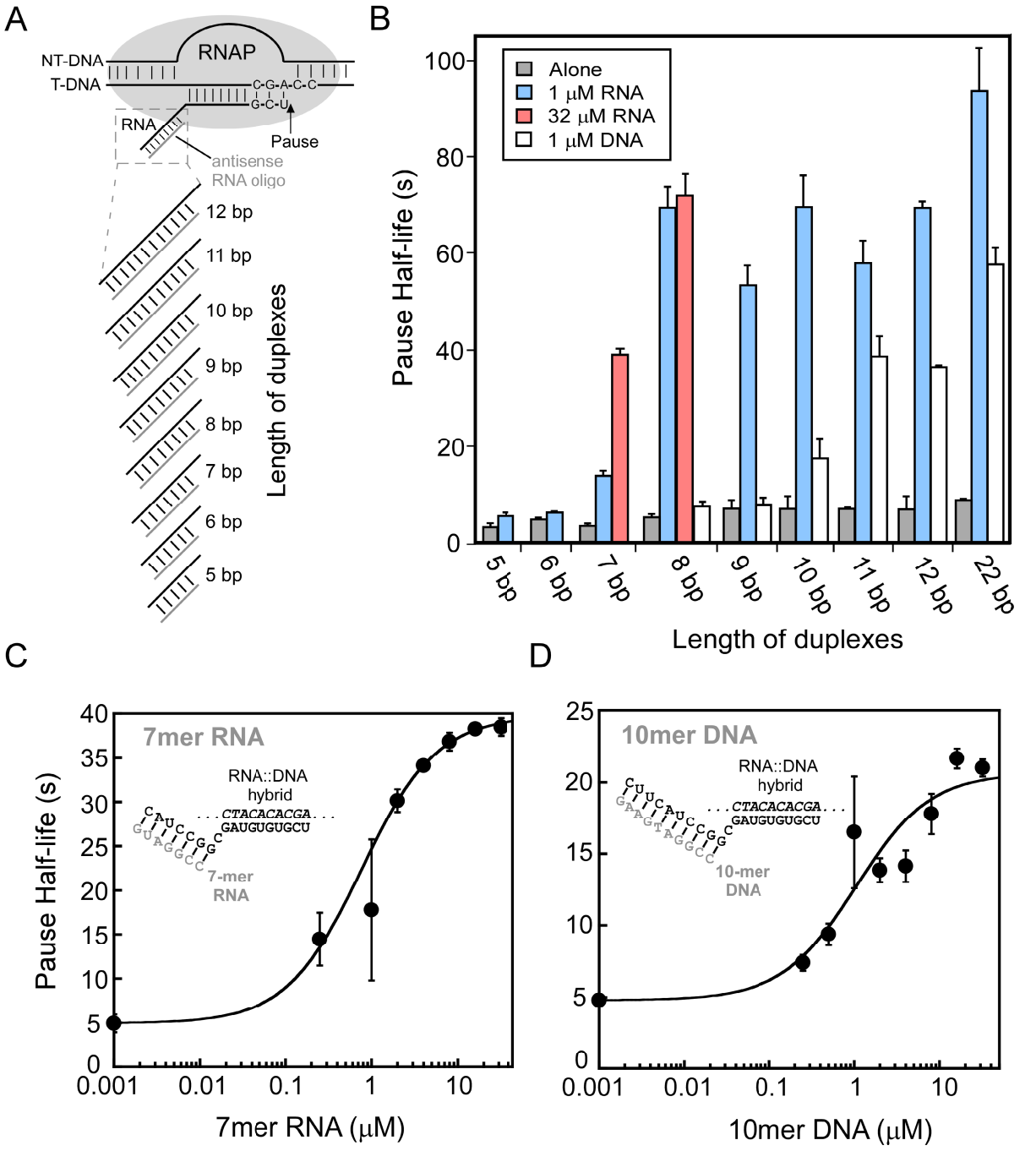


Figure 3-4. Duplexes of sufficient length in the RNA exit channel of RNAP increase pause duration.

Figure 3-4 legend.

(A) ECs with nascent RNA annealed to oligonucleotides of varying lengths (Table 3-4).

(B) Pause half-life of ECs incubated without oligos (alone) or with different lengths of 1 μM RNA oligos, or 1 μM DNA oligos. The maximum pause half-lives of 7-nt and 10-nt antisense RNA oligos at saturating concentration (32 μM) are shown in red.

(C) 7mer RNA oligo-concentration dependence of pause half-life.

(D) 10mer DNA oligo concentration-dependence of pause half-life.

Exit-channel DNA:RNA duplexes increase PEC lifetime less than RNA:RNA duplexes

We next tested whether DNA oligos targeting the exiting RNA at positions identical to the RNA oligos could also stimulate pause duration, and found DNA oligos exhibited significantly less effect on PEC lifetime than RNA duplexes (Figure 3-4B). Further, at 1 μM, DNA oligos ≥11 nt in length were required for maximal effect (in contrast to ≥8mer for RNA oligos). Further, the maximal effect of DNA oligos on pause duration was about half that observed for the comparable length RNA oligo (Figure. 3-4B). We conclude that exit-channel DNA:RNA duplexes increase PEC lifetime less than RNA:RNA duplexes. Since a 10mer DNA oligo at 1 μM gave about the same partial effect as a 7mer RNA oligo, we next tested whether its binding also was submaximal by measuring pause duration at varying 10mer DNA oligo concentrations. We found that pause duration saturated at ~4-fold stimulation (less than the ~7-fold seen for the 7mer RNA oligo) but that half-maximal binding again occurred at ~1 μM oligo (Figure 3-4D). These results establish (*i*) that DNA:RNA exit-channel duplexes are both more difficult to form than RNA:RNA duplexes (*e.g.* a 10mer DNA oligo gives only half-maximal effect at 1 μM, whereas 1 μM 8mer RNA is saturating); (*ii*) that once formed, DNA:RNA exit-channel duplexes stimulate pause duration significantly less than a RNA:RNA duplex of comparable length, and (*iii*) that DNA:RNA exit channel duplexes must be ≥11-bp to give the maximum possible effect.

The approximate equivalence of 10-11-bp DNA:RNA duplexes with 7-8 RNA:RNA duplexes for pause stimulation (even if not to similar levels) is consistent with calculations of DNA:RNA *vs*. RNA:RNA stabilities in our assay conditions (Table 3-1). Thus, these differences in length requirements likely reflect basic properties of the nucleic acid duplexes rather than differences in duplex-RNAP contacts. However, the reduced maximal stimulation of pause duration of all length DNA:RNA vs. RNA:RNA duplexes cannot be similarly explained and

must reflect differences in the way these two types of duplexes affect the enzyme (see Discussion).

TABLE 3-1. Predicted thermodynamics of RNA:RNA or RNA:DNA duplex formation

Sequence	Length	ΔG°_{37} (1M NaCl)[a]	ΔG°_{37} (100 mM NaCl)[b]	T_m[c]
RNA:RNA duplex	nt	kcal mol^{-1}		°C
r(CCUGA) /r(GGACU)	5	-6.14 ± 0.26	-5.53 ± 0.56	23.6
r(CCUGAC) /r(GGACUG)	6	-7.95 ± 0.26	-6.67 ± 0.68	35.9
r(CCUGACU) /r(GGACUGA)	7	-10.46 ± 0.27	-8.25 ± 0.84	45.3
r(CCGGAUGA) /r(GGCCUACU)	8	-12.70 ± 0.30	-9.67 ± 0.98	54.5
r(CCGGAUGAA) /r(GGCCUACUU)	9	-14.06 ± 0.29	-10.52 ± 1.1	55.8
r(CCGGAUGAAG) /r(GGCCUACUUC)	10	-15.71 ± 0.30	-11.56 ± 1.20	59.3
r(CCGGAUGAAGC) /r(GGCCUACUUCG)	11	-19.13 ± 0.31	-13.71 ± 1.40	66.2
r(CCGGAUGAAGCA) /r(GGCCUACUUCGU)	12	-21.67 ± 0.32	-15.31 ± 1.56	70.4
r(CCGGAUGAAGCUCUACAAAUGC) /r(GGCCUACUUCGAGAUGUUUACG)	22	-39.82 ± 0.38	-26.75 ± 2.73	80.2
RNA:DNA duplex				
d(CCGGATGA) /r(GGCCUACU)	8	-9.1 ± 0.52	-7.4 ± 0.75	35.5
d(CCGGATGAA) /r(GGCCUACUU)	9	-10.1 ± 0.58	-8.0 ± 0.82	35.6
d(CCGGATGAAG) /r(GGCCUACUUC)	10	-11.9 ± 0.68	-9.2 ± 0.93	37.8
d(CCGGATGAAGC) /r(GGCCUACUUCG)	11	-14.6 ± 0.83	-10.8 ± 1.11	45.7
d(CCGGATGAAGCA) /r(GGCCUACUUCGU)	12	-15.5 ± 0.88	-11.4 ± 1.17	49
d(CCGGATGAAGCTCTACAAATGC) /r(GGCCUACUUCGAGAUGUUUACG)	22	-31 ± 1.76	-21 ± 2.15	59.2

[a]ΔG°_{37} values were calculated by using nearest-neighbor parameters in 1 M NaCl for RNA:RNA duplex (Xia et al., 1998) and DNA:RNA duplex (Sugimoto et al., 1995).

[b]ΔG°_{37} values in 100 mM NaCl were determined from the equation given in Nakano *et al.* (Nakano et al., 1999).

[c]T_m values were calculated from using the Oligo Analyzer tool from IDT website (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer) at 130 mM NaCl and 5 mM $MgCl_2$ and total strand concentration of 1 μM. Errors are estimated to be ± 1.3 °C for RNA/RNA duplexes and ± 2.7 °C for DNA/RNA duplexes (Owczarzy et al., 2008; Owczarzy et al., 2004).

A maximal NusA effect on PEC lifetime requires RNA:RNA duplexes ≥10-bp

NusA stimulates the duration of some pauses by factors of 3 or more, and this effect of NusA depends on both the formation of an RNA structure in the exit channel and the NusA N-terminal domain that contacts RNAP at the flap tip, near the RNA exit point (Ha et al., 2010). Thus, we next asked how the effect of NusA varied as a function of different exit channel duplexes using 1 μM RNA oligos (Figure 3-5). NusA increased the duration of oligo-stabilized pausing, and this effect depended on the length and structure of the exit-channel duplex, but with patterns that differed from those observed for oligos in the absence of NusA (Figure 3-5A). The maximal effect of NusA (~10-fold stimulation of pause duration relative to oligo alone) required longer RNA duplexes (≥ 10-bp instead of ≥8-bp). As a control, we also tested the effect of NusA on PECs lacking an exit-channel duplex (single-stranded RNA, ssRNA, in exit channel; Figure 3-5C). We found no significant NusA effect on pause duration for PECs lacking exit-channel duplexes (Figure 3-5C), consistent with the idea that NusA acts through structured RNA in the RNAP exit channel to affect pausing. Consistent with previous results (Ha et al., 2010), NusA required only an extended duplex and not a ssRNA hairpin loop to stimulate pausing to extents even greater than observed for the native *his* pause hairpin (Figures 3-2D and F).

Given that the 7mer RNA oligo required significantly higher than 1 μM concentration of oligo to achieve maximal effect in the absence of NusA, we also tested whether NusA affected the oligo concentration dependence of pause stimulation by comparing effects of 1 μM and 32 μM oligo (Table 3-2). Although high oligo concentration gave the expected increase in pause duration in the absence of NusA, 1 μM 7mer RNA oligo proved to be a saturating concentration in the presence of NusA. This result demonstrates that NusA stabilizes RNA oligo binding to the exiting RNA. The reduced stimulation of pause lifetime by 7-bp relative to 8-bp exit-channel

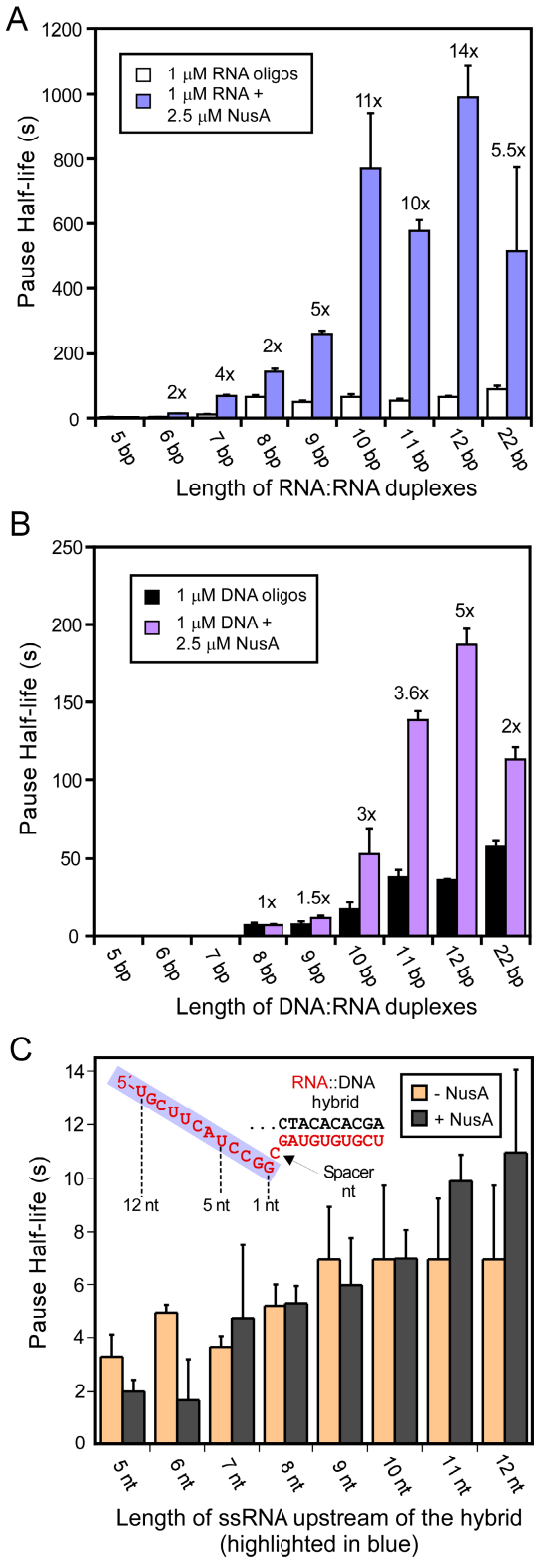


Figure 3-5. NusA increases pausing stimulated by oligos of varying length.

Figure 3-5 legend.

(A) NusA effect on RNA oligo stimulation of pause lifetime was measured using ECs reconstituted with DNAs #5069 and #5420 and different sizes of RNAs (Table 3-4; see Materials and Methods). Antisense RNA oligos were added at 1 μM to 50 nM ECs as illustrated in Figure 3-2A. For samples with NusA, 1 μM antisense oligo and 2.5 μM full-length NusA protein were added at the same time. The enhancement of pause half-life by NusA was calculated and the fold effect is indicated above the bars.

(B) Bar graph of pause half-life for ECs with 1 μM antisense DNA oligos and in the presence or absence of 2.5 μM full-length NusA. The experiment was performed identically to that shown in panel A except that DNA oligos with various lengths were added to generate a range of ECs containing exit-channel RNA:DNA heteroduplexes.

(C) NusA has no effect on the pause half-life of ECs when antisense RNA oligos are absent. ECs were assembled on nucleic acid scaffolds (shown in the inset) with different RNAs by changing the length of RNA highlighted in the blue box located upstream from the CMP spacer nucleotide. Pause half-lives of ECs with 2.5 μM full-length NusA (gray) or without NusA (orange) were determined as described in Materials and Methods.

TABLE 3-2. Pause enhancement by NusA for different types of duplexes and oligo concentrations

Types of duplex	[oligo]	Pause half-life - NusA	Pause half-life + NusA	+NusA/-NusA
	μM		s	
7-bp RNA:RNA duplex	1	14 ± 1	62 ± 5	4.4 ± 0.5
	32	38 ± 1	59 ± 5	5.2 ± 0.3
10-bp DNA:RNA duplex	1	15 ± 0.6	48 ± 2	3.2 ± 0.2
	32	26 ± 2	88 ± 7	3.4 ± .04
12-bp DNA:RNA duplex	1	35 ± 3	170 ± 15	4.8 ± 0.6
	32	36 ± 3	186 ± 17	5 ± 0.6

The pause half-life was calculated from the rate of pause escape on the scaffolds containing different exit channel duplexes using an *in vitro* transcriptional pause assay as described in Methods. 2.5 μM of full-length NusA was used in all experiments.

RNA duplexes at saturating concentrations in both the absence and presence of NusA suggests that contacts of the exit-channel duplex to both RNAP and to NusA can extend to positions >18 nt from the transcript 3′ end. Further, stabilization of exit-channel duplexes by NusA when no ssRNA is present upstream from the duplex argues strongly against the indirect NusA action model in which the duplex-stabilizing effect is proposed to result from NusA displacement of the upstream hairpin arm from a sequestering ssRNA binding site on RNAP (Gusarov and Nudler, 2001; see Discussion).

Exit-channel DNA:RNA duplexes reduce NusA stimulation of PEC lifetime

We next tested whether NusA would have comparable effects on pause duration for PECs containing exit DNA:RNA duplexes. Using the same set of exit-channel DNA:RNA duplexes that elicited maximal NusA effect with a ≥10-bp RNA:RNA duplex (Figures 3-4A and B), we found that NusA had no significant effect on pause duration for DNA:RNA <10-bp, and that the maximal NusA effect of 3.5- to 5-fold increase in pause duration occurred with DNA:RNA duplexes of 11-12-bp (in contrast to ≥10-fold stimulation for 11- and 12-bp RNA:RNA duplexes; Figure 3-5B). To test whether DNA oligos simply bind more weakly even in the presence of NusA, we compared the effect of 10mer and 12mer DNA oligos at 1 μM and 32 μM with and without NusA (Table 3-2). In contrast to the result with a 7mer RNA oligo, 1 μM 10mer DNA oligo proved to be subsaturating even in the presence of NusA, whereas a 12mer DNA oligo gave saturating effects at 1 μM in the absence or presence of NusA. Thus, not only are RNA:RNA duplexes more effective at stimulating pausing than DNA:RNA duplexes even at saturating oligo concentrations, but the formation of RNA:RNA duplexes also appears to be stabilized by NusA in ways that DNA:RNA duplexes are not.

Single- or double-stranded RNA upstream of the exit-channel duplex does not contribute to the direct effect of exit-channel duplexes on pause duration

When the *his* pause hairpin or other exit-channel structures form during natural transcription, the exit channel must accommodate three RNA strands: the two strands of the RNA duplex and the third RNA strand that connects the bottom of the hairpin stem near the lid domain to the upstream RNA transcript. This third RNA strand is thought to exit under the RNAP flap domain and is not required for effects on pause duration of the *his* pause hairpin or of artificial exit-channel duplexes (Ha et al., 2010; Kyzer et al., 2007). However, it remains unclear whether RNA upstream from the exit-channel duplex, in single- or double-stranded form, may positively or negatively affect formation of exit-channel duplexes, for instance by affecting clamp or flap positioning, since the *his* pause hairpin may form robustly and obscure such effects. To determine the impact of upstream RNA on duplex formation, we designed RNA oligos that could form 8-bp exit channel duplexes identical to those formed by our 8mer RNA oligo but that contained additional non-complementary nt upstream of the duplex (Figure 3-6). In one case, 8 non-complementary Us were present upstream (U-tail RNA) and in the other a short hairpin structure was present (hairpin-tail RNA; 7-bp stem, 4-nt loop). At high oligo concentration (32 μM), we found that the U-tail RNA and hairpin-tail RNA stimulated pause duration equivalently to the 8mer RNA oligo (Figure 3-6). Thus, neither the U-tail nor hairpin-tail altered the effect of the exit channel duplex per se. However, at 1μM the hairpin-tail RNA, but not the U-tail RNA, increased pausing much less effectively than the 8mer RNA previously found to be saturating at 1 μM (see above). Thus, the hairpin tail, but not single-stranded upstream RNA, appears to inhibit formation of the exit-channel duplex.

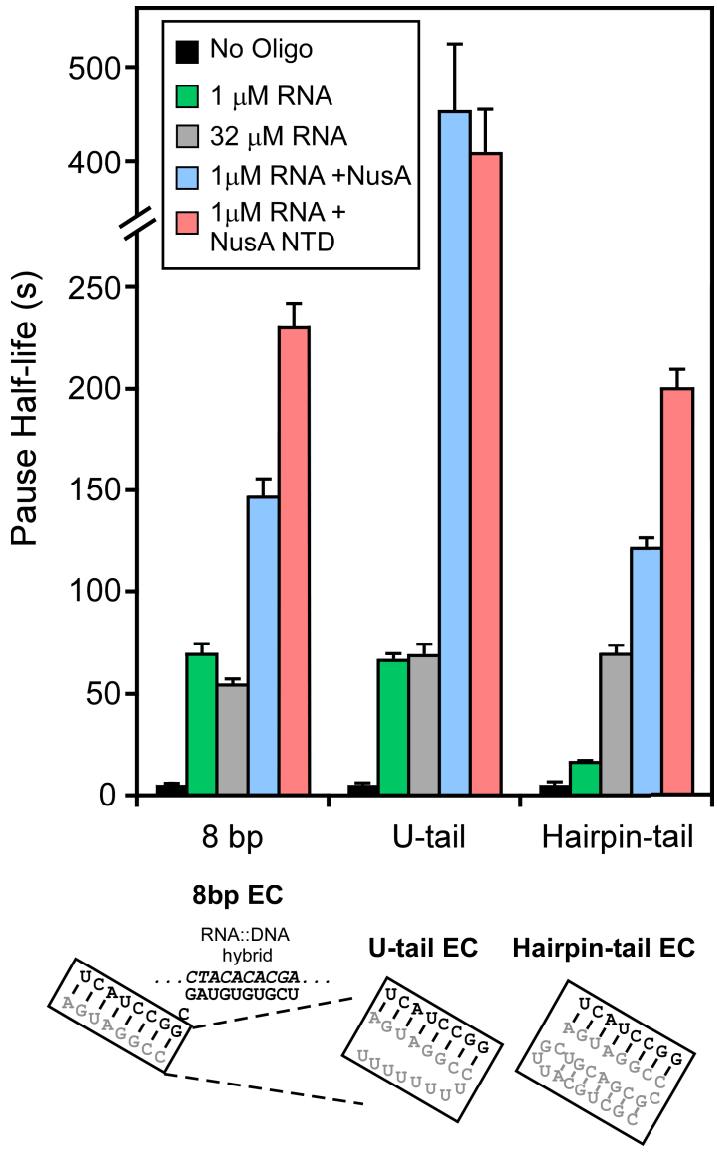


Figure 3-6. Full-length NusA favors but does not require exit-channel duplexes with upstream ssRNA for enhancement of PEC lifetime.

Figure 3-6 legend.

Pause stimulation by the antisense 8mer RNA oligo, RNA oligos with single-strand tails (U-tail EC), and RNA oligos with double-strand tails (Hairpin-tail EC). ECs were shown in Figure 3-2B (see Materials and Methods). Antisense 8mer RNA oligo, U-tail antisense RNA oligo, and hairpin-tail antisense RNA oligo were added to 1 μM or 32 μM to generate the 8-bp EC, U-tail EC, and hairpin-tail EC, respectively. Pause half-lives were measured as in preceding figures (see Materials and Methods) in the absence of NusA, with full-length NusA (2.5 μM) or NusA-NTD (residues 1-137; 4 μM).

ssRNA upstream from an exit-channel duplex enhances the effect on pause duration of full-length NusA but not of NusA-NTD

Ha *et al.* (Ha et al., 2010) found that full-length NusA required about 2-fold higher concentration to stimulate pausing on a his pause scaffold relative to on a promoter-initiated template for which RNA is present upstream of the pause hairpin, but that this difference was insignificant for the NusA-NTD alone, which binds PECs more weakly (20- to 40-fold higher apparent K_i). Thus, we wondered if ssRNA upstream from an exit-channel duplex could contribute to NusA stimulation of pause duration. To address this question, we tested NusA enhancement of pause duration using the U-tail and hairpin-tail RNAs oligos. Although the effects of 8mer RNAs with or without the U-tail and the NusA effect on an 8-bp exit channel duplex saturate at 1 μM oligo, pause duration increased for the U-tail RNA relative to the 8mer RNA in the presence of NusA (Figure 3-6). Consistent with findings of Ha *et al.* (Ha et al., 2010), some of this effect disappeared when using the NusA-NTD. Although upstream RNA is not needed for NusA action, single-stranded, upstream RNA can contribute to full-length NusA action. In contrast, PECs containing the hairpin-tail RNA reached about the same pause duration in the presence of NusA as observed for the 8mer oligo. Hence, duplex RNA upstream of the exit-channel duplex does not appear to aid NusA action. These results are consistent with an interaction of the NusA KH1 and KH2 domains with single-stranded RNA upstream of the exit-channel duplex that contributes to NusA enhancement of pausing, as proposed by Ha *et al.* (Ha et al., 2010).

RfaH pause suppression can be overcome by high concentrations of antisense RNA

RfaH, a NusG paralog found in some enterobacteria, was found previously to inhibit pausing at the *his* pause site (Artsimovitch and Landick, 2002; Belogurov et al., 2007). This effect of RfaH appears to be mediated by RfaH contacts to the RNAP clamp and lobe domains that prevent clamp opening required for exit-channel duplex formation (Figsure 3-1 and 3-7A; (Sevostyanova et al., 2011). If this model is correct, then binding of antisense oligos in the RNAP exit channel and binding of RfaH to RNAP may compete indirectly; bound RfaH would be expected to inhibit oligo binding by locking the clamp shut whereas bound oligo would be expected to inhibit RfaH binding by favoring an open clamp position that precludes simultaneous RfaH contacts to the clamp and lobe. Alternatively, RfaH may act through effects on the active site even when an exit-channel duplex is formed; in this latter case, high concentrations of antisense oligo should not outcompete pause suppression by RfaH.

To test whether RfaH competes with antisense oligo binding or acts on ECs even in the presence of an exit-channel duplex, we tested pause-enhancement by the 8mer antisense RNA and pause suppression by the isolated N-terminal domain of RfaH (RfaH-NTD). RfaH-NTD lacks the autoinhibitory C-terminal RfaH domain and is capable of RfaH action even in the absence of an *ops* sequence on the EC nontemplate strand (Ref. Belogurov et al., 2010; Belogurov et al., 2007; Methods and Materials). RfaH-NTD decreased pause dwell time at the *his* pause by a factor of ~5 at 1 μM antisense 8mer (pause half-life ~69 s without RfaH-NTD *versus* 14 s with 500 nM RfaH-NTD, and required only 125 nM RfaH-NTD for a large effect; Figures 3-7B and C; Table 3-3). Before testing for competition between RfaH and antisense RNA, we next asked if NusG might exhibit a similar effect. Although the NusG-CTD does not inhibit NusG binding to RNAP (Mooney et al., 2009) and full-length NusG has little effect on

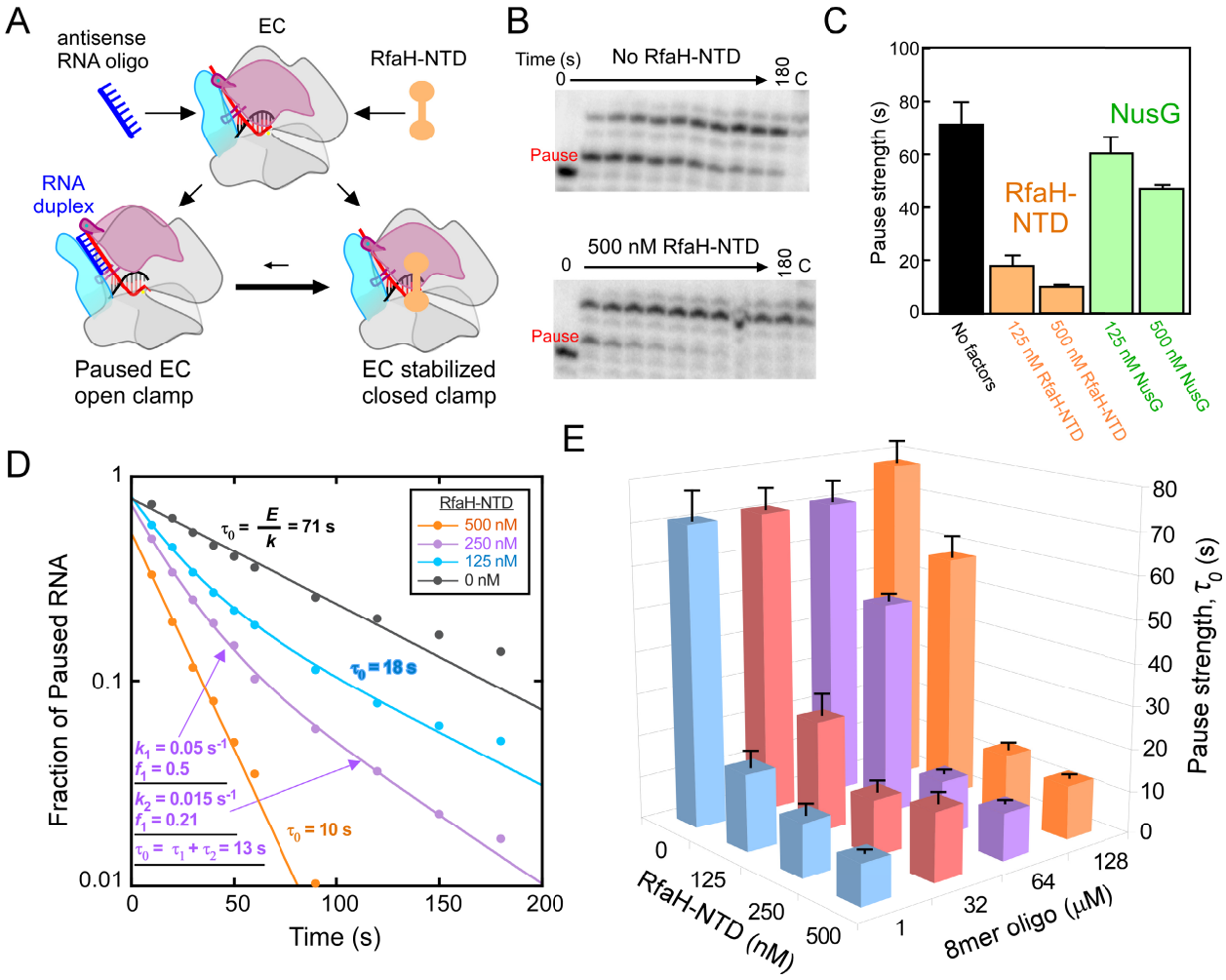


Figure 3-7. RfaH-NTD and the 8mer RNA oligo compete for binding to ECs.

(A) A model of competition between RfaH and oligo binding through the movement of the RNAP clamp domain (pink). RNA is red, antisense RNA oligo is blue, and RfaH is orange. Formation of RNA duplex requires opening of the clamp domain, which is inhibited when RfaH is bound to EC and holds the clamp in the closed position.

Figure 3-7 legend (cont.)

(B) Representative pause assay gel panels for ECs assembled on the nucleic acid shown in Figure 3-2B. ECs (C18; 50 nM) were elongated through the *his* pause site in the presence of 1 μM 8-mer antisense RNA oligo with or without 500 nM RfaH-NTD at 100 μM UTP and 10 μM GTP (see Materials and Methods).

(C) Pause strengths derived from the experiment shown in panel B and similar experiments with NusG (antisense RNA at 1 μM). Pause strengths were calculated as shown in panel D.

(D) Derivation of pause strengths. The fractions of pause RNA for the experiment shown in panel B and similar experiments with 125 nM and 250 nM RfaH-NTD were plotted as a function of reaction time. Apparent fraction paused (E) and pause escape rate (k) were determined by nonlinear regression. Pause strength (τ_0) was calculated from E and k ($\tau_0 = E/k$). Two pause rate components were evident at 125 and 250 nM RfaH-NTD, necessitating use of a double-exponential fit. In these cases, pause strengths were calculated from the two rate components and two fractions present ($\tau_0 = f_1/k_1 + f_2/k_2$; $\tau_0 = \tau_2 + \tau_2$) as illustrated for 250 nM RfaH-NTD in the figure (see Materials and Methods).

(E) A 3-D plot showing pause strengths as a function of concentrations of antisense 8mer and RfaH-NTD present in the pause assays.

TABLE 3-3. Antagonistic effects of RfaH and antisense 8mer RNA oligo on pausing

	1 μM 8mer		32 μM 8mer		64 μM 8mer		128 μM 8mer	
RfaH-NTD	half-life	fraction	half-life	fraction	half-life	fraction	half-life	fraction
(nM)	(s)		(s)		(s)		(s)	
500	14 ± 0.3	0.48 ± 0.01	17 ± 3	0.67 ± 0.07	14 ± 1	0.55 ± 0.01	17 ± 1	0.52 ± 0.01
250	14 ± 2 46 ± 9	0.51 ± 0.07 0.22 ± 0.09	14 ± 2 46 ± 9	0.53 ± 0.08 0.23 ± 0.09	16 ± 1	0.55 ± 0.01	20 ± 2	0.51 ± 0.01
125	14 ± 3 58 ± 9	0.47 ± 0.07 0.32 ± 0.09	16 ± 7 69 ± 14	0.35 ± 0.1 0.43 ± 0.1	44 ± 1	0.84 ±0.02	62 ± 6	0.64 ± 0.01
0	69 ± 5	0.71 ± 0.07	71 ± 6	0.69 ± 0.01	72 ± 6	0.68 ± 0.004	79 ± 7	0.68 ± 0.01
125 nM RfaH-NTD + 127 μM non-comp. oligo	14 ± 5 60 ± 15	0.3 ± 0.1 0.38 ± 0.1	- -	- -	- -	- -	- -	- -

Pause half-life and fraction paused species were calculated from pause assays using the 8-bp scaffold (Figure 3-2B), as described in Materials and Methods. Errors are standard deviations from experimental triplicates. -, not determined.

hairpin-stabilized pausing (Artsimovitch and Landick, 2000), we reasoned that high concentrations of NusG might compete for antisense RNA binding. Indeed, we found that 500 nM NusG partially suppressed pause stimulation by 1 μM antisense RNA (Figure 3-7C), but the weak effect and absence of any effect at lower concentration precluded using NusG for our planned experiment. We concluded that NusG likely binds the closed clamp EC more weakly than RfaH-NTD, and therefore used RfaH-NTD for subsequent experiments.

To test whether the 8mer antisense RNA and RfaH-NTD compete as a function of their relative concentrations, we conducted the oligo-stimulated pause assay at various RfaH-NTD and 8mer concentrations. In some conditions, the rate of pause RNA disappearance was biphasic (*e.g.*, at 250 nM RfaH-NTD and 1 μM 8mer; Figure 3-7D; Table 3-3). When present, the faster pause escape rate was relatively constant with a $t_{1/2}$ of ~14 s (Table 3-3). In contrast, the $t_{1/2}$ of the slow species, when evident, was inversely related to the concentration of RfaH-NTD (Table 3-3). To simplify analysis given the presence of the two pause species, we calculated pause strengths (τ_0; fraction paused / escape rate; Figure 7D; Materials and Methods) because τ_0 of multiple pause species at one position can be added. We found that high concentrations of RfaH-NTD could outcompete the effect of even high concentrations of 8mer antisense oligo. However, at intermediate concentrations of RfaH (125 nM RfaH-NTD), high concentrations of 8mer could overcome pause suppression by RfaH-NTD even though 1 μM 8mer was sufficient to give the full effect of the 8mer in the absence of RfaH-NTD. To verify that this effect was specific to formation of an exit-channel duplex, we also tested high concentrations of a noncomplementary 8mer RNA in the presence of 1 μM antisense 8mer and 125 nM RfaH-NTD (Table 3-3). We found that high concentrations of noncomplementary 8mer did not compete for pause suppression by RfaH, verifying that competition occurs through formation of an exit-channel

duplex. In our experiments, we added RfaH-NTD and the oligo at the same time. To test whether the order of addition is important, we added RfaH-NTD followed by the addition of oligo or vice versa. We found that the competition between RfaH and the oligo was unaffected by the order of RfaH and oligo addition. We conclude that RfaH-NTD and 8mer antisense oligo compete in a concentration-dependent manner consistent with the model that RfaH stabilizes the closed clamp, which inhibits 8mer binding, whereas exit-channel duplexes stabilize the open clamp, which inhibits RfaH binding. Our results suggest *E. coli* NusG acts similarly, but with reduced ability to favor clamp closing (see Discussion).

Discussion

Our study provides several new insights into the mechanism of pause enhancement by exit-channel duplexes and the mechanisms of the elongation factors NusA and RfaH. We found that the length and type of the exit-channel duplexes affect duplex-stabilized pausing and NusA enhancement of pausing differently. Beyond simply being harder to form, DNA:RNA exit channel duplexes are less effective at stabilizing the paused state even in the absence of NusA. The maximal NusA stimulation of pausing requires duplexes 2-bp longer than the maximal exit-channel duplex effect without NusA (≥10-bp vs. ≥8-bp RNA:RNA duplexes, respectively). Our results suggest that NusA functions through direct NusA-duplex interaction, rather than indirect displacement of ssRNA-binding to RNAP, and that RfaH and exit-channel duplexes compete for opposite effects on RNAP clamp conformation.

The bacterial RNAP exit channel appears to recognize RNA:RNA duplexes preferentially

Although many studies provide evidence for direct RNA hairpin-RNAP interaction in the stimulation of transcriptional pausing by *E. coli* RNAP (Artsimovitch and Landick, 1998; Chan and Landick, 1993, 1997; Toulokhonov et al., 2001; Toulokhonov and Landick, 2003; Wang et al., 1997), little is known about the structural requirements for these effects. Chan *et al.,* 1993 (Chan and Landick, 1993) found that a UUCG tetraloop could replace the 8-nt loop normally present on the 5-bp stem of the *his* pause hairpin without loss of pause stabilization, but that the tetraloop *his* pause hairpin reduced the effect of NusA. Toulokhonov *et al.,* 2001 (Toulokhonov et al., 2001) found that increasing the length of the *his* pause hairpin stem to 8-bp gave a 50% increase in pause duration, about the same effect as increasing the distance of the hairpin from the RNA 3′ end from 11 nt to 12 nt. Our results here suggest that the 8-bp RNA:RNA stem is

sufficient for the maximal effect of an exit-channel duplex, whereas a 7-bp stem is less effective and a ≤6-bp stem has little effect. The requirement for ≥7-bp for pause stimulation is consistent with the recent report of a "rule-of-seven" for rapid pairing of oligonucleotides (Cisse et al., 2012); for both RNA and DNA, and both *in vivo* and *in vitro* seven contiguous bp greatly increase k_{on} for pairing. We infer that the same rule applies to pairing of nascent RNA in the RNAP exit channel to oligos in solution.

Interestingly, the concentration dependencies of pause stimulation by 7mer antisense RNA oligo and 10mer antisense DNA oligo (Figures 3-4C and D) are not much different than those predicted for pairing of the pure nucleic acids at 37 °C and the buffer conditions used in our assay (50% of 50 nM transcript predicted to be paired by ~2 μM 7mer RNA or ~0.7 μM 10mer DNA based on data in Table 3-1). Thus, the RNAP exit channel does not appear to enhance or inhibit oligo binding dramatically, although direct measures of oligo binding will be required to verify this conclusion.

However, the natural 5-bp stem of the *his* pause hairpin is more effective than our 7-bp artificial duplex. This result suggests that the hairpin loop, which is likely to assume some degree of secondary structure, contributes to the effect of the *his* pause hairpin even though a single-stranded loop is not required for the effect of artificial duplexes. Thus, the natural *his* pause hairpin may resemble the artificial 8-bp exit-channel duplex in overall shape, evident in a model of the *his* pause hairpin loop (Ha et al., 2010), and this duplex length apparently is required for the full-effect of the exit-channel duplex.

Interestingly, DNA:RNA duplexes of any size are unable to mimic the full effect of the *his* pause hairpin (Figure 3-4). Although DNA:RNA duplexes are expected to be less stable than RNA:RNA duplexes (Nakano et al., 1999; Owczarzy et al., 2008; Owczarzy et al., 2004),

DNA:RNA duplexes predicted to be as stable as somewhat shorter RNA:RNA duplexes (Refs. Sugimoto et al., 1995;Table 3-2; Xia et al., 1998) still fail to increase pause duration to the same extent as those RNA:RNA duplexes (Figure 3-4). We suggest that this reduced effect of DNA:RNA exit channel duplexes is likely to reflect some extent of specificity in the interactions of pause-enhancing duplexes with the RNAP exit channel. In crystal structures, a 10-bp DNA:RNA duplex is at least as large in both helix diameter and length as an 8-bp RNA:RNA duplex (Figure 3-8). Hence, based simply on steric consequences for the exit channel, the DNA:RNA duplex should be as effective as the RNA:RNA duplex in altering RNAP clamp conformation. The small differences evident are in the positions of phosphate charges relative to the minor groove. Although differences in interactions within the grooves are possible, we favor the view that differences in locations of the phosphates on the RNA:RNA duplex increase interactions with positive charges on RNAP that either slow translocation or alter interactions with charges on RNAP that result in a greater degree of clamp opening. Distinguishing these two possible effects remains an important research question. A PEC crystal structure that includes an exit-channel duplex would be highly informative, but all efforts to date to obtain such a structure have resulted either in insufficient resolution to detect the duplex or, more likely, degradation of the duplex during formation of the crystal lattice (Tagami et al., 2010; Weixlbaumer et al., 2013).

NusA remodels RNAP exit-channel interaction specificity

The *E.coli* hexadomain NusA regulator (NTD-S1-KH1-KH2-AR1-AR2) is known to interact with the RNAP exit channel *via* an essential NusA NTD-RNAP-flap tip contact (Ha et al., 2010; Toulokhonov et al., 2001; Yang et al., 2009), alone to promote hairpin-stabilized pausing and intrinsic termination (Farnham et al., 1982; Landick and Yanofsky, 1984), and in concert with

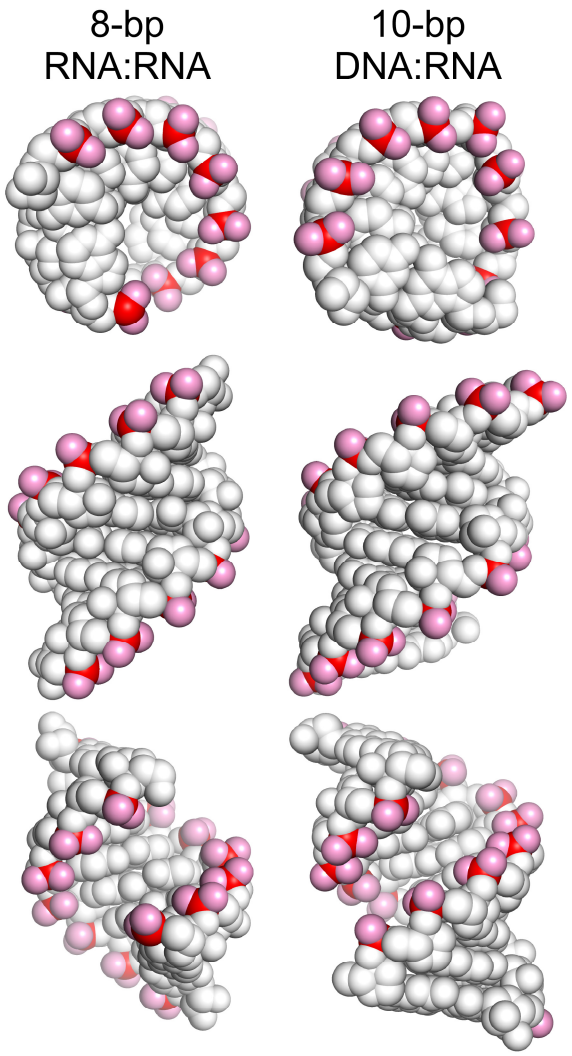


Figure 3-8. Structures of 8-bp RNA:RNA and 10-bp DNA:RNA duplexes.

Representations of duplex structures (PDB 1rna and 1fix; Refs. (Dock-Bregeon et al., 1989; Horton and Finzel, 1996) shown end on (*top*), from major groove side (*middle*), and from minor groove side (*bottom*) with phosphates in red and phosphate oxygens in pink.

antiterminator proteins like λN or λQ to suppress termination by inhibiting formation of exit-channel duplexes (Gusarov and Nudler, 2001; Shankar et al., 2007). Based on nascent RNA-NusA crosslinking, Gusarov and Nudler argue that alone NusA stimulates the exit-channel duplex formation indirectly by binding to and displacing nascent RNA from a ssRNA-specific binding site near the exit channel (Gusarov and Nudler, 2001). In a variant of this model, Yang and Lewis propose that the NusA AR1-2 interacts with the upstream arm of the exit-channel hairpin to displace it from the ssRNA-binding site (Yang and Lewis, 2010). Our results contradict the simple versions of either model. Pause enhancement by NusA requires the potential to form an exit-channel duplex, but remains strong even at high concentrations of antisense oligos that could both fully saturate a ssRNA-binding site and form the exit-channel duplex. NusA-NTD alone can achieve this enhancement, but it requires a 10-bp rather than 8-bp duplex. These observations are most consistent with a direct interaction of NusA-NTD with the exit-channel duplex that is optimal with a 10-bp duplex and either inhibits translocation or stabilizes clamp opening. This direct model of NusA interaction also explains the observations that either single-stranded nascent RNA or nascent hairpin RNA crosslinks to NusA-NTD (Ha et al., 2010) and that a replacement of the 8-nt loop of the *his* pause hairpin with a UUCG tetraloop sequence (which stabilizes the hairpin but reduces the apparent duplex length) decreases the effect of NusA (Chan and Landick, 1993).

Thus, we favor the view that NusA remodels the RNA binding specificity of the RNAP exit channel to include additional duplex contacts extending up to -21 from the RNA 3′ end (a 10-bp duplex from -12 to -21). However, greater insight into the nature of direct NusA-nascent RNA interactions awaits PEC crystal structures that include both exit-channel duplexes and NusA.

RfaH and exit-channel duplexes may compete indirectly by tightest binding to different RNAP conformations

Both RfaH and its general paralog NusG, the only elongation factor conserved in all three domains of life, are thought to stabilize the closed-clamp, elongation-efficient conformation of RNAP (Belogurov et al., 2009; Klein et al., 2011; Martinez-Rucobo et al., 2011). RfaH requires interaction of its CTD with a specific nontemplate strand DNA sequence called *ops* to undergo a structural rearrangement and liberate the RNAP-binding activity of the RfaH-NTD (Artsimovitch and Landick, 2002; Belogurov et al., 2007), but both RfaH-NTD and NusG-NTD alone bind ECs though multiple contacts and exhibit antipausing activity (Belogurov et al., 2007; Mooney et al., 2009). However, RfaH-NTD, which contacts the EC on the clamp helices, the gate loop, and potentially the NT DNA strand (Figure 3-1), is more potent and can override the pausing-enhancing effects of exit-channel duplexes (Belogurov et al., 2010; Belogurov et al., 2007; Sevostyanova et al., 2011). Our findings clarify the action mechanisms of RfaH/NusG-class regulators with the key insight that RfaH and exit-channel duplex-generating oligos compete in a concentration-dependent manner for their opposite effects on pausing by RNAP (Figure 3-7). This concentration-dependence necessitates mutually exclusive binding of RfaH-NTD *versus* oligo to the EC states in which they exert their largest effects on pausing. In principle, mutually exclusive binding could occur because RfaH-NTD directly blocks exit-channel duplex formation through contacts to nascent RNA or steric exclusion in the exit channel. However, such a direct effect of RfaH-NTD on exit-channel duplexes is highly unlikely because structural models of RfaH binding, which are well constrained by mapped contacts to the clamp helices, gate loop, and NT DNA, place all of RfaH-NTD ≥25 Å from the exiting RNA with the bulky and negatively charged NT DNA strand and upstream DNA duplex between RfaH-NTD and RNA

(Belogurov et al., 2007; Sevostyanova et al., 2011). Thus, the simplest interpretation of the concentration-dependent competition of RfaH-NTD and antisense oligo for effects on pausing is that RfaH binds tightest to the closed-clamp RNAP conformation and that duplexes form more readily in the open-clamp conformation. We note that the results do not preclude RfaH binding weakly to open-clamp RNAP (*e.g.*, through a single-contact to the clamp helices) or antisense oligo binding to the closed-clamp conformation, as long as such binding results in weaker effects on pausing. Indeed, NusG, which binds ECs avidly (Mooney et al., 2009) but competes with duplex formation much less effectively than RfaH (Figure 3-7C), may primarily bind the clamp helices and make only weak clamp-closing contacts with other parts of RNAP. Interestingly, *B. subtilis* NusG (Yakhnin et al., 2008), like RfaH (Artsimovitch and Landick, 2002), can even enhance pausing at some sites, suggesting significant regulatory plasticity exists for this class of regulators.

The concentration-dependent competition of RfaH-NTD and antisense oligos also has an important implication for antitermination, regardless of the precise mechanism. In addition to antipausing, RfaH, λN, and λQ can suppress transcription termination for at least a subset of intrinsic terminators; λN and λQ appear to achieve suppression by blocking formation of the terminator hairpin through contacts also involving NusA (Gusarov and Nudler, 2001; Shankar et al., 2007; Svetlov et al., 2007). For λN and λQ a number of mechanisms for inhibiting hairpin formation have been considered, including direct antiterminator contacts to nascent RNA. However, our finding that the closed clamp may inhibit exit-channel duplex formation suggests that indirect suppression of terminator hairpin formation could be achieved simply by RfaH-, λN-, or λQ-stabilization of a closed-clamp conformation. Such a view is also consistent with crystal structures of open- and closed-clamp RNAP conformations because only the open-clamp

conformation appears able to accommodate RNA secondary structures in the nascent RNA exit channel (Tagami et al., 2010; Weixlbaumer et al., 2013).

The idea that different effects of regulators and nascent RNA structures on transcript elongation by RNAP may be mediated by competing effects on clamp position is attractive because it would allow synergy or antagonism between different regulators and structures interacting at different locations of the clamp. Depending on the strengths of interactions with open- or closed-clamp conformations, possible direct contacts among regulators and RNA structures, and the consequences of clamp conformations for RNAP activity, multiple regulatory inputs to elongation could be integrated. Much work remains, however, to understand specific interactions of regulators with the clamp and the consequences of different clamp conformations for different steps in the nucleotide addition cycle, such as translocation or catalysis mediated by trigger-loop folding.

Materials and Methods

Reagents and Oligonucleotides

All DNA and RNA oligonucleotides (Table 3-4) were obtained from IDT (Corvalville, IA) and purified by denaturing PAGE before use. [α-^{32}P]CTP was from PerkinElmer Life Sciences and NTPs were from GE Healthcare (Piscataway, NJ).

Proteins

Core *E. coli* RNAP was purified as described previously (Toulokhonov et al., 2007; Vassylyeva et al., 2002). His-tagged full-length and N-terminal domain of NusA were purified from BL21 λDE3 cells containing a NusA overexpression plasmid by a two-step purification protocol using HisTrap HP affinity chromatography followed by gel filtration on Superdex 200, as described previously (Ha et al., 2010; Kyzer et al., 2007). RfaH-NTD was obtained by overexpression of an RfaH derivative containing a TEV protease cleavage site between the NTD and CTD and containing a C-terminal hexa-histidine tag from plasmid pIA777 (kindly provided by I. Artsimovitch). *E.coli* cells (BL21 λDE3) harboring pIA777 were grown at 37 °C in LB medium containing kanamycin (25 μg/mL) until apparent OD_{600} reached 0.6. The temperature was then lowered to 20°C and isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.2 mM. Cells were harvested after overnight induction and pelleted at 2800 × *g* for 15 min at 4 °C. Cell pellets were resuspended in buffer A (50 mM Tris-HCl, pH 7.9, 0.05 mM EDTA, 8% glycerol, 500 mM NaCl, 2 mM β-mercaptoethanol) supplemented with 0.1 mg PMSF/mL and a protease inhibitor cocktail (final concentrations of 0.0125 mg benzamide/mL, 2×10^{-4} mg chymostatin/mL, 2×10^{-4} mg leupeptin/mL, 4×10^{-5} mg pepstatin/mL, 4×10^{-4} mg aprotonin/mL, 4×10^{-4} mg antipain/mL) and lysed by sonication. After centrifugation at 27,000 × g at 4 °C for 15 min, the supernatant was applied to a 5 ml HisTrap HP affinity column (GE

Healthcare). The eluted fractions containing the full-length RfaH were combined and dialyzed in buffer A overnight at 4 °C. The full-length RfaH was then cleaved with ~500 U of His_6-TEV protease in buffer A for 2 hr at 20 °C, as described previously (Belogurov et al., 2007). The cleavage products were then re-applied to the HisTrap HP column and RfaH-NTD was recovered in the flow-through. The recovered fractions containing RfaH-NTD were pooled, dialyzed against storage buffer (10 mM Tris-HCl, pH 7.9, 50% glycerol, 250 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT), and analyzed by SDS-PAGE.

In Vitro EC Reconstitution

Nucleic-acid scaffolds for reconstituting ECs were assembled in reconstitution buffer (RB; 10 mM Tris·HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$) by heating RNAs, tDNAs, and ntDNAs (0.5, 0.6, and 0.6 μM, respectively; Table 3-4) to 95 °C for 2 min, rapidly cooling to 45 °C, and then cooling to room temperature in 2 °C steps of 2 min each, as described previously (Kyzer et al., 2007). ECs reconstituted by incubating core *E. coli* RNAPs with the nucleic-acid scaffold (3:1 RNAP:scaffold) in elongation buffer (EB; 25 mM HEPES-KOH, pH 8.0, 130 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.15 mM EDTA, 5% glycerol, and 25 μg of acetylated bovine serum albumin/ml) for 15 min at 37 °C.

In vitro transcriptional pause assays

The nascent RNA in the reconstituted ECs was first radiolabeled by incubation with [α-^{32}P]CTP (2 μM; 5 Ci/mmol) in EB buffer at 50 nM EC for 1 min at 37 °C to give C18 ECs (Kyzer et al., 2007). RB buffer (as control) or antisense DNA or RNA oligos (1 μM final concentration, if not indicated otherwise) were then added to the solution and incubation was continued at 37 °C for 10 min. The C18 ECs were then incubated with 10 μM GTP and 100 μM UTP. Reaction samples were removed at the indicated times and quenched with an equal volume of 2X stop buffer (8 M

urea, 50 mM EDTA, 90 mM Tris-borate buffer, pH 8.3, 0.02% bromphenol blue, and 0.02% xylene cyanol). After the final time point, high concentrations of GTP and UTP (500 μM each) were added to chase all active ECs out of the paused state. The radiolabeled RNA reaction products were then separated by electrophoresis using a denaturing 20% polyacrylamide gel in 0.5X TBE buffer (Hein et al., 2011). Experiments were conducted with three variations: (*i*) with no oligonucleotide (RB solution); (*ii*) with RNA oligos ranging from 5 to 12 or 22 nt in length; or (*iii*) with DNA oligos ranging from 8 to 12 or 22 nt in length. The experiments were repeated with full-length NusA (~2.5 μM), NusA-NTD (4 μM) or RfaH-NTD (125 nM, 250 nM or 500 nM), which were added immediately after [α-^{32}P]CTP radiolabeling and allowed to bind for 10 min at 37 °C.

Data Quantitation and Analysis

Gels were exposed to phosphorimager screens, scanned using a Typhoon PhosphorImager, and quantitated using the ImageQuant Software (GE Healthcare). The RNA present in each lane was quantitated as a fraction of the total RNA in each lane (Figure 3-2) and corrected for the unreacted fraction remaining in the chase lane. The rates of escape from the *his* pause site (*k*; reported as half-lives, $t_{1/2}$) and the fractions EC in the one or two pause states (*E*) were obtained by fitting the data to either a single- or double-exponential decay equation. In most experiments the slow pause fraction, which may represent backtracked C18 or U19 ECs (Kyzer et al., 2007), was negligible and was not included in the calculation of pause half-life. For experiments with RfaH-NTD, two significant pause fractions were present in some conditions. Thus, for the experiments with RfaH-NTD (Figure 3-7), we calculated and reported pause strength, τ_0, as E/k where E is the pause efficiency when only one paused state existed or as $(f_1/k_1) + (f_2/k_2)$ when two paused state existed (Landick et al., 1996). f_1, f_2 are the fractions of paused RNAPs that enter

the paused states with escape rates k_1 and k_2 (see Figure 3-7D). All measurements of pause half-lives or pause strengths are reported as means ± standard deviations of experimental triplicates.

TABLE 3-4. Oligonucleotides used in this study

Stock number	Description	Sequence (5′ → 3′)
DNA		
5420	Template DNA	CTCTGAATCTCTTCCAGCACACATCAGGACGTACTGACC
5069	Non-template DNA	GGTCAGTACGTCCATTCGATCTTCGGAAGAGATTCAGAG
6601	8-bp antisense DNA oligo	CCGGATGA
7579	12-bp antisense DNA oligo	CCGGATGAAGCA
7660	11-bp antisense DNA oligo	CCGGATGAAGC
7663	10-bp antisense DNA oligo	CCGGATGAAG
7666	9-bp antisense DNA oligo	CCGGATGAA
8048	22-bp antisense DNA oligo	CCGGATGAAGCTCTACAAATGC
RNA		
6593	8-bp PEC RNA	UCAUCCGGCGAUGUGUG
6598	8-bp antisense RNA oligo	CCGGAUGA
7576	12-bp PEC RNA	UGCUUCAUCCGGCGAUGUGUG
7578	12-bp antisense RNA oligo	CCGGAUGAAGCA
7658	11-bp PEC RNA	GCUUCAUCCGGCGAUGUGUG
7659	11-bp antisense RNA oligo	CCGGAUGAAGC
7661	10-bp PEC RNA	CUUCAUCCGGCGAUGUGUG
7662	10-bp antisense RNA oligo	CCGGAUGAAG
7664	9-bp PEC RNA	UUCAUCCGGCGAUGUGUG
7665	9-bp antisense RNA oligo	CCGGAUGAA
7694	7-bp PEC RNA	AGUCAGGCGAUGUGUG
7695	7-bp antisense RNA oligo	CCUGACU
7696	6-bp PEC RNA	GUCAGGCGAUGUGUG
7697	6-bp antisense RNA oligo	CCUGAC
7698	5-bp PEC RNA	UCAGGCGAUGUGUG
7699	5-bp antisense RNA oligo	CCUGA
8043	8-bp U-tail RNA oligo	UUUUUUUUCCGGAUGA
8045	8-bp hairpin-tail RNA oligo	CGCUGCAUUGCUGCAGCGCCCGGAUGA
8046	22-bp PEC RNA	GCAUUUGUAGAGCUUCAUCCGGCGAUGUGUG
8047	22-bp antisense RNA oligo	CCGGAUGAAGCUCUACAAAUGC

Acknowledgements

We thank members of the Landick laboratory for many helpful discussions and for comments on the manuscript, Rachel Mooney for providing plasmids, and Irina Artsimovitch for providing pIA777 plasmid and the RfaH purification protocol.

References

Artsimovitch, I., and Landick, R. (1998). Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. Genes Dev *12*, 3110-3122.

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

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell *109*, 193-203.

Artsimovitch, I., Svetlov, V., Murakami, K., and Landick, R. (2003). Co-overexpression of *E. coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. J Biol Chem *278*, 12344-12355.

Belogurov, G.A., Mooney, R.A., Svetlov, V., Landick, R., and Artsimovitch, I. (2009). Functional specialization of transcription elongation factors. EMBO J *28*, 112-122.

Belogurov, G.A., Sevostyanova, A., Svetlov, V., and Artsimovitch, I. (2010). Functional regions of the N-terminal domain of the antiterminator RfaH. Mol Microbiol *76*, 286-301.

Belogurov, G.A., Vassylyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassylyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. Mol Cell *26*, 117-129.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R.*, et al.* (2012). Opening and closing of the bacterial RNA polymerase clamp. Science *337*, 591-595.

Chan, C.L., and Landick, R. (1993). Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. J Mol Biol *233*, 25-42.

Chan, C.L., and Landick, R. (1997). Effects of neutral salts on RNA chain elongation and pausing by Escherichia coli RNA polymerase. J Mol Biol *268*, 37-53.

Cisse, II, Kim, H., and Ha, T. (2012). A rule of seven in Watson-Crick base-pairing of mismatched sequences. Nat Struct Mol Biol *19*, 623-627.

Dock-Bregeon, A.C., Chevrier, B., Podjarny, A., Johnson, J., de Bear, J.S., Gough, G.R., Gilham, P.T., and Moras, D. (1989). Crystallographic structure of an RNA helix: [U(UA)6A]2. J Mol Biol *209*, 459-474.

Farnham, P.J., Greenblatt, J., and Platt, T. (1982). Effects of NusA protein on transcription termination of the tryptophan operon of *Escherichia coli.* Cell *29*, 945-951.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Mol Cell *3*, 495-504.

Gusarov, I., and Nudler, E. (2001). Control of intrinsic transcription termination by N and NusA: the basic mechanisms. Cell *107*, 437-449.

Ha, K.S., Toulokhonov, I., Vassylyev, D.G., and Landick, R. (2010). The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. J Mol Biol *401*, 708-725.

Hein, P.P., Palangat, M., and Landick, R. (2011). RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. Biochemistry *50*, 7002-7014.

Horton, N.C., and Finzel, B.C. (1996). The structure of an RNA/DNA hybrid: a substrate of the ribonuclease activity of HIV-1 reverse transcriptase. J Mol Biol *264*, 521-533.

Kassavetis, G.A., and Chamberlin, M.J. (1981). Pausing and termination of transcription within the early region of bacteriophage T7 DNA *in vitro*. J Biol Chem *256*, 2777-2786.

Kireeva, M., Kashlev, M., and Burton, Z.F. (2010). Translocation by multi-subunit RNA polymerases. Biochim Biophys Acta *1799*, 389-401.

Klein, B.J., Bose, D., Baker, K.J., Yusoff, Z.M., Zhang, X., and Murakami, K.S. (2011). RNA polymerase and transcription elongation factor Spt4/5 complex structure. Proc Natl Acad Sci U S A *108*, 546-550.

Kyzer, S., Ha, K.S., Landick, R., and Palangat, M. (2007). Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. J Biol Chem *282*, 19020-19028.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.

Landick, R., Carey, J., and Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the trp operon leader region. Proc Natl Acad Sci U S A *82*, 4663-4667.

Landick, R., Wang, D., and Chan, C. (1996). Quantitative analysis of transcriptional pausing by RNA polymerase: the *his* leader pause site as a paradigm. Meth Enzymol *274*, 334-352.

Landick, R., and Yanofsky, C. (1984). Stability of an RNA secondary structure affects *in vitro* transcription pausing in the *trp* operon leader region. J Biol Chem *259*, 11550-11555.

Malinen, A.M., Turtola, M., Parthiban, M., Vainonen, L., Johnson, M.S., and Belogurov, G.A. (2012). Active site opening and closure control translocation of multisubunit RNA polymerase. Nucleic Acids Res *40*, 7442-7451.

Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. EMBO J *30*, 1302-1310.

Mooney, R.A., Schweimer, K., Rosch, P., Gottesman, M., and Landick, R. (2009). Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. J Mol Biol *391*, 341-358.

Nakano, S., Fujimoto, M., Hara, H., and Sugimoto, N. (1999). Nucleic acid duplex stability: influence of base composition on cation effects. Nucleic Acids Res *27*, 2957-2965.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. Mol Cell *50*, 882-893.

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

Opalka, N., Brown, J., Lane, W.J., Twist, K.A., Landick, R., Asturias, F.J., and Darst, S.A. (2010). Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. PLoS Biol *8*.

Owczarzy, R., Moreira, B.G., You, Y., Behlke, M.A., and Walder, J.A. (2008). Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. Biochemistry *47*, 5336-5353.

Owczarzy, R., You, Y., Moreira, B.G., Manthey, J.A., Huang, L., Behlke, M.A., and Walder, J.A. (2004). Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures. Biochemistry *43*, 3537-3554.

Pan, T., Artsimovitch, I., Fang, X.W., Landick, R., and Sosnick, T.R. (1999). Folding of a large ribozyme during transcription and the effect of the elongation factor NusA. Proc Natl Acad Sci U S A *96*, 9545-9550.

Pan, T., and Sosnick, T. (2006). RNA folding during transcription. Annu Rev Biophys Biomol Struct *35*, 161-175.

Peters, J.M., Vangeloff, A.D., and Landick, R. Bacterial transcription terminators: the RNA 3'-end chronicles. J Mol Biol *412*, 793-813.

Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science *328*, 504-508.

Sevostyanova, A., Belogurov, G.A., Mooney, R.A., Landick, R., and Artsimovitch, I. (2011). The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Mol Cell *43*, 253-262.

Shankar, S., Hatoum, A., and Roberts, J.W. (2007). A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. Mol Cell *27*, 914-927.

Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995). Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry *34*, 11211-11216.

Svetlov, V., Belogurov, G.A., Shabrova, E., Vassylyev, D.G., and Artsimovitch, I. (2007). Allosteric control of the RNA polymerase by the elongation factor RfaH. Nucleic Acids Res *35*, 5694-5705.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., and Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. Nature *468*, 978-982.

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

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

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.

Vassylyeva, M.N., Lee, J., Sekine, S.I., Laptenko, O., Kuramitsu, S., Shibata, T., Inoue, Y., Borukhov, S., Vassylyev, D.G., and Yokoyama, S. (2002). Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from Thermus thermophilus. Acta Crystallogr D Biol Crystallogr *58*, 1497-1500.

Wang, D., Severinov, K., and Landick, R. (1997). Preferential interaction of the his pause RNA hairpin with RNA polymerase beta subunit residues 904-950 correlates with strong transcriptional pausing. Proc Natl Acad Sci U S A *94*, 8433-8438.

Weixlbaumer, A., Leon, K., Landick, R., and Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. Cell *152*, 431-441.

Wickiser, J.K., Winkler, W.C., Breaker, R.R., and Crothers, D.M. (2005). The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. Mol Cell *18*, 49-60.

Worbs, M., Bourenkov, G.P., Bartunik, H.D., Huber, R., and Wahl, M.C. (2001). An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. Mol Cell *7*, 1177-1189.

Xia, T., SantaLucia, J., Jr., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C., and Turner, D.H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry *37*, 14719-14735.

Yakhnin, A.V., Yakhnin, H., and Babitzke, P. (2008). Function of the Bacillus subtilis transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader. Proc Natl Acad Sci U S A *105*, 16131-16136.

Yang, X., and Lewis, P.J. (2010). The interaction between RNA polymerase and the elongation factor NusA. RNA Biol *7*.

Yang, X., Molimau, S., Doherty, G.P., Johnston, E.B., Marles-Wright, J., Rothnagel, R., Hankamer, B., Lewis, R.J., and Lewis, P.J. (2009). The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. EMBO Rep *10*, 997-1002.

Zhang, J., and Landick, R. (2009). Substrate loading, nucleotide Addition, and translocation by RNA polymerase. In RNA Polymerase as Molecular Motors, H. Buc, and T. Strick, eds. (London: Royal Society of Chemistry), pp. 206-235.

Zhang, J., Palangat, M., and Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. Nat Struct Mol Biol *17*, 99-104.

Chapter Four

RNA polymerase pausing and nascent RNA structure formation are energetically linked through clamp domain movement

This chapter will be published as a research article with me as first author.

Pyae P Hein, Kellie E Kolb, Tricia Windgassen, Michael J Bellecourt, Rachel Anne Mooney, and, Robert Landick

Abstract

The rate of RNA synthesis, controlled by transcriptional pausing, and folding of nascent RNA into biologically active structures are kinetically interdependent and are influenced by elongation factors. In bacteria, pausing is required for proper folding of some RNAs. Depending on location, structures in the nascent RNA can increase pausing by interacting with RNA polymerase (RNAP), dissociate RNAP, or decrease pausing by preventing backtracking. In archaea and eukaryotes, the interplay of RNA folding and pausing is thought to be equally important, but is less well understood. In bacterial RNAP, opening of the clamp domain has been proposed to mediate effects of nascent RNA structures. However, the connections among RNA structure formation, clamp movement, translocation, and catalytic activity remain poorly understood. Using a direct fluorescence measurement of exit-channel duplex formation, disulfide crosslinks that trap the clamp in the closed or open conformation, and RNAP substitutions and deletions that affect pausing, we report that nascent RNA structures affect the RNAP active site by stabilizing an open-clamp RNAP conformation. Our results also suggest that clamp conformations linked to pausing can control the rates of nascent RNA structure formation.

Introduction

Although it has been appreciated for decades that folding of RNA into biologically active structures is controlled by the rate of RNA synthesis from its 3′ end in the active site of RNA polymerase (RNAP; Brehm and Cech, 1983; Nussinov and Tinoco, 1981), the mechanisms by which RNAP may guide RNA folding remain poorly understood. The locations and durations of pauses by both bacterial RNAP and eukaryotic RNAPII can guide proper RNA folding (Lai et al., 2013; Matysiak et al., 1999; Pan and Sosnick, 2006). Abrogation of pausing can eliminate kinetic windows essential for formation of metastable structures involved in riboswitch or ribozyme function (Pan et al., 1999; Perdrizet et al., 2012; Wickiser et al., 2005; Wong and Pan, 2009), and synthesis of RNA using pause-resistant T7 RNAP can produce biologically inactive RNA *in vivo* (Lewicki et al., 1993). Conversely, RNA structures can influence the rate of transcript elongation either by stimulating prolonged pausing through interactions in the RNA exit channel (Toulokhonov et al., 2001) or by preventing RNAP backtracking (Nudler, 2012; Zamft et al., 2012).

The possible interplay of RNA folding and transcriptional pausing and its modulation by direct effects of RNAP or associated factors has been studied most extensively in bacteria, especially using *Escherichia coli* as a model. RNA structures are known to form with their duplex stems within the RNA exit channel of RNAP at sites of transcriptional pausing. Transient pauses that occur ubiquitously during RNA synthesis (Neuman et al., 2003) are thought to arise by a rearrangement in the RNAP active site that blocks loading of the template base into the active site, allowing time for RNA structure formation, for backtracking, or for regulator and ribosome interactions (Herbert

et al., 2006; Landick, 2006; Weixlbaumer et al., 2013). These "elemental" paused elongation complexes (ePECs) thus serve as regulatory intermediates whose duration can be modified in a variety of ways, including the nascent RNA stem-loop structures. When its stem extends to -13 or -12 (positions relative to -1 or the RNA 3´ end), a nascent RNA structure can increase pause dwell times by factors of 10 or more (Figure. 4-1A and B; Toulokhonov et al., 2001; Yakhnin and Babitzke, 2010). This dwell time is increased through interactions with RNAP thought to require the flap tip on one side of the RNA exit channel (Toulokhonov and Landick, 2003) and other interactions thought to affect the RNAP active site by stabilizing an open conformation of the clamp domain (Nayak et al., 2013; Toulokhonov et al., 2007; Weixlbaumer et al., 2013). When the stem extends closer to the active site, the pause stimulating effect of the hairpin is lost and can drive dissociation of the elongation complex (EC). Elongation regulators, such as the multidomain NusA, also play a role in stabilization of the nascent RNA structure. The NusA modulates the interactions of nascent structures with RNAP, both by enhancing exit-channel duplex stimulation of pausing or termination via contacts of its N-terminal domain with the exiting duplex and the flap tip (Ha et al., 2010) and by promoting formation of biologically active nascent RNA structures for ribozymes and riboswitches (Pan et al., 1999; Perdrizet et al., 2012; Wickiser et al., 2005). Despite the role for RNA structure in pausing, questions remain regarding how these exit-channel duplexes stimulate pausing in addition to whether RNAP actively promotes RNA structure either through direct contacts or is kinetically driven.

The hairpin-stimulated pause in the *E. coli his* operon leader region has provided a particularly useful model to investigate these issues. Here, a 5-bp stem, 8-nt loop "pause

hairpin" that forms 12 nt from the RNA 3′ end and delays RNAP, serving to synchronize transcription and translation during attenuation has been found to prevent formation of a helical hairpin form of the trigger loop (TL) required for rapid catalysis in the RNAP active site (Nayak et al., 2013). An open-clamp conformation of RNAP stabilized by the pause hairpin is postulated to trap the TL in a partially folded, paused conformation through intermediate effects on the bridge helix (BH). However, the hairpin-promoted paused state appears to be trapped in the pretranslocated register by fraying of the RNA 3′ end. Both exit-channel RNA:RNA duplex-RNAP and NusA-paused elongation complex (PEC) interactions may inhibit translocation to allow NTP binding, and thus the relative contributions of inhibited translocation and inhibited TL folding to hairpin-promoted pausing remain unknown.

To investigate the interplay of nascent RNA structure formation, RNAP interactions, and transcriptional pausing, we have applied a combination of stopped-flow fluorescence assays that can detect exit channel duplex formation or translocation, disulfide crosslinks that can trap RNAP in the closed- or open-clamp conformations, and substitutions in RNAP that abrogate hairpin-stimulated pausing. Our results define roles of RNAP, NusA, and RfaH in mediating nascent structure formation and establish that clamp opening itself can stimulate pausing. Our findings also demonstrate that inhibition of both TL folding and translocation make significant contributions to the effects of exit-channel duplexes on the rate of pause escape.

Results

The RNAP exit channel allows efficient formation of RNA structures

To investigate how the RNAP exit channel affects formation of nascent RNA structures, we first compared binding of an 8-nt antisense RNA oligonucleotide (asRNA) to form an 8-bp nascent RNA:RNA duplex in the exit channel, mimicking the *his* pause hairpin or to form a duplex with the same RNA free in solution. We monitored duplex formation by measuring fluorescence quenching of pyrrolo-cytosine (PC; 3-[β-D-2-ribofuranosyl]-6-methylpyrrolo [2,3-d]pyrimidin-2(3H)-one) (Berry, 2004; Tinsley and Walter, 2006) positioned near the 5′ end of the nascent RNA (Figure 4-1A). PC forms Watson-Crick base pairs with G similarly to unmodified C and is quenched by stacking interactions with adjacent bases in duplex structures (Berry, 2004; Liu and Martin, 2001, 2002; Tinsley and Walter, 2006). To reduce stacking in the ssRNA and increase fluorescence of unpaired PC, we positioned PC between pyrimidines (5′-U-PC-C-3′). RNA:RNA duplex formation was readily detectable by ~15% decrease in PC fluorescence upon addition of the antisense 8mer, with little to no change upon addition of a noncomplementary 8mer RNA (Figures 4-2A, B, C and D). This result is consistent with prior measurements of PC fluorescence in DNA and in DNA:DNA duplexes (Johnson et al., 2005) or in RNA (Dash et al., 2004).

To determine whether RNAP promotes or inhibits nascent RNA:RNA duplex formation, we formed ePECs with 1.5 μM *E. coli* RNAP on 0.5 μM nucleic acid scaffold containing the PC RNA with sequences from the *his* operon leader region (Figure 4-1A). The pause duration of the *his* ePEC is ordinarily increased ~10-fold by the *his* pause RNA hairpin that forms 12 nt from the nascent RNA 3′ end, between the flap and clamp domains

where it is thought to stabilize an open-clamp conformation of RNAP (Chan et al., 1997; Weixlbaumer et al., 2013). The effect of the pause hairpin can be replicated by an 8-nt asRNA bound to PC-containing pause RNA lacking the upstream arm of the hairpin (Figures 4-1A and B; (Ha et al., 2010; Kolb et al., 2013). *E. coli* RNAP was bound to a "bubble" ePEC scaffold containing two complementary DNA oligos with a non-complementary bubble at template bases that bound the nascent RNA oligo. We measured the rate of the duplex RNA formation by mixing the ePECs with 1 μM asRNA and followed the decrease in fluorescence signal in a stopped-flow spectrofluorimeter (Figure 4-1B; Materials and Methods). PC fluorescence decreased rapidly with an observed rate of 0.94 ± 0.03 s^{-1} (Figure 4-1C). To obtain the association rate (k_{on}) of asRNA binding to the ePEC, we determined observed PC fluorescence quenching rates at increasing concentrations of asRNA (0.4 ± 0.02 μM^{-1} s-1; Figure 4-1C, inset). This association rate was only ~2-fold slower than the rate we measured for asRNA binding to the scaffold RNA in the absence of RNAP (0.76 ± 0.02 μM^{-1} s^{-1}; Figure 4-1F). This association rate of asRNA for free scaffold RNA is comparable to previous estimates for rates of formation for duplexes of 8-20 bp (~0.23-1.1 μM^{-1} s^{-1}; Braunlin and Bloomfield, 1991; Cisse et al., 2012; Kinjo and Rigler, 1995; Wetmur, 1991).

We next measured the dissociation rate (k_{off}) for asRNA oligo release from the nascent RNA:RNA duplex or free scaffold duplex using a competitor RNA lacking PC. We first formed the duplex-containing RNAs by annealing asRNA at 1 μM, and then mixed the duplex species with 400-fold excess competitor RNA in the stopped-flow instrument (Figure 4-1D). A time-dependent increase in PC fluorescence was observed, reflecting regeneration of single-stranded PC-containing RNA as the asRNA transferred

to the unlabeled competitor RNA (k_{off} = 0.10 ± 0.01 s^{-1}; Figure 4-1E). RNAP had no effect on the off-rate, as the parallel experiment using PC scaffold alone gave a comparable dissociation rate constant (k_{off} = 0.10 ± 0.03 s^{-1}; Figure 4-1F). Thus, the affinity of asRNA for scaffold *vs*. the PEC differs by only a factor of 2 (K_d of 130 ± 40 nM *vs.* 250 ± 30 nM, respectively). Thus, RNAP alone does not substantially stabilize nascent RNA structure. The slower on-rate could be explained simply by steric restriction of routes for asRNA-nascent RNA encounters, but a balance among steric, electrostatic, and noncovalent bonding effects also is possible. Thus, duplex formation, rather than dissociation, appears to be the step modulated by RNAP. It also is the biologically relevant step since RNA structures, once formed, typically are stable on the time scale of transcription (Lai et al., 2013). For these reasons and to simplify experiments, we measured on-rates to explore the features of RNAP that control nascent RNA structure formation.

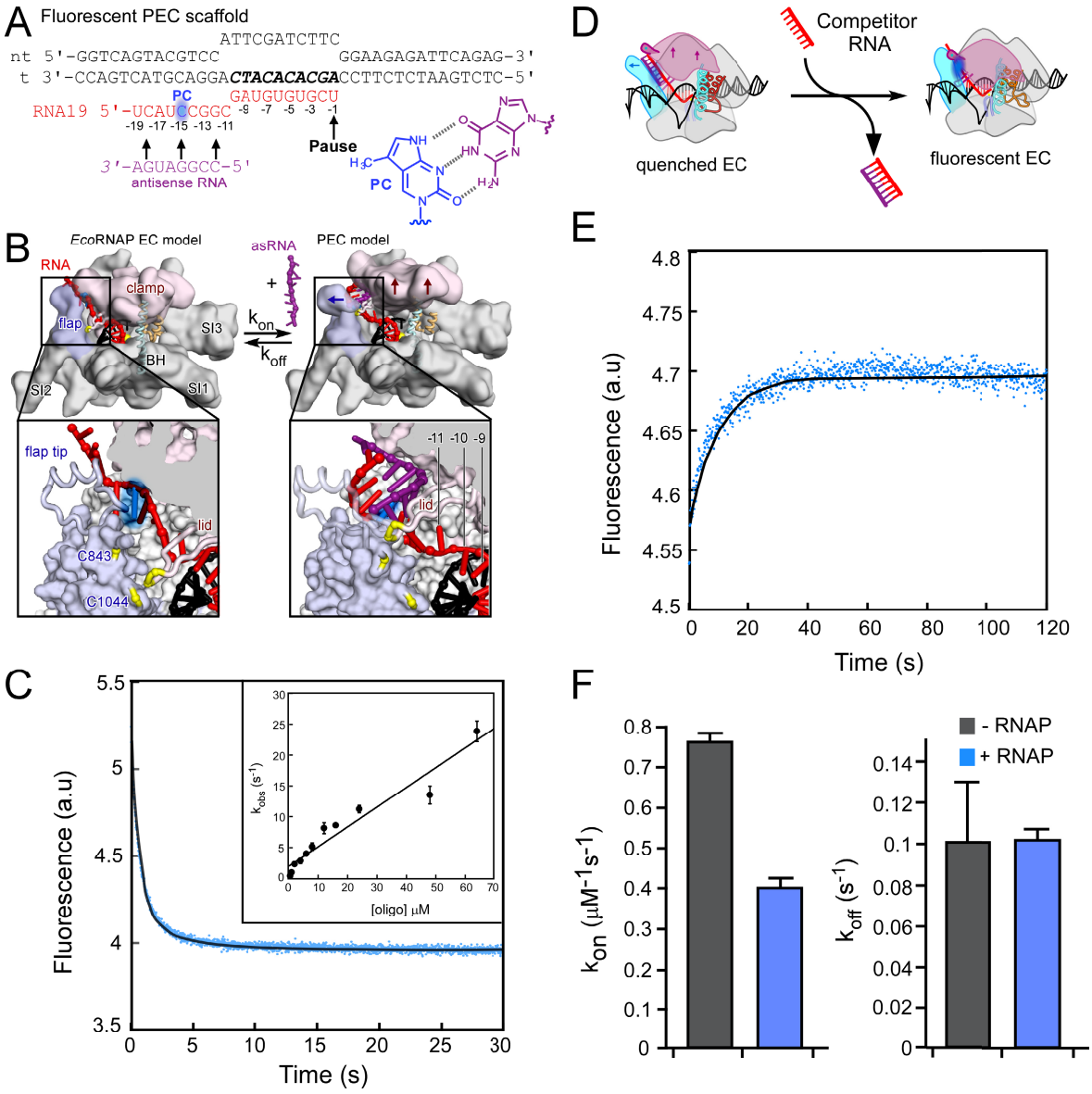


Figure 4-1. Slight inhibition of the RNA:RNA duplex formation by RNAP.

(A) A nucleic acid scaffold (ntDNA-5069, tDNA-5420, RNA-7604; Table 4-4). RNA (red) containing PC (highlighted in blue) containing RNA (red) used for EC reconstitution and for the exit-channel duplex formation studies. The position of the pause (U_{19}) is indicated by an arrow. The 8-mer asRNA oligo is shown in purple. Base-pairing between PC and GMP (purple) is shown below the nucleic acid scaffold.

Figure 4-1 legend (cont.)

(B) Models of *E. coli* RNAP EC and PEC. Left, closed-clamp, active EC model (PDB 2o5i (Vassylyev et al., 2007a)) showing clamp (light pink), flap (light blue), bridge helix (BH), and location of major sequence insertions (SI1, SI2, SI3). The RNA:DNA hybrid (red and black), the exiting RNA (red), and 8-nt antisense RNA (purple) are indicated. Right, open-clamp paused EC (PDB 4gzy(Weixlbaumer et al., 2013)) with arrows indicating clamp and flap movement upon RNA:RNA duplex formation. The blow-up shows the location of PC (blue) in the exiting RNA and a decrease in fluorescence of PC upon formation of the RNA:RNA duplex in the exit channel.An 8-bp exit channel duplex (purple and red) is modeled into the expanded RNA exit channel.

(C) A representative time trace at 250 nM ECs and 1 μM 8mer asRNA oligo. PECs were formed by assembling *E. coli* RNAP on the scaffold shown in A. Fluorescence of PC decreases with a rate of 0.94 ± 0.03 s^{-1} (k_{obs}) upon addition of the 8mer asRNA oligo. Inset; the oligo concentration dependence of the observed rate constant (k_{obs}) upon addition of excess 8mer asRNA oligo to PECs containing PC RNA. The slope of the linear regression line yields the on-rate (k_{on}; Materials and Methods).

(D) Schematic of an assay to determine the off-rate (k_{off}). Colors are as in panel B.

(E) Addition of 100 μM competitor RNA to quenched ECs (500 nM) with the exit-channel duplex RNA:RNA increases the PC fluorescence, indicating the release of an 8-mer asRNA oligo from the duplex in the EC. The off-rate was determined by fitting data points to a single exponential equation (Materials and Methods).

(F) The bimolecular rate constants of the 8-bp RNA:RNA formation or the dissociation rate of the 8-mer asRNA oligo from the duplex with and without RNAP was calculated and plotted.

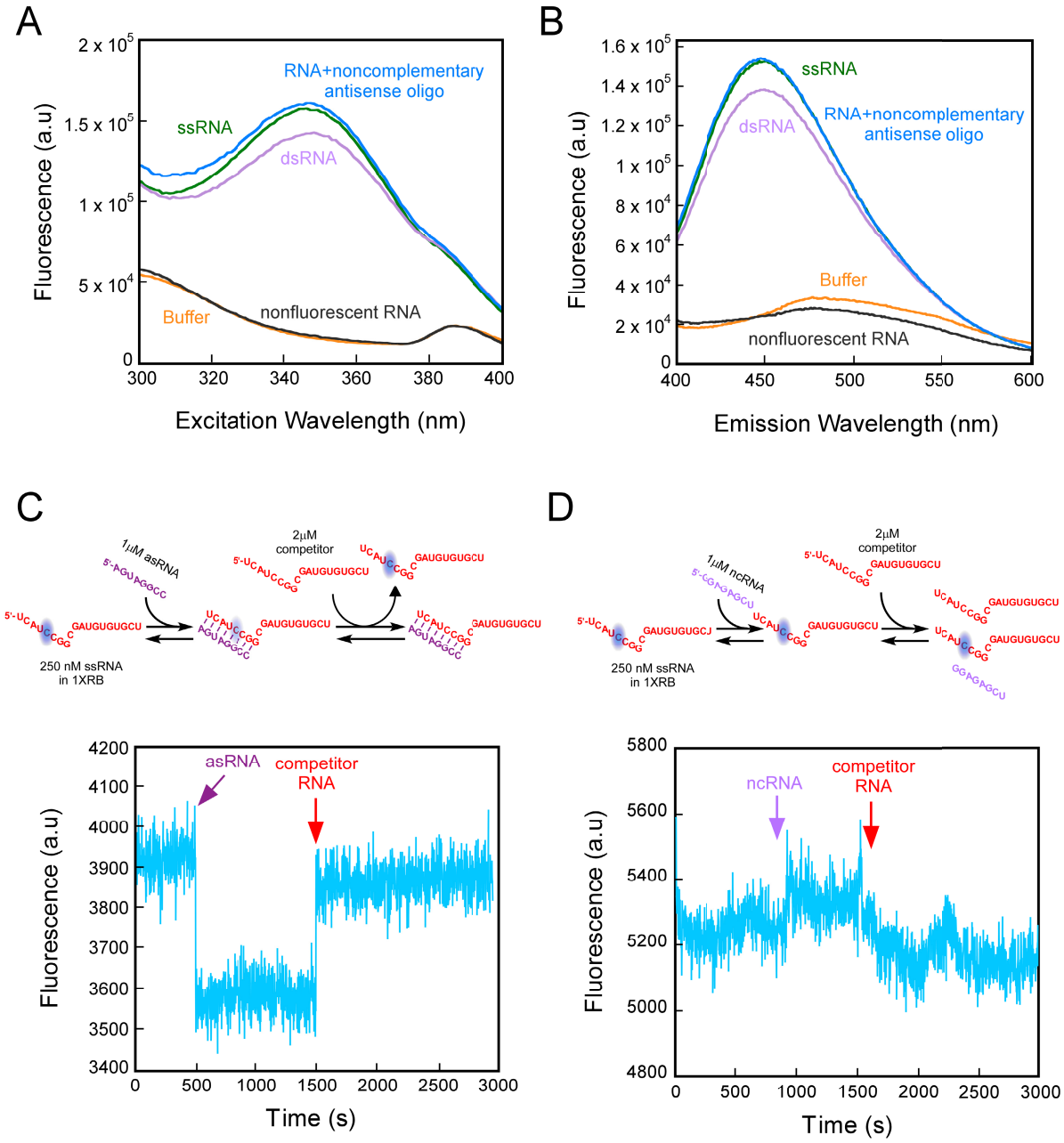


Figure 4-2. Formation of the 8-bp RNA:RNA duplex by pairing the 8mer asRNA oligo to the PC containing nascent RNA.

Figure 4-2 legend.

(A) Fluorescence excitation spectra of ssRNA (green), dsRNA (purple), ssRNA with 8mer noncomplementary RNA (blue), ssRNA without PC (black), and buffer (orange), are shown.

(B) Fluorescence emission spectra of ssRNA, dsRNA, ssRNA with 8mer noncomplementary RNA, ssRNA without PC, and buffer, are indicated and colored as in A.

(C) Kinetic studies of the association of an 8mer RNA oligo (dark purple; RNA-6598; Table 4-4) to the PC containing ssRNA (red; RNA-7604; Table 4-4). Dissociation of the 8mer RNA from the ssRNA is triggered by the presence of competitor RNA containing no PC (RNA-7418; Table 4-4).

(D) Quenching of PC fluorescence is specific to the duplex formation. Addition of the non-complementary RNA oligo (purple) does not change the fluorescent signal.

The RNAP flap tip is required for duplex-stimulated pausing but not for duplex formation

We next asked how the RNAP flap domain affected formation and effects of nascent RNA structures. The flap domain forms one wall of the RNA exit channel near the RNA from nt -11 to -17, but makes little direct contact with the exiting RNA (Vassylyev et al., 2007a). However, deletion of the flexible flap tip (β subunit residues 890-914 in *E. coli* RNAP; ΔFT1, Figure 4-3A) prevents pause stabilization by the *his* pause hairpin even though the hairpin appears to form normally (Toulokhonov and Landick, 2003). This result is disputed by Kuznedelov *et al.* (Kuznedelov et al., 2006), who report that a slightly larger flap-tip deletion (884-914; ΔFT2, Figure 4-3A) does not affect pausing, and argue that the 890-914 deletion interferes with pause hairpin formation. To clarify the effect of the flap tip on RNA structure formation and pausing, we purified both flap-tip deletion RNAPs and tested their effects on duplex-stimulation of pausing and on duplex formation (Figure 4-3). Neither flap-tip deletion affected the duration of the elemental pause, but both deletions eliminated duplex-stimulation of pausing either with or without NusA present (Figure 4-3B); NusA is known to enhance the effect of exit-channel duplexes with wild-type RNAP (Ha et al., 2010; Kolb et al., 2013). Based on these observations, we predicted that NusA would promote the exit-channel duplex formation. Indeed, we found that NusA increased asRNA oligo binding to the exiting RNA of the WT-ECs by a factor of ~2 (k_{obs} = 3.3 ± 0.2 s^{-1} with NusA *vs.* 1.8 ± 0.3 s^{-1} without NusA; Figure 4-3C). Interesting, the N-terminal domain of NusA (NusA-NTD) is sufficient for promoting the formation of exit-channel duplex (Figure 4-3C).

We next tested the effects of flap-tip deletions on the exit-channel duplex formation. Neither deletion interfered with formation of the exit-channel duplex, as detected in the PC quenching assay (Figure 4-3C). The smaller 890-914 flap-tip deletion allowed a slightly faster rate of asRNA binding, close to that observed for scaffold RNA alone and consistent with removal of a steric block to duplex formation. The larger 884-914 deletion exhibited an intermediate rate of asRNA binding, possibly because the bulky side chains or altered endpoints disrupted flap folding and interfered with asRNA binding. We noted that the observed rates of duplex formation (k_{obs}) observed for WT RNAP in the presence of NusA are similar to rates observed for scaffold RNA alone and for flap-tip deletion RNAPs (Figure 4-3C). The fact that NusA does not promote k_{obs} to be faster than scaffold RNA alone or ΔFT RNAPs suggest that NusA removes a steric block to duplex formation excreted by the flap-tip through direct interactions with the flap-tip. We conclude that the flap-tip is essential for the effect of exit channel duplexes on pausing, but does not promote the formation of exit channel structures. Instead, an interaction of duplexes with the flap tip, after the duplex forms, is necessary for pause stabilization. These findings are consistent with the model that interaction of an exit channel duplex with the flap tip creates clash that promotes clamp opening and consequent effects on the RNAP active site (Toulokhonov and Landick, 2003; Toulokhonov et al., 2007).

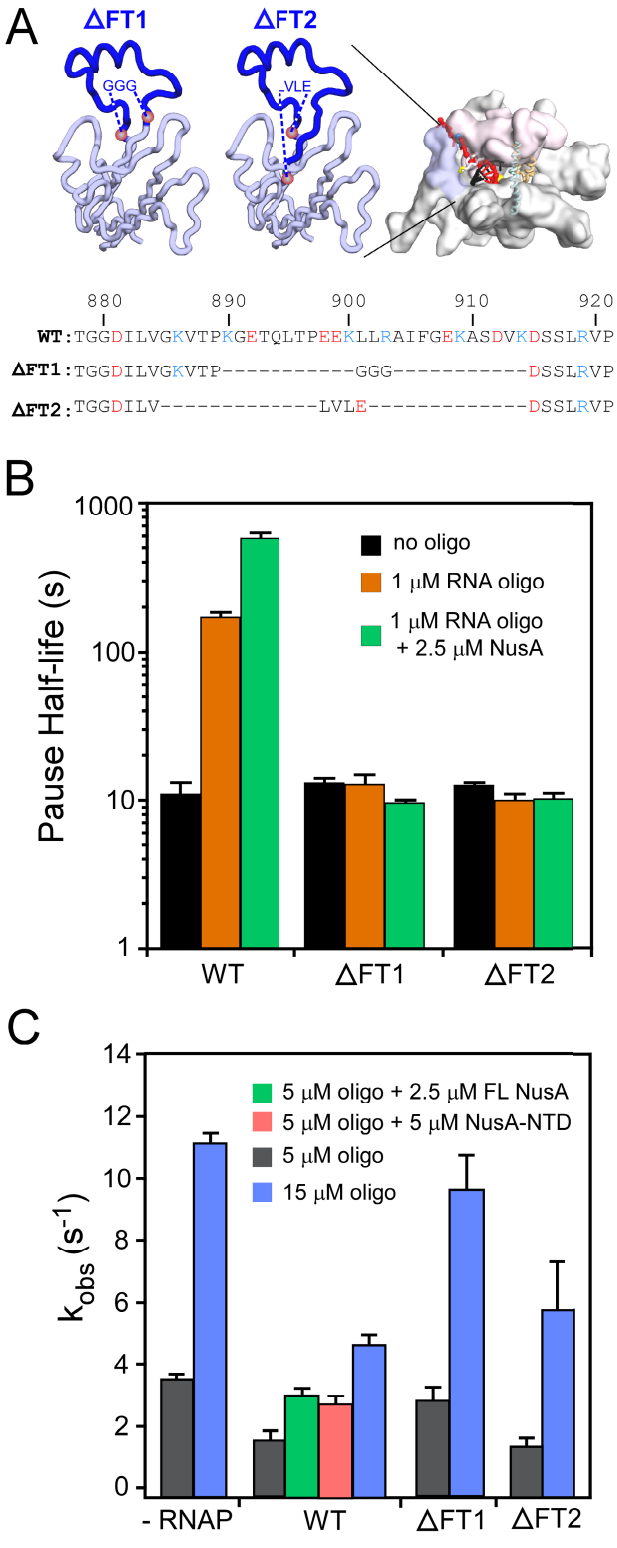


Figure 4-3. Effects of the flap-tip deletion on oligo-mediated pausing and the exit-channel duplex formation.

Figure 4-3 legend.

(A) Structure of *E. coli* RNAP EC showing the β flap region (light blue) and the location of two β flap-tip deletions (dark blue). Endpoints of these β flap-tip deletions are marked by red spheres (top). The β flap-tip sequence of WT and two flap-tip deletion variants are shown (bottom). The acidic amino acid residues are shown in red and the basic residues in blue

(B) Effect of flap-tip deletion on oligo-stabilized pausing. Pause half-life analysis for ECs reconstituted with wild-type and ΔFT *E. coli* RNAPs. ECs were reconstituted on the PEC scaffold that lacks PC (Figure 4-1A) at 2 nt before the pause site yielding EC_{G17} and halted at C_{18} position by incubating with 2 μM [α-^{32}P]-CTP. ECs were elongated through the *his* pause site (U_{19}) and pause half-lives of ECs were measured (Materials and Methods) in the absence or presence of 1 μM 8mer asRNA oligo, with or without full-length NusA (2.5 μM).

(C) RNA:RNA duplex formation is not substantially affected by the deletion of flap-tip. The observed rate of the duplex formation (k_{obs}) for the nucleic acid scaffold lacking RNAP or ECs containing wild-type in the presence or absence of 2 μM NusA or 5 μM Nus-NTD and ΔFT RNAPs were determined with 5 μM and 15 μM asRNA oligo at 37 °C.

Clamp-flap disulfide crosslinks: clamp opening is necessary for pause stimulation by exit-channel duplexes

The RNA duplex-flap interaction required for hairpin-stimulation of pausing could be explained by consequences other than effects on clamp opening. For instance, flap-duplex interactions could inhibit translocation of the RNA/DNA scaffold through RNAP. Thus, we next sought to test directly for effects of exit-channel duplexes on clamp conformation and for effects of clamp conformation on the formation of exit channel duplexes. For this purpose, we inserted Cys residues at locations that were predicted to form disulfide boinds between the lid and flap domains in either the closed clamp conformation (β′258iC-β1044C; based on *Tth*RNAP EC, pdb 2o5i) (Vassylyev et al., 2007a) or the open clamp conformation (β′258iC-β843C; based on *Tth*RNAP ePEC, pdb 4gzy) (Weixlbaumer et al., 2013) of RNAP (Figures 4-1B, 4-4A and 4-4C). β′258iC is an insertion of Cys between β′ G257 and β′ G258 in the lid domain; hereafter, β′258iC-β1044C or β′258iC-β843C are referred to as closed or open, respectively.

We first tested for formation of the expected crosslinks as a function of redox potential by non-reducing SDS-PAGE of reconstituted ePECs using mixtures of cystamine (CSSC) and DTT (Figures 4-4A and C; (Nayak et al., 2013). The closed and open crosslinks were readily detected by the slower electrophoretic mobility of crosslinked β′-βsubunits. As the redox potential decreased (less negative values), the fraction of crosslinked species increased as expected (Figures 4-5A and B). We chose -0.24 V to study the effect of clamp conformation on pausing and duplex formation because it gave the highest fraction of crosslinking (~0.6 for both open and closed crosslinks); higher concentrations of CSSC decreased crosslinking by driving mixed

disulfide formation (Figure 4-5B). We next tested the ability of another oxidizing agent, diamide, on the fraction crosslinked for closed and open in ECs by varying the concentration of diamide. We found that diamide increased the fraction crosslinked (up to ~ 90%) for both closed and open relative to CSSC (~ 65%). However, a shift in mobility of the crosslinked species during SDS-PAGE was observed using diamide, becoming more apparent at higher diamide concentrations (Figure 4-6A). Despite a higher percentage of crosslinked observed with diamide, we chose to use CSSC in our fluorescent assays described below because diamide interfered with the fluorescent signal.

We next tested whether stabilizing the clamp in the closed or open conformation with disulfide crosslinks affected duplex stimulation of pausing. To test the effect of inhibiting clamp opening, we formed PECs on the ePEC scaffold (Figure 4-1A) by addition of the asRNA and both CTP and UTP to halt RNAP at the *his* pause site. PECs containing wild-type RNAP exhibited similar pause durations after addition of GTP with or without CSSC treatment, verifying that CSSC treatment did not affect RNAP elongation kinetics significantly (Table 4-1; see also Nayak et al., 2013). Similar pausing behavior with and without asRNA also was observed for PECs containing noncrosslinked closed RNAP (Table 4-1; 1 mM DTT), establishing that the Cys substitutions alone had at most modest effects on pause kinetics. However, PECs containing ~50% crosslinked closed RNAP displayed two distinct populations of PECs. One fraction gave a pause dwell time >100 s, about equal to that observed with PECs containing the exit-channel duplex (Figures 4-4B and 4-5C). The other fraction, ~50% of the PECs, exhibited a much shorter dwell time (~10 s) about equal to that of PECs lacking the exit-channel duplex (*i.e.*, the ePEC). Thus, formation of the closed-clamp disulfide crosslink appeared to

eliminate stimulation of pausing by the exit-channel duplex. We verified that the closed-clamp crosslink eliminates the effect of an exit-channel duplex on pausing by crosslinking PECs containing closed RNAP and asRNA oligo with diamide (Table 4-1). Although diamide could not be used in our oligo-binding fluorescence experiments because it is inherently fluorescent, diamide gives a high fraction of flap-lid disulfide crosslinking because it does not form mixed disulfides. PECs containing ~90% diamide-crosslinked closed RNAP gave a uniform pause dwell time of ~20 s, consistent with the behavior of the elemental pause on the *his*P scaffold (Table 4-1; Figures 4-6A, B, and C).

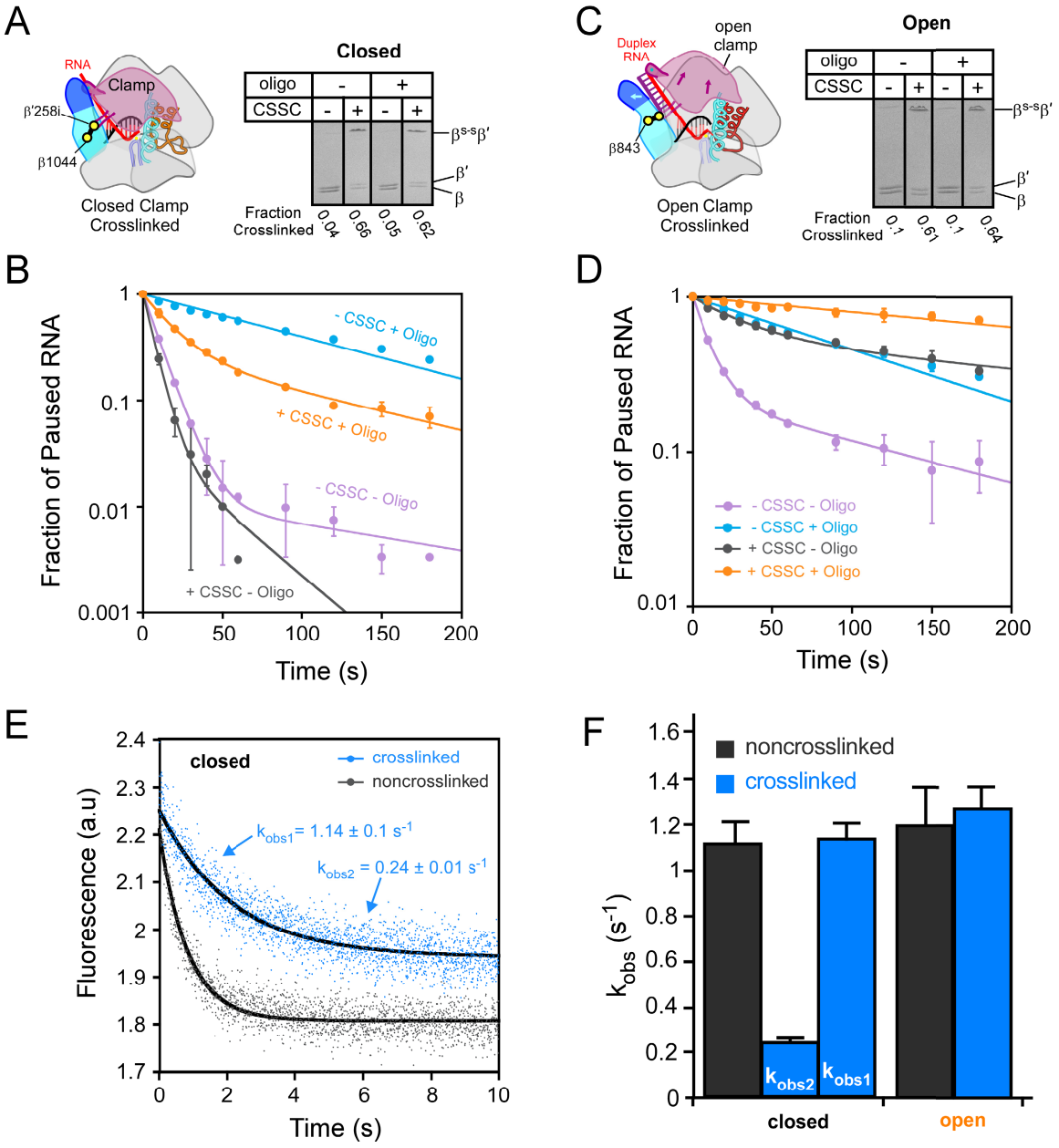


Figure 4-4. Effects of closed and open crosslinking on oligo-stabilized pausing and formation of paused RNA duplex in the RNA exit channel.

Figure 4-4 legend.

(A) Diagrammatic depiction of closed cys-pair crosslink, which stabilize the closed clamp conformation (left). SDS-PAGE of crosslinked or noncrosslinked samples. ECs (U_{19}) (250 nM) were assembled on a nucleic acid scaffold shown in Figure 4-1A that lacks PC under reducing conditions. Disulfide bond formation was initiated by addition of a mixture of 1 mM CSSC and 1 mM DTT. ECs were incubated with or without 5 μM 8-mer asRNA oligo. After quenching reactions with iodoacetamide, RNAP subunits were separated on a 4-15% precast non-reducing SDS-PAGE Phast gel (GE Biosciences) and silver-stained. Crosslinked fraction in each lane was determined and indicated (right and underneath).

(B) ECs (G_{17}) containing closed RNAP were formed and crosslinked with 1 mM CSSC and 0.8 mM DTT. G_{17} ECs were elongated to U_{19} by addition of 100 mM UTP and 2 μM [α-^{32}P]CTP. ECs halted at U_{19} position were extended to G_{21} in steps 1-6 (Figure 4-5C; Materials and Methods).The fraction of paused RNA at each was plotted as a function of reaction time. Pause escape rates (k_{pe}) were determined by fitting data points to either a single exponential (-CSSC + Oligo) or a double exponential equation (-CSSC –Oligo; -CSSC +Oligo; +CSSC +Oligo). Pause escape by crosslinked closed ECs in the presence of oligo was biphasic, and in each case gave a fast k_{pe} for a majority fraction (~70%), which reflects the crosslinked EC (fraction crosslinked ~68%) and a slower k_{pe} could be assigned to the noncrosslinked EC (See also Table 4-1).

(C) Representative time traces showing a single-exponential (noncrosslinked) and a double-exponential (crosslinked) decrease in fluorescence of PC upon addition of 5 μM 8mer asRNA oligo to 250 nM noncrosslinked and crosslinked closed ECs, respectively.

For crosslinked samples, ECs were treated with a mixture of 1 mM CSSC and 0.8 mM DTT (E_h = -0.24 V) giving the estimated crosslinked fraction of 68%.

Figure 4-4 legend (cont.)

(D) Diagrammatic depiction of open cys-pair crosslink, which stabilize the open clamp conformation (left). PAGE of crosslinked or noncrosslinked samples. The experiment was performed identically to that shown in Figure 4-4A except that the experiment was performed with open RNAP.

(E) ECs containing open RNAPs were formed, crosslinked, and elongated to G_{21} through the pause site U_{19} (Figure 4-5D; Materials and Methods). Fraction of pause RNA (U_{19}) was determined and plotted as a function of reaction time. The rate of escape, $k_{pe,}$ was obtained by fitting the disappearance of RNA 19 to a single exponential (for –CSSC +Oligo and +CSSC +Oligo) and to a double exponential equation (for –CSSC –Oligo and +CSSC -Oligo).

(F) The observed rate of the decrease in fluorescence of PC by binding of 5 μM 8mer asRNA oligo to the nascent RNA reflecting the duplex formation in ECs formed with closed or open RNAPs incubated with or without 1 mM CSSC and 0.8 mM DTT.

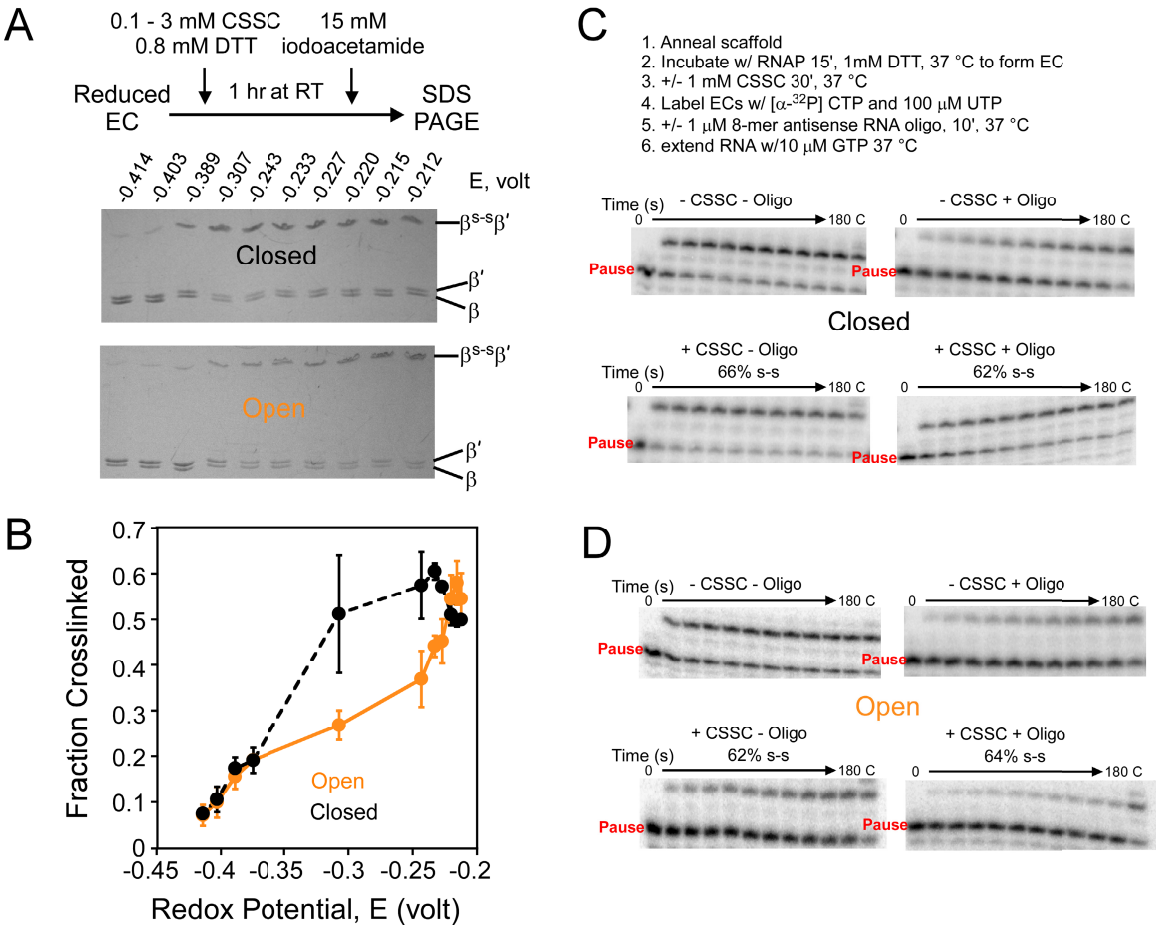


Figure 4-5. Formation of crosslinks using CSSC as an oxidant to stabilize open and closed clamp ECs and the effects of crosslinking on oligo-stabilized pausing.

(A) Crosslink formation of closed (top) and open (bottom) in ECs detected by non-reducing SDS-PAGE. Crosslinking was measured at different redox potentials generated with 0.8 mM DTT and varying the concentrations of CSSC.

(B) Crosslinked fraction of closed and open in ECs plotted as a function of redox potential. The fraction crosslinked was determined as the fraction of retarded-mobility $\beta^{S-S}\beta'$ relative to the total ($\beta + \beta' + \beta^{S-S}\beta'$).

Figure 4-5 legend (cont.)

(C) ECs containing closed RNAP were formed, crosslinked, and extended to G_{21} in steps 1-6 (Experimental Procedures). Prior to reaction with CSSC (-CSSC), ECs were reconstituted on the nucleic acid scaffold shown in Figure 4-1A. For crosslinked samples, ECs were oxidized with 1 mM CSSC for 30 min at 37 °C (+CSSC; ~68% crosslinked). G_{17} RNA was labeled and extended to U_{19} by addition of [α-^{32}P]CTP and 100 μM UTP. The complexes were then incubated with or without 1 μM 8mer antisense RNA oligo for 10 min at 37 °C. ECs were elongated through the *his* pause site (U_{19}) with 10 μM GTP. Samples were removed at 10, 20, 30, 40, 50, 60, 90, 120, 150, and 180 s and separated by denaturing PAGE. Remaining samples were incubated with 0.5 mM GTP and UTP each for an additional 3 min at the end of the reaction ("C" lane).

In contrast to the crosslinked closed PEC, the crosslinked open PEC appeared to be permanently trapped in a state similar to the hairpin-stabilized (or duplex-stabilized) PEC. If the effects of exit-channel duplexes on pausing are caused by clamp opening and the resultant effects on the active site (*e.g.*, inhibition of TL folding), then crosslinking the clamp in the open conformation should mimic these effects even in the absence of an exit-channel duplex. To investigate this prediction, we measured the pause durations of noncrosslinked and crosslinked PECs containing open RNAP in the absence and presence of asRNA. As expected, the asRNA increased pause dwell times of noncrosslinked PECs >10-fold (Table 4-1, 1 mM DTT). However, after formation of the open crosslink, ~60% of the PECs exhibited a long pause duration in the absence of asRNA that was comparable to the dwell time observed with noncrosslinked PECs in the presence of asRNA (Figures 4-4D and 4-5D). Further, addition of asRNA to the open crosslinked PECs converted the remaining fast fraction to a long-duration pause, presumably by acting on the uncrossinked fraction (Table 4-1, 1 mM CSSC). We confirmed these conclusions using diamide as an oxidant to generate larger fractions of crossinked PECs (Table 4-1; Figure 4-6D). Importantly, the crosslinked PECs were not dissociated or inactivated whether or not asRNA was bound because addition of high concentrations of NTPs at the end of the pause assay caused extension of essentially all the pause RNA (Figures 4-6C and D).

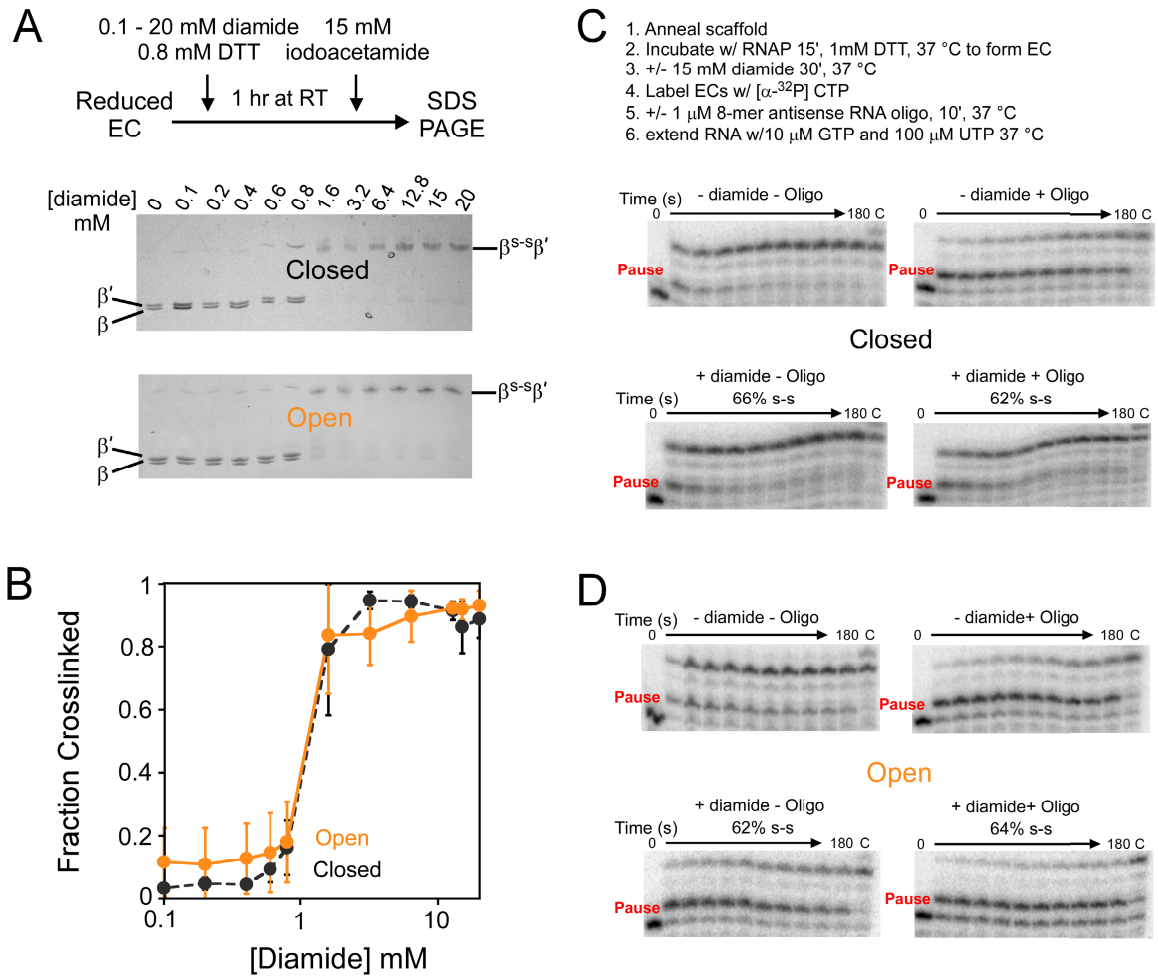


Figure 4-6. Formation of crosslinks using diamide as an oxidant to stabilize open and closed clamp ECs and the effects of crosslinking on oligo-stabilized pausing.

(A) Crosslink formation of closed (top) and open (bottom) in ECs detected by non-reducing SDS-PAGE. Crosslinking was measured at different concentrations of diamide generated with 0.8 mM DTT and varying the concentrations of diamide.

(B) Crosslinked fraction of closed and open in ECs as a function of diamide concentrations. The fraction crosslinked was determined as the fraction of retarded-mobility $\beta^{S\text{-}S}\beta'$ relative to the total ($\beta + \beta' + \beta^{S\text{-}S}\beta'$).

Figure 4-6 legend (cont.)

(C) ECs containing closed RNAP were formed, crosslinked, and extended to G_{21} in steps 1-6 (Experimental Procedures). Prior to reaction with diamide (-diamide), ECs were reconstituted on nucleic acid scaffold shown in Figure 4-1A. For crosslinked samples, ECs were oxidized with 15 mM diamide for 30 min at 37 °C (+diamide; ~85% crosslinked). G_{17} RNA was labeled and extended to C_{18} by addition of [α-^{32}P]CTP. The complexes were then incubated with or without 1 μM 8mer antisense RNA oligo for 10 min at 37 °C. ECs were elongated through the *his* pause site (U_{19}) with 10 μM GTP and 100 μM UTP. Samples were removed at 10, 20, 30, 40, 50, 60, 90, 120, 150, and 180 s and separated by denaturing PAGE. Remaining samples were incubated with 0.5 mM GTP and UTP each for an additional 3 min at the end of the reaction ("C" lane).

TABLE 4-1. Pause dwell time of ECs formed by closed or open RNAPs oxidized with either CSSC or diamide

		WT RNAP		closed crosslinked				open crosslinked	
Condition	asRNA	Dwell time (s)	fraction	Fraction xlinked	Dwell time (s)	fraction	Fraction xlinked	Dwell time (s)	fraction
1 mM DTT	-	10 ± 0.4	1	-	10 ± 0.1	0.98 ± 0.01	-	11 ± 0.4	0.77 ± 0.1
				0.04	175 ± 58	0.02 ± 0.01	0.1	160 ± 16	0.23 ± 0.1
	+	133 ± 7	1	0.05	108 ± 6	1	0.1	128 ± 6	1
1 mM CSSC + 0.8 mM DTT[a]	-	11 ± 0.1	1	0.66	6.6 ± 0.3	0.95 ± 0.1	0.61	312 ± 97	0.60 ± 1
				-	34 ± 11	0.05 ± 0.1	-	33 ± 10	0.40 ± 1
	+	160 ± 8	1	0.62	18 ± 1	0.72 ± 0.04	0.64	450 ± 32	1
				-	125 ± 15	0.28 ± 0.04	-		
Diamide[b]	-	29 ± 1	0.45 ± 0.01	0.86	19 ± 1	0.45 ± 0.01	0.85	322 ± 24	0.7 ± 0.01
	+	270 ± 7	0.68 ± 0.01	0.86	26 ± 1	0.44 ± 0.1	0.86	1500 ± 220	0.72 ± 0.01

The pause dwell-time was calculated from the rate of pause escape on the scaffold, which can form the 8-bp RNA:RNA duplex when the asRNA oligo is present, using *in vitro* transcription pause assay as described in Materials and Methods.

[a]ECs halted at the pause site (U_{19}) were treated with 1mM CSSC and 0.8 mM DTT (Redox Potential = ~0.243 V) with fraction crosslinked of ECs containing clsoed and open crosslinked RNAPs being ~0.7 and 0.61, respectively (Figures 4-4A and C). The fraction of complexes attributable to each species was obtained by fitting U_{19} pause RNA concentration as a function of time to a two-exponential decay equation.

[b]In these experiments, ECs halted at 2 nt before the pause (C_{18}) were treated with 15 mM diamide and 0.8 mM DTT and elongated through the pause site (U_{19}) on the *his* pause scaffold (Figure 4-1A). Fraction represents the pause efficiency (fraction of RNAPs that recognize the pause).

Taken together, our results demonstrate that much, and perhaps all, of the pause-stimulating effect of exit-channel duplexes can be explained by stabilization of the open clamp conformation by the duplex, with consequent effects on the RNAP active site. Blocking clamp opening with a disulfide crosslink blocks the effect of the duplex on pausing, whereas opening the clamp artificially using a disulfide crosslink instead of a duplex stimulates pausing to a level similar to the effect of the exit-channel duplex.

A closed clamp conformation inhibits but does not prevent formation of exit-channel duplexes

Suppression of duplex effects by the closed clamp crosslink (but not simulation of the effects by the open-clamp crosslink) might be explained if the closed clamp prevents formation of the exit-channel duplexes. To test whether clamp opening is required for duplex formation, we tested whether asRNA binding is inhibited in PECs containing the closed crosslink using the PC fluorescence-quenching assay as described above (Figures 4-4E and F). We first verified that noncrosslinked PECs containing the closed or open Cys substitutions bound asRNA comparably to wild-type PECs and found the Cys substitutions alone did not interfere with duplex formation (Figures. 4-3C and 4-4F; Table 4-2). We next determined the observed rates of asRNA binding to noncrosslinked ePECs and crosslinked closed ePECs (~55% crosslinked; Figure 4-4E; Table 4-2). We found that asRNA bound crosslinked clsoed ePECs more slowly and less efficiently than both noncrosslinked ePECs and wild-type PECs (Figure 4-4F; Table 4-2). For the noncrosslinked sample, we observed a rate of asRNA binding similar for wild-type ePECs ($1.2 \pm 0.2\ s^{-1}$); for the crosslinked sample, we observed two rate components of

about equal amplitude with the faster one similar to noncrosslinked (k_{obs1} = 1.14 ± 0.1 s^{-1}) and the slower one (k_{obs2} = 0.24± 0.01 s^{-1}; Figures. 4-4E and F; Table 4-2). To confirm that the slow component was attributable to the crosslinked ePEC, we varied the redox potential (CSSC concentration) to vary the crosslinked fraction. The fraction of slow species decreased when the crosslinked fraction decreased without affecting k_{obs2}, verifying that k_{obs2} resulted from asRNA binding to the crosslinked ePEC (Table 4-2). We also found that the amplitude of the fluorescence decrease (difference between initial and final fluorescence) was affected by crosslinking; the more highly crosslinked ePEC gave a lower amplitude (Table 4-2). This result is consistent with the observed rates, since 5 μM asRNA is predicted to bind only a fraction of crosslinked ePECs at these rates (K_d ~ 3.6 μM; a fraction of closed crosslinked PECs bound to asRNA ~ 0.58). To verify that differences in amplitude were due to incomplete asRNA binding to the crosslinked ePECs, we compared the amplitude of fluorescence change of noncrosslinked ePECs and crosslinked ePECs at 15 μM asRNA and found the amplitude difference largely disappeared (Table 4-2; again consistent with the calculated K_d). These results establish that the closed clamp is inhibitory for exit channel-duplex formation, but only by a factor of ~5 relative to RNAP in which the clamp can open or is trapped in the open conformation. Thus, the inability of exit-channel duplexes to stimulate pausing when the clamp is crosslinked in the closed conformation results from inability of the duplex to stimulate clamp opening, not from inability of the exit-channel duplex to form in the closed-clamp conformation.

TABLE 4-2. Effect of the RNAP lid-clamp cys-pair crosslinking to restrict clamp opening on the rate of the asRNA oligo binding to the nascent RNA

[oligo]	[CSSC] mM	crosslinked fraction[a]	k_{obs} (s^{-1})[b]	Fraction	Amplitude of fluorescence decrease[c]
5 μM	0	0.05	1.2 ± 0.02	1	0.33 ± 0.07
	0.6 (-0.375 V)	0.23	0.25 ± 0.02	0.17	0.28 ± 0.04
			1.14 ± 0.1	0.83	
	1 (-0.243 V)	0.56	0.24 ± 0.01	0.47	0.25 ± 0.03
			1.14 ± 0.1	0.53	
15 μM	0	0.05 ± 0.04	4.2 ± 0.1	1	0.31 ± 0.07
	0.4 (-0.389 V)	0.15 ± 0.05	0.9 ± 0.1	0.16	0.30 ± 0.07
			5 ± 0.2	0.84	
	1 (-0.243 V)	0.56 ± 0.06	1.25 ± 0.2	0.30	0.29 ± 0.07
			5 ± 0.5	0.70	

[a]Crosslinked fraction was determined from the quantification of the non-reducing SDS-PAGE samples as described in Materials and Methods. The concentration of EC in all experiments was 250 nM.

[b]The observed rate of fluorescence decrease upon binding of the asRNA oligo to the nascent RNA in noncrosslinked or crosslinked ECs with different concentrations of CSSC.

[c]Amplitude was calculated by subtracting the final fluorescence when the fluorescence signal plateau from the initial fluorescence before the addition of the asRNA.

RNAP active-site conformation and exit-channel conformation are energetically linked

The finding that the open-clamp crosslink alone, in the absence of an exit channel duplex, mimics duplex stimulation of pausing strongly supports the view that the RNAP exit channel and active site are energetically linked through clamp movements. To test this idea further, we next examined the effects of pause-suppressing substitutions near the RNAP active site on asRNA binding in the exit channel to test the energetic linkage in the opposite direction. In other words, if exit-channel-driven clamp opening can alter the conformation of the active site to promote pausing, can alterations of the active site that suppress pausing inhibit clamp opening in the exit channel? To address this question, we chose two previously identified substitutions known to inhibit pausing at the hairpin-stabilized *his* pause site: β´F773V in the N-terminal portion of the bridge helix (BH) and βT563I in the conserved βDloopII, which interacts with both the BH and nucleotides in the RNAP active site (Figure 4-7A). Both substitutions strongly suppress the *his* pause (Artsimovitch et al., 2003; Landick et al., 1990; Toulokhonov et al., 2007), possibly by altering BH conformation in ways that prevent formation of the paused TL conformation, which could affect clamp energetics (Nayak et al., 2013). T563I was originally described as part of the β P560S/T563I double substitution (Landick et al., 1990). To simplify analysis, we tested the single substitutions and determined that T563I alone exhibited the same effects seen previously with the βP560S/T563I substitutions (Figure 4-8A).

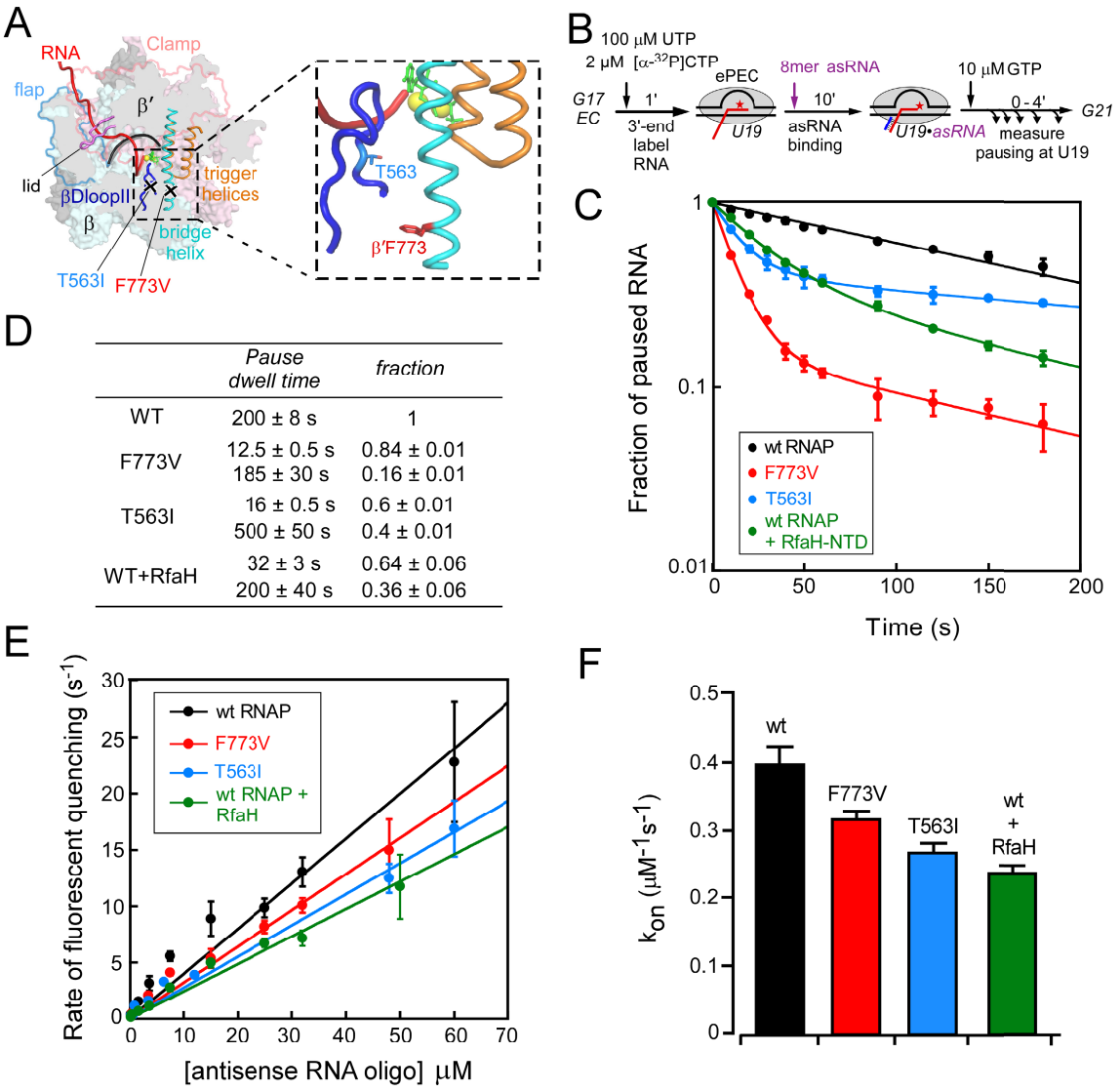
To test whether the F773V and T563I substitutions inhibit exit-channel duplex formation and clamp opening, we first verified the effect of the substitutions on asRNA-stimulated pausing on the ePEC (Figure 4- 1A). Consistent with earlier reports on hairpin-stabilized

pausing (Svetlov et al., 2007; Toulokhonov et al., 2007), both the F773V and T563I mutant RNAPs reduced the fraction of ePECs that convert to hairpin-stabilized PECs significantly (from 100% of ePECs for wild-type RNAP to 16% or 40% of ePECs for F773V and T563I, respectively; Figures. 4-7C and D). These results are consistent with reduced stability of the open-clamp conformation in the F773V and T563I RNAPs (Toulokhonov et al., 2007).

To determine how these substitutions affect exit-channel duplex formation, we measured the on-rate for asRNA oligo binding to ePECs reconstituted with T563I, F773V and wild-type RNAPs using the PC fluorescence assay (Figure 4-7E and Materials and Methods). Consistent with clamp-mediated energetic linkage between the active site and the exit channel, the F773V and T563I substitutions decreased the rate of asRNA oligo binding to the exiting RNA in ECs (Figures 4-7E and F). These effects were relatively modest, however, a little less than 2-fold. We tested whether the effects were greater at 15 °C, but found about the same relative magnitudes of binding rates (Figure 4-8B). Nonetheless, the results were reproducible and show that changes in the active site of RNAP, far from the exit-channel can affect formation of exit-channel duplexes. The simplest explanation is that the substitutions affect clamp conformation, which affects duplex formation through the linkage between the RNA exit channel and the active site.

To confirm that inhibiting clamp opening inhibits duplex formation and test for the biological relevance of this inhibition, we next tested whether the *E. coli* elongation regulator RfaH affects the rate of duplex formation using the PC fluorescence assay. RfaH is a NusG-like regulator, inhibits hairpin-stabilized pausing at the *his* pause site, and is thought to stabilize the closed-clamp conformation through contacts to the β′clamp

helices and β gate loop of RNAP (Sevostyanova et al., 2011). We recently reported that the N-terminal domain of RfaH (RfaH-NTD), which makes all the RNAP contacts and is sufficient for pause suppression, exhibits concentration-dependent competition with asRNA for opposing effects on pausing (Hein et al., 2011). This result predicts that RfaH-NTD inhibits asRNA binding. To test this prediction, we measured exit-channel duplex formation using the PC fluorescence assay in the presence of RfaH-NTD. Consistent with expectation, ePECs bound by RfaH-NTD exhibited a slower rate of exit-channel duplex formation (Figures 4-7E and F). Together, our results strongly support the model that pause-suppressing alterations to RNAP outside the exit channel (F773V, T563I, and RfaH-NTD) inhibit asRNA binding to the nascent RNA. These findings support the view that clamp conformation generates an energetic link between formation of secondary structures in the RNAP exit channel and active site conformation.



	Pause dwell time	*fraction*
WT	200 ± 8 s	1
F773V	12.5 ± 0.5 s 185 ± 30 s	0.84 ± 0.01 0.16 ± 0.01
T563I	16 ± 0.5 s 500 ± 50 s	0.6 ± 0.01 0.4 ± 0.01
WT+RfaH	32 ± 3 s 200 ± 40 s	0.64 ± 0.06 0.36 ± 0.06

Figure 4-7. Effects of RNAP substitutions and RfaH on oligo-mediated pausing and the exit-channel duplex formation.

Figure 4-7 legend.

(A) A structure of the *E. coli* EC based on *Tth*EC bound to NTP (pdb 2o5j)(Vassylyev et al., 2007b) showing the location of RNAP substitutions tested in this work. Clamp (magenta), bridge helix (cyan), trigger helices (orange), flap (light blue), flap-tip (dark blue), βDloop 2 (blue), the exiting RNA (red), RNA:DNA hybrid (red and black), active-site Mg^{2+} (yellow), and lid (dark pink) are shown. Most nucleic acids in the EC are omitted for clarity. The blow-up shows the location of residues T563 and F773 with respect to the active site magnesium and NTP (yellow and green, respectively).

(B) A schematic of the transcriptional pause assay. Preformed G_{17} ECs (50 nM) were elongated to the U_{19} position by addition of [α-^{32}P]-CTP and 100 μM UTP. The complexes were incubated with an 8-nt asRNA oligonucleotide (purple). The pause assay was initiated by addition of 10 μM GTP and aliquots were removed at the indicated times.

(C) The fraction of U_{19} pause RNA present was plotted as a function of reaction time. The rate of escape, k_{pe}, was determined by fitting the disappearance of pause RNA to a single exponential (for WT RNAP) or to a double exponential equation (for mutant RNAPs and wtRNAP + 250 nM RfaH-NTD).

(D). Pause dwell-times for wild-type, F773V, T563I, and wild-type with 250 nM RfaH-NTD PECs (50 nM) in the presence of 1 μM 8mer asRNA oligo (RNA-6598; Materials and Methods).

Figure 4-7 legend (cont.)

(E) The oligo-concentration dependence of the observed rate of PC fluorescence quenching (k_{obs}) for wild-type, F773V, or T563I RNAPs are plotted in comparison to values measured for ECs with 2.5 μM RfaH.

(F) A bar graph showing on-rates (k_{on}) of the 8mer asRNA oligo binding to the nascent RNA with the PC in 500 nM ECs containing wild-type, F773V, or T563I RNAPs, and wild-type RNAP in the presence of 2.5 μM RfaH-NTD.

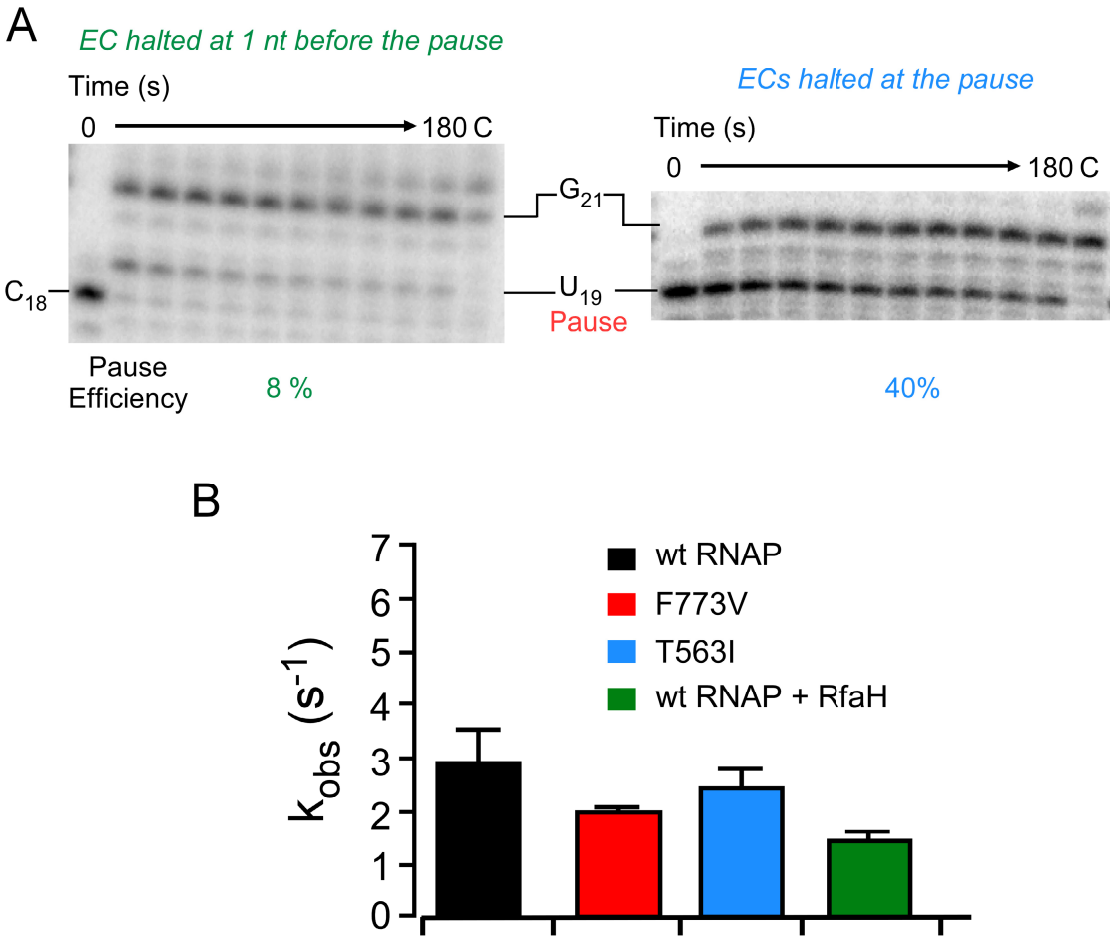


Figure 4-8. The exit-channel duplex formation measurements at lower temperature for different ECs and pause kinetics of T563I RNAP.

(A) ECs reconstituted with *E. coli* T563I RNAP were halted either at C_{18} position (left; green) or at U_{19} position (right; blue) and the pause efficiency (fraction of RNAPs that recognize the pause) is shown.

(B) The observed rates of the decrease in PC fluorescence (k_{obs}) upon addition of 27 μM 8mer antisense RNA oligo for ECs assembled on a scaffold (Figure 4-1A) with wild-type, F773V, or T563I RNAPs at 15 °C were shown in comparison to values measured for ECs with 2.5 μM RfaH-NTD.

Exit-channel duplexes inhibit forward translocation of RNAP through interaction with the flap tip

In principle, exit-channel duplexes could slow pause escape by two separate types of effects: inhibition of nucleic-acid translocation through RNAP or clamp comformation-mediated inhibition of TL folding required for catalysis (Nayak et al., 2013; Toulokhonov et al., 2007; Zhang et al., 2010). To determine the relative contributions of these two effects, we sought to measure the effect of exit-channel duplexes on translocation rate using a fluorescence reporter assay that relies on unquenching of the fluorescent G analog 6-methylisoxanthopterin (6-MI) upon translocation (Figure 4-9A(Malinen et al., 2012). We incorporated 6-MI in two different versions of the ePEC scaffold, either in the nontemplate DNA (ntDNA) strand at +2 position or in the template DNA (tDNA) strand at -8 position relative the pause RNA 3′ nt at -1 (Figure 4-9B). In both cases, we adjusted the sequence context so that translocation of the RNA 3′ U from the *i* site to the *i+1* site (required for substrate GTP binding) would melt a bp adjacent to 6-MI in which G was stacked on 6-MI, and leave 6-MI in a terminal bp stacked on a pyrimidine. This change in sequence context is thought to maximize the change in 6-MI fluorescence (Hawkins, 2008; Rist and Marino, 2002). We formed ECs one nt upstream from the pause site on the 6-MI-containing ePEC scaffolds (nascent RNA ending at C_{18}; Figure 4-9A). Upon addition of UTP and asRNA oligo, the PEC is formed. Forward translocation of RNAP exposes the *i+1* site for GTP binding and would unquench 6-MI fluorescence (Figure 4-9B). We determined the rate at which ePECs and PECs containing the exit-channel duplex increased 6-MI fluorescence (Figures. 4-9C, D, and E). To determine the contribution of the rate of U_{19} addition to the

rate of fluorescence increase and to deconvolute of the translocation rate from the U_{19} addition rate, we measured U_{19} addition using a quench-flow apparatus (Figure 4-10; Materials and Methods). After subtracting the contribution of U_{19} addition rate, two effects of the exit-channel duplex were evident. First, the exit-channel duplex reduced the rate of translocation (from ~30 s^{-1} to ~8 s^{-1}; Figures. 4-9C, D, E, and F; Table 4-3). Second, the exit-channel duplex reduced the amplitude of the fluorescence increase by 35-40% of that observed on the ePEC scaffold without the exit-channel duplex (Figures. 4-9C,D, and E). The simplest explanation for the decrease in signal amplitude is that most of the duplex-containing PEC remains in the pre-translocated register. This interpretation is consistent with multiple findings. A crystal structure of an ePEC formed on these same sequences using *Tth*RNAP is post-translocated (Weixlbaumer et al., 2013), whereas multiple lines of biochemical evidence establish that the duplex-stabilized PEC returns to the pre-translocated register and causes the RNA 3´ nt to fray off the DNA template (Toulokhonov et al., 2007; Weixlbaumer et al., 2013). We conclude that the exit-channel duplex significantly inhibits translocation of the PEC from the pre-translocated to the post-translocated register, by about a factor of 4 in the rate of forward translocation.

The RNAP flap-tip interacts with the *his* pause hairpin RNA and is required for duplex-enhancement of pausing even though it does not aid duplex formation (Figure 4-3(Toulokhonov et al., 2001; Toulokhonov and Landick, 2003). Based on these observations, we hypothesized that the interaction of the flap tip and exit-channel duplex might be required for duplex inhibition of translocation. Therefore, we tested the effect of exit-channel duplexes on translocation by ePECs containing flap-tip deletion RNAPs.

Consistent with the hypothesis, deletion of the flap tip abolished the exit channel-duplex effect on translocation (Figure 4-9F). In concert with the results described above, these findings indicate a dual function for flap-tip interaction with the exit-channel duplex: (1) promotion of clamp domain opening, with consequent effects on catalysis in the RNAP active site and (2) inhibition of translocation.

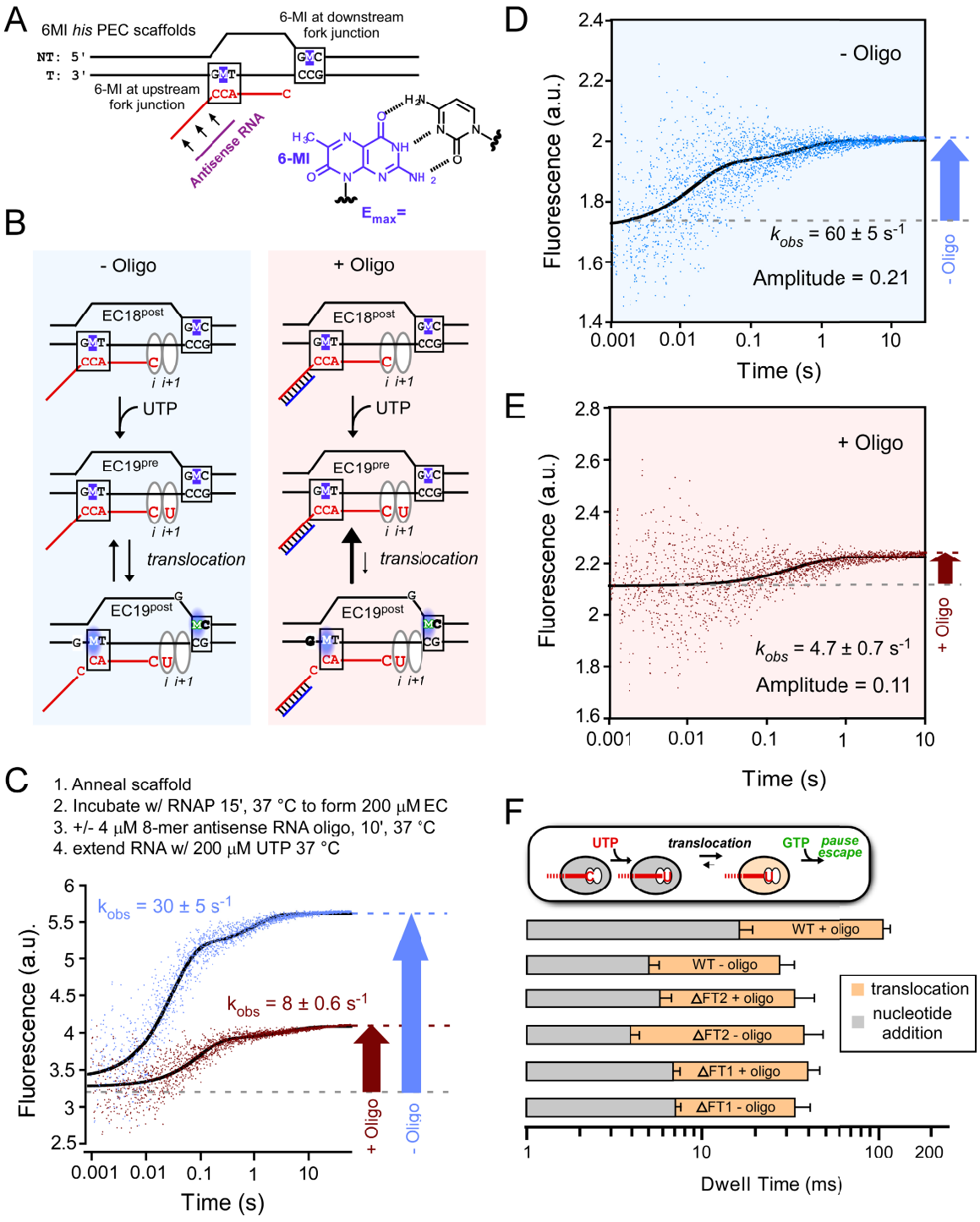


Figure 4-9. Oligo inhibits forward translocation and the flap-duplex interaction is required for translocation inhibition by the asRNA oligo.

Figure 4-9 legend.

(A) A representative *his* scaffold (Table 4-4) containing 6-MI (M; purple) incorporated at either at +2 position or in the template DNA (tDNA) strand at -8 position relative the RNA 3´ nt (CMP) at -1. This scaffold was used for reconstitution of ECs (yields EC_{C18}) and kinetic analysis of translocation following UMP incorporation. Guanine located adjacent to 6-MI serves as a quencher. DNA is shown in black, RNA in red, and 8-mer asRNA oligo in dark purple. The upstream and downstream fork junctions of the transcription bubble are shown.

(B) Inhibition of oligo on forward translocation of RNAP. Changes that occur in fork junctions of the transcription bubble and the increase in 6-MI fluorescence upon UTP addition and translocation in the absence (left; blue box) or presence (right; red box) of an 8-mer asRNA oligo are shown. RNAP's two active center sub-sites are indicated by gray ovals.

(C) Time-resolved measurements of an increase in 6-MI incorporated at +2 position (in the downstream fork junction) fluorescence upon UMP incorporation and subsequent RNAP translocation for ECs incubated with or without 4 μM 8mer asRNA oligo. The rate of 6-MI fluorescence increase, which represents the composite rate of UMP incorporation and translocation, is obtained by fitting the data points to a double exponential equation (Materials and Methods). The amplitudes of fluorescence change in the presence or absence of oligo are indicated by solid arrows.

Figure 4-9 legend (cont.)

(D,E) Time-resolved measurements of 6-MI (at -8 position in the upstream fork junction) fluorescence increase upon UMP incorporation and subsequent RNAP translocation for ECs in the absence (D) or the presence (E) of an 8mer asRNA oligo. The observed rate of the 6-MI fluorescence increase (k_{obs}), which represents the composite rate of UMP incorporation and translocation, is obtained by fitting the data points to a double exponential equation. The amplitudes of fluorescence change in the presence or absence of oligo were determined and indicated.

(F) Dwell time of nucleotide addition and translocation of wild-type and ΔFT RNAPs in the presence or absence of asRNA oligo (Materials and Methods).

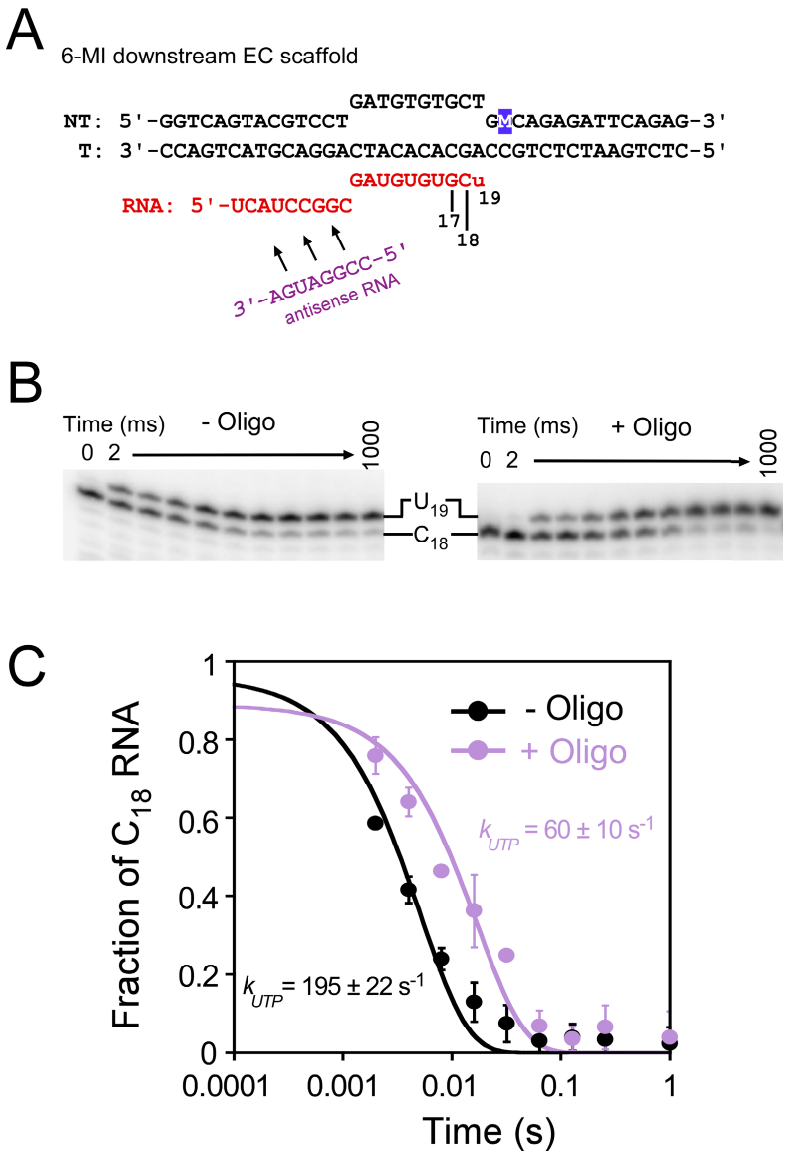


Figure 4-10. UTP addition rate measurements using the rapid-quench flow analysis.

Figure 4-10 legend.

(A) EC scaffold (NT-8244, T-8245, RNA-8246; Table S1) used for reconstitution of ECs (yields EC_{C18}) and kinetic analysis of translocation following UMP incorporation, which is denoted by a lowercase letter. 6-MI was incorporated at the -8 position in the tDNA, indicated by M, and shown as white-on-purple.

(B) Denaturing PAGE showing the UMP incorporation by ECs incubated with or without oligo. Samples were removed and quenched at 0.002, 0.004, 0.008, 0.016, 0.032, 0.064, 0.128, 0.256, and 1 s.

(C) The 18mer RNA (C_{18}) present in each lane in panel B was quantitated as a fraction of the total RNA in each lane and plotted as a function of reaction time. The rate of single round UTP addition, $k_{utp,}$ was obtained by fitting the disappearance of RNA 19 to a single exponential equation

TABLE 4-3. Effect of the 8mer asRNA oligo on the rate of translocation by ECs containing WT RNAP, WT RNAP and regulators, and mutant RNAPs

		Overall Rate[a] (s^{-1})	UTP addition rate[b] (s^{-1})	Translocation Rate (s^{-1})
WT RNAP	- oligo	30 ± 5	195 ± 22	31 ± 6
	+ oligo	8 ± 0.6	60 ± 10	7.2 ± 2
ΔFT-933 RNAP	- oligo	24 ± 5	137 ± 6	29 ± 6
	+ oligo	21 ± 3	148 ± 6	25 ± 4
ΔFT-942 RNAP	- oligo	24 ± 5	249 ± 27	27 ± 7
	+ oligo	25 ± 4	165 ± 27	29 ± 9
WT RNAP + RfaH	- oligo	27 ± 9	275 ± 34	30 ± 11
	+ oligo	14 ± 1.5	146 ± 21	16 ± 3
WT RNAP + NusA	- oligo	24 ± 7	198 ± 34	27 ± 10
	+ oligo	1.3 ± 0.3	7 ± 1.5	1.6 ± 0.7
WT RNAP reconstituted on a scaffold with longer dsDNA	- oligo	29 ± 9	151 ± 22	35 ± 14
	+ oligo	11 ± 2	70 ± 11	13 ± 4

[a]Overall rate represents the composite rate of translocation and UTP addition. The overall rate was calculated from the fluorescence increase of 6-MI incorporated at the +2 position in the ntDNA upon addition of UTP (Figure 4-9A).

[b]The UTP addition rates of different complexes were obtained from the rate of appearance of U_{19} RNA using a rapid-quench flow method as described in Materials and Methods. Errors are standard deviations from experimental triplicates.

Effect of regulators on translocation at the his pause site

Some transcriptional regulators have been proposed to affect translocation of RNAP, based on their effects on nucleotide addition and synergy with force effects on nucleotide addition in single-molecule force clamps. NusA reduces elongation rate, slows translocation, and prolong pausing at sites stimulated by nascent RNA structures (Ha et al., 2010; Zhou et al., 2011). RfaH, a paralog of NusG, accelerates elongation, inhibits backtracking, and weakly suppresses duplex-stabilized pausing (Figure 4-11A; (Herbert et al., 2010; Kolb et al., 2013). To test for direct effects of RfaH-NTD and NusA on forward translocation by RNAP, we measured translocation rates of ePECs with or without the exit-channel duplex, RfaH-NTD, or NusA using the 6-MI fluorescence assay. Consistent with previous results (Zhou et al., 2011), NusA significantly decreased the translocation rate, but only in the presence of asRNA (Figure 4-11B; Table 4-3). In contrast, RfaH increased translocation rate by decreasing duplex-inhibition of translocation (Figure. 4-11B; Table 4-3). These results confirm that NusA and RfaH affect translocation in ways that at least partially account for their synergy with exit-channel duplexes effects on transcriptional pausing (see Discussion).

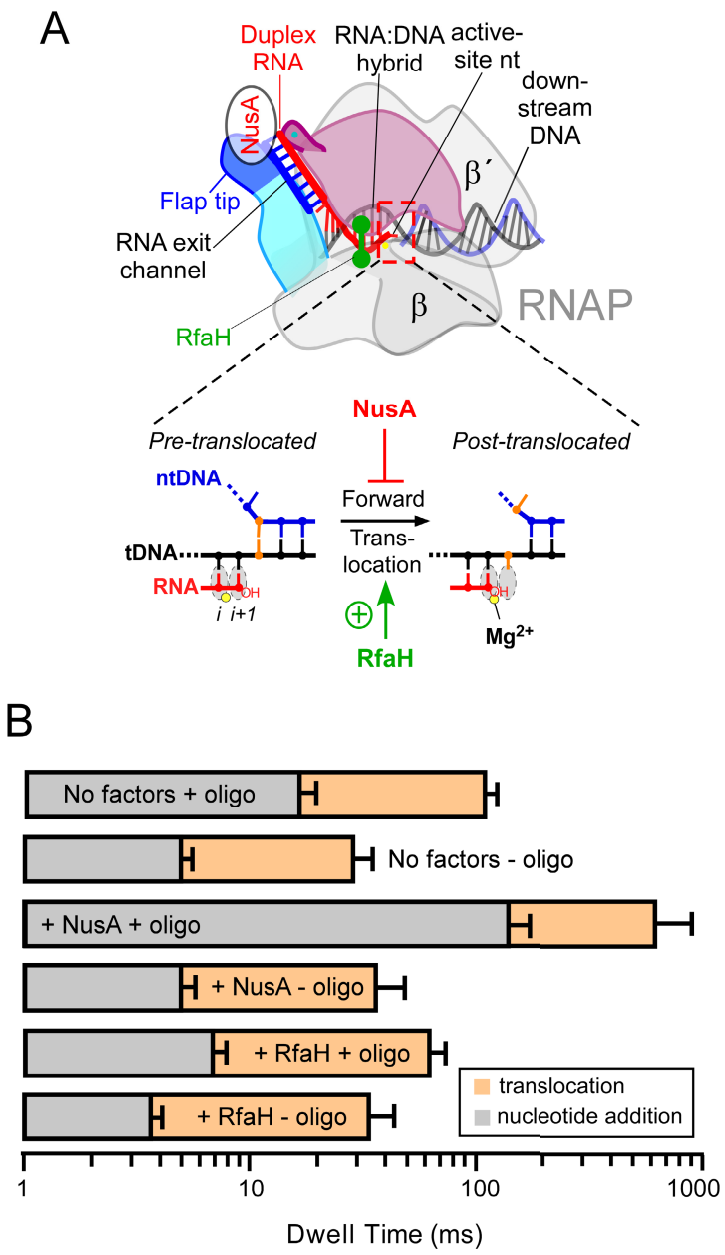


Figure 4-11. NusA inhibits forward translocation whereas RfaH-NTD stimulates translocation at the *his* pause site.

Figure 4-11 legend.

(A) Top: a model of the effects on translocation by elongation factors. RfaH is shown in green, NusA in red, RNAP in gray, flap-tip in dark blue, the flap domain in light blue, clamp in magenta, ntDNA in blue, tDNA in black, and RNA in red. Bottom: magnified view of the RNAP active site showing the forward translocation step, which is inhibited by NusA, but promoted by RfaH.

(B) Translocation rates were measured in the presence or absence of corresponding elongation factors, with or without 4 μM 8mer asRNA oligo. In experiments with elongation factors, ECs were incubated with 2 μM full-length NusA or 1 μM RfaH-NTD for 10 min at 37 °C prior to stopped-flow fluorescent or quench-flow experiments. Dwell time of nucleotide addition and translocation were determined from the composite rate and UMP incorporation rate obtained from stopped-flow and quench-flow experiments, respectively (Materials and Methods).

Discussion

Our results offer three new insights into the interplay between pausing by RNAP and formation of nascent RNA structures. First, although RNAP alone permits RNA structure formation without significantly altering duplex stability, the transcriptional machinery, in combination with regulators such as NusA and RfaH actively modulates the rate at which RNA duplexes form by altering the properties of the RNA exit channel. This modulation of structure formation depends in part on clamp domain conformation. Second, changes in the position of the clamp domain alter the properties of the RNAP active site even when they are not driven by the formation of RNA structures in the exit channel. Third, effects of exit-channel duplexes on nucleotide addition are mediated both by effects on the active site through the clamp and inhibition of translocation mediated by interactions of RNA:RNA duplexes with the flap domain.

The transcriptional machinery actively modulates RNA structure formation by altering RNA exit-channel conformation

The RNA exit channel of bacterial RNAPs is a crucial regulatory site for the enzyme, where multiple regulators appear to modulate formation and effects of nascent RNA structures. NusA can promote structure formation in the exit channel (Ha et al., 2010), but its effects can be reversed by antitermination proteins like λN or λQ that require NusA to suppress formation of terminator hairpins in the exit channel (Shankar et al., 2007). A key question is whether these effects work by controlling active stabilization of RNA structures by the RNAP exit channel, by modulating postulated inhibitory sites

on RNAP that repress RNA structure formation (Gusarov and Nudler, 2001), or by directly interacting with RNA structures.

Our results favor the direct interaction model for regulator modulation of structure formation with contributions of RNAP being driven by its interaction with regulators. We found that RNAP itself slowed duplex formation only 2-fold and did not stabilize exit-channel duplexes (Figure 4-1). Removing the flap tip domain, which potentially contacts RNA structures, increased rather than decreased the rate of duplex formation. Further, crosslinking the RNA exit channel in a closed conformation slowed duplex formation by only a factor of ~5 and crosslinking it in the open conformation did not significantly affect the rates of duplex formation (Figure 4-4).

Thus, the contribution of direct RNAP-RNA contacts on the rate of nascent RNA structure formation is modest. It can be increased by regulators like NusA that may directly stabilize structures (Figure 4-3) or like RfaH and NusG that may indirectly slow structure formation by stabilizing the closed clamp conformation (Figure 4-7). Although not tested here, we report elsewhere that NusG inhibits duplex-stimulated pause duration similarly to RfaH but requires much higher concentrations, suggesting that NusG has a weaker affinity than RfaH for the closed-clamp conformation of RNAP (Kolb et al., 2013). Thus the major effects of RNAP itself on RNA structure formation are likely kinetic via pauses that set kinetic windows for structure formation, with direct effects requiring additional regulators like NusA. This interpretation is consistent with observations that pausing alone contributes to proper RNA folding, but that the effects are much greater in the presence of NusA (Pan et al., 1999; Wickiser et al., 2005).

Clamp conformations energetically link the RNA exit channel and the active site

Previous results establishing effects of RNA structures on pausing have supported a mechanism involving clamp domain opening, but could not distinguish whether inhibition of nucleotide addition by exit-channel duplexes was caused by changes in the clamp conformation per se or by duplex-RNAP contacts that require clamp opening to allow duplex formation and establishment of the contacts. Our results clarify this picture in two ways. First, the effects of the duplex can be mimicked simply by crosslinking the clamp in an open conformation even when no duplex is present. Thus, the open clamp alone is responsible for inhibiting nucleotide addition in the RNAP active site. Importantly, we find the reverse is also true: altering the BH/TL module with amino-acid substitutions that suppress pausing decreases the rate of duplex formation in the exit channel, presumably by favoring a more closed clamp conformation (Figure 4-7). These findings strongly support the view that clamp opening alters the conformational dynamics of the BH/TL module and prevents TL folding in the hairpin-stabilized PEC (Nayak et al., 2013; Toulokhonov et al., 2007). It also raises the possibility that regulators that enhance pausing might do so simply by stabilizing an open clamp. An important unanswered question is the extent of clamp opening required to enhance pausing.

The second clarification is that the duplex can still form, albeit more slowly, even when opening of the clamp is blocked by a disulfide bond but cannot stimulate pausing. This result is evident both from the existence of a slower rate component of duplex formation when the closed clamp crosslink is present in a fraction of PECs (Figures. 4-4E and F) and a distinct component of the slightly slower pause escape in the crosslinked PECs whose relative contribution is correlated with the fraction of PECs trapped in the

closed-clamp state (Figure 4-4B; Table 4-1). Thus, the duplex can form and modestly slow pause escape but requires the associated clamp opening to slow pause escape by the factor of ~10 seen in wild-type PECs.

Exit-channel duplexes inhibit both translocation and TL folding via clamp opening

A crucial finding of this work is a direct demonstration that exit-channel duplexes inhibit translocation of PECs, but that an equal if not greater effect of the duplex is stabilization of the open-clamp conformation and its consequent effects on the active site. This conclusion can be drawn directly from the ~4-fold inhibition of translocation by the duplex. Since the overall inhibition of nucleotide addition by the duplex is ~10-fold, the fraction of the effect coming from the steps after translocation (NTP binding and catalysis) can be calculated from the reciprocal additivity of sequential rates ($1/k_{obs} = 1/k_1 + 1/k_2$, where k_1 would be the rate constant for translocation). Since the duplex decreases k_{obs} ≥10x and k_1 ~4x, it must decrease k_2 by ≥7x. This relationship may explain the limited effect of the duplex when the clamp is crosslinked shut, since the effect on translocation without an effect on the clamp may have only modest effect on the overall rate. Importantly the fact that translocation and catalysis occur on similar time-scales in the PEC means that changes in either translocation or catalysis can affect the overall rate of nucleotide addition, a conclusion supported by prior kinetic, biochemical, and single-molecule studies (Dangkulwanich et al., 2013; Malinen et al., 2012; Toulokhonov et al., 2007). Notably, it means that regulators can modulate transcript elongation and pausing by modulating any of the steps in the nucleotide addition cycle (translocation, NTP binding, or catalysis). Thus, the most significant component of the effect of NusA on

pausing appears to be inhibition of translocation (Figure 4-11). This balancing among the relative rates of steps in the nucleotide addition cycle may reflect evolutionary selection of enzyme parameters optimized for regulation. It also suggests that regulators could enhance pausing through direct stabilization of an open-clamp conformation, inverse to the effect of RfaH and NusG in suppressing pausing. To date, such pause-enhancing regulators acting though the clamp have not been identified.

Perspectives

Our results establish an active interplay between RNA structure formation in the RNAP exit channel and nucleotide addition in the active site of bacterial RNAP, mediated by clamp opening and modulated by elongation regulators. An important question is the extent to which the clamp domain mediates a similar interplay of nascent RNA structure formation and pausing in archaeal RNAP or eukaryotic RNAPII. The potential for clamp movements in the archaeal and eukaryotic enzymes appears to be restricted by the presence of RPB4/7-like subunits. However, it is unclear whether partial and sufficient clamp opening may be possible even when RPB4/7-like subunits are bound, or if RBP4/7 may dissociate to allow coupling in some situations (Mosley et al., 2013). The experimental approaches we have developed for bacterial RNAP should allow these questions to be addressed in both archaea and eukaryotes.

Materials and Methods

Reagents and oligonucleotides

All DNA and RNA oligonucleotides (Table 4-4) were obtained from IDT (Corvalville, IA) and purified by denaturing PAGE before use. [α-^{32}P]-CTP was from PerkinElmer Life Sciences, NTPs were from GE Healthcare (Piscataway, NJ); PC containing oligos, from TriLink Biotechnologies (San Diego, CA); 6-MI, from Fidelity Systems (Gaithersburg, MD); cystamine dihydrochloride (CSSC), from MP Biomedicals; DNA-modifying enzymes, NEB or Agilent; and DNA purification reagents, from Promega. All other reagents were from Sigma Aldrich or Fisher Scientific.

Construction and Purification of wildtype or variant *E. coli* RNAPs

RNAPs were expressed from pRM795, pRM843, or derivatives, in *E. coli* strain BLR λDE3 (Table 4-4). Plasmid pRM795 is a T7 RNAP-based overexpression plasmid containing the *E. coli rpoA*, *rpoZ*, *rpoB*, *rpoC* (PKA, His_{10}) that yields core RNAP ($\alpha_2\beta\beta'\omega$) with Protein Kinase A (PKA) and His_{10} tags on the C-terminus of β′. Plasmid pRM843 is also a T7 RNAP-based overexpression plasmid containing the *E. coli rpoA*, *rpoZ*, (His_{10})*rpoB*, *rpoC*(PKA, Strep) that yields core RNAP ($\alpha_2\beta\beta'\omega$) with His_{10} tag on the N-terminus of β and Protein Kinase A (PKA) and Strep tags on the C-terminus of β′. Wild-type and variant core *E. coli* RNAPs were purified from *E.* coli BLR IDE3 transformed with the appropriate pRM795 derivative plasmids (Table 4-4) as previously described (Nayak et al., 2013) with the following modifications to the purification protocol: upon elution of samples from a heparin sepharose column, protein-containing fractions were pooled and dialyzed against 2 L buffer B (100 mM Tris HCl [pH 7.9], 150 mM NaCl, 0.1 mM EDTA, 5 mM β-mercaptoethanol) for 4 hr at 4 ºC. Dialysed RNAP

was loaded onto a StrepTactin FPLC column (5 ml Strep-Tactin Superflow high capacity, iba; GE Pharmacia AKTA Purifier, GE Biosciences), washed with 25 ml buffer B, and eluted with buffer B + 2.5 mM D-desthiobiotin. Fractions containing purified RNAPs were then dialyzed into RNAP storage buffer (20 mM Tris-HCL [pH 8.0], 250 mM NaCl, 20 μM $ZnCl_2$, 1 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 50% glycerol) and stored at −80 ºC in aliquots.

His-tagged full-length NusA was purified from BL21 λDE3 cells containing a NusA overexpression plasmid by a two-step purification protocol using HisTrap HP affinity chromatography followed by gel filtration on Superdex 200 as described previously (Ha et al., 2010; Kolb et al., 2013; Kyzer et al., 2007). RfaH-NTD was purified from *E. coli* BL21 λDE3 cells harboring an overexpression plasmid, pIA777 (kindly provided by I. Artsimovitch) as described previously (Belogurov et al., 2007; Kolb et al., 2013). The plasmid encodes an RfaH derivative with a TEV protease cleavage site between the NTD and CTD and contains a C-terminal hexa-histidine tag.

Assembly of ECs

Nucleic acid scaffolds used for reconstituting ECs in *in vitro* transcriptional pause assays, stopped-flow fluorescent oligo binding measurements, and crosslinking assays were assembled as described previously (Kyzer et al., 2007). Briefly, 10 μM RNA, 12 μM tDNA, and 15 μM ntDNA (Table 4-4) were annealed in reconstitution buffer (RB; 10 mM Tris·HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$). Scaffolds were diluted with elongation buffer (EB; 25 mM HEPES-KOH, pH 8.0, 130 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.15 mM EDTA, 5% glycerol, and 25 μg of acetylated bovine serum albumin/ml) and incubated with RNAP for 15 min at 37 °C to form ECs used for

transcriptional pause assays and crosslinking assays, whereas ECs used for stopped-flow fluorescent oligo binding experiments were reconstituted in RB. For ECs used in crosslinking assays, RNAP was the limiting component at 1 μM RNAP, and RNA was used at 2 μM (1:2 RNAP: nucleic acid scaffold), whereas for ECs used in stopped-flow fluorescent oligo binding measurements, RNA was limiting at 500 nM, and RNAP was added at 1.5 μM (3:1 RNAP: nucleic acid scaffold). Similarly, ECs used for pause assays were preformed on a scaffold (T-5420; NT-5069; nM RNA-6593) with excess scaffold (150 nM RNA, 100 nM RNAP final) to ensure that almost every RNAP assembled into ECs.

ECs containing fully complementary DNA strands for translocation measurements with 6-MI were assembled using the modified reconstitution method described above by first annealing RNA and tDNA (2:1 RNA: tDNA) without ntDNA in RB buffer supplemented with 5 mM $MgCl_2$. Nucleic acid scaffolds lacking the ntDNA strand were incubated with RNAP (3:1 RNAP: scaffold) in EB buffer at 37 °C for 15min. The ntDNA was then added and further incubated at 37 °C for an additional 15 min. For ECs containing 6-MI at the ntDNA +2, ntDNA was limiting at 0.5 μM and added to 0.7 μM ECs lacking ntDNA, whereas, in ECs with 6-MI located in tDNA within RNA:DNA hybrid, tDNA was a limiting component at 500 nM, and RNA and tDNA are at 1 μM and 2 μM, respectively (tDNA: RNA: RNAP: ntDNA = 1: 2: 3: 4). In UMP incorporation measurements (quench-flow experiments), RNA was limiting at 50 nM, tDNA at 100nM, RNAP at 150 nM, and ntDNA at 200 nM (1:2:3:4 RNA: tDNA: RNAP: ntDNA).

Cys-pair disulfide bond crosslinking assay

ECs (1 μM) were reconstituted on a scaffold (6 μM T-5420; 8 μM NT-5069; 4 μM RNA-7418) with *E. coli* RNAP in EB buffer. Crosslinking of 1 μM ECs were measured over a range of redox potentials (from -0.212 to -0.414 V) using a mixture of 0.8 mM DTT and CSSC (0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.25, 1.5, 2, 2.5, 3 mM; Figure 4-3A) or 0.8 mM DTT and diamide (0.1, 0.2, 0.4, 0.6, 0.8, 1.6, 3.2, 6.4, 12.8, 15, 20 mM final; Figure 4-4A) as an oxidizing agent. These crosslinking experiments were performed at room temperature (RT) for 60 min. Samples were stopped by addition of iodoacetamide to 15 mM. Formation of Cys-pair crosslinks was evaluated by non-reducing SDS-PAGE (4-15% gradient Phastgel; GE) as described previously (Nayak et al., 2013). Gels were stained with Coomassie R and imaged with a CCD camera. Fraction crosslinked was quantified using ImageJ software. Experimental error was determined as the standard deviation in measurements from 3 independent replicates.

Oligo-mediated transcriptional pause assay

Transcriptional pause assays in the presence or absence of an 8-nt asRNA oligo were performed as described previously (Kolb et al., 2013). ECs were reconstituted 2 nt upstream from the *his* pause site at 50 nM scaffold (yields EC_{G17}; Figure 4-7B) with 200 nM core RNAPs, advanced to the U_{19} position by incubation with 2 μM [α-^{32}P]CTP and 100 μM UTP for 1 min at 37 °C in EB buffer. The complexes were then incubated with RB buffer (as control), or the 8mer antisense RNA oligonucleotide (1 μM final concentration) for 10 min at 37 °C. The pause assay was initiated by addition of 10 μM GTP to the U_{19} ECs. Reaction samples were removed at 10, 20, 30, 40, 50, 60, 90, 120, 150, 180 s and quenched with an equal volume of 2X stop buffer (8 M urea, 50 mM

EDTA, 90 mM Tris-borate buffer, pH 8.3, 0.02% bromophenol blue, and 0.02% xylene cyanol). After the final time point, all active ECs were chased out of the paused state with high concentrations of GTP and UTP (500 μM each).

For samples with elongation factors, ECs were incubated with 2.5 μM full-length NusA, or 250 nM RfaH-NTD, which were added immediately after [α-^{32}P]-CTP radiolabeling. In experiments testing the effects of cys-pair crosslinking on oligo-mediated pausing, ECs were reconstituted as described above followed by oxidization of RNAPs in the ECs with 1 mM CSSC and 0.8 mM DTT for 15 min at 37 °C. ECs were then radiolabeled, and the pause assay time course was conducted as described above. The reaction products were then separated on a denaturing 20% polyacrylamide gel in 0.5x TBE buffer (Hein et al., 2011). Gels were exposed to phosphorimager screens, scanned using a Typhoon PhosphorImager, and quantitated using the ImageQuant Software (GE Healthcare). The RNA present in each lane was quantitated as a fraction of the total RNA in each lane (Figures. 4-4B, D, and 4-7C) and corrected for the unreacted fraction remaining in the chase lane. The rates of escape from the *his* pause site (k) determined by single- or double-exponential fits of the fraction of $U_{19}/(>U_{19})$ as a function of time using KaleidaGraph (Synergy Software). Pause efficiencies or the fractions of EC in the one or two pause states (F) were estimated from the extrapolated intercept at time 0 as described previously (Kolb et al., 2013; Kyzer et al., 2007).

Stopped-flow fluorescence experiments

ECs (500 nM) assembled on a scaffold (1 μM T-5420; 1.5 μM NT-5069; 500 nM RNA-7604) containing PC RNA in RB was loaded in one syringe of the stopped-flow instrument (SF-300X, KinTek Corporation, Austin, TX, USA) and an 8mer asRNA oligo

(RNA-6598) was loaded in the other syringe. After allowing the two solutions to mix rapidly at 37 °C, PC fluorescence was measured in real time using 400-nm long-pass filter (Edmond optics) by exciting at 337 nm (4-nm bandwidth). The kinetics of fluorescence decrease was measured at various [asRNA] (*e.g.,* Figure 4-1E), and time traces were fit with a single exponential equation (Eq. 1):

$$F_t = F_0 + Ae^{-k_{obs}t} \qquad (1)$$

where F_0 and F_t are the initial and final fluorescence intensity of PC, respectively. A is the amplitude of the fluorescent change and k_{obs} is the pseudo-first-order observed rate constant of the decrease in PC fluorescence. To obtain a binding rate constant for the 8mer asRNA oligo, k_{obs} was determined at various [oligo] from 0.5 to 64 μM and plotted as a function of [oligo]. The resulting concentration dependence was fit to Eq. 2 and the bimolecular rate constant (on-rate; k_{on}) was calculated from the slope of the line (see also Figure 4-1C).

$$k_{obs} = k_{on}\,[oligo] + 0.1 \qquad (2)$$

The y-intercept value (0.1), which is equal to k_{off}, was obtained from off-rate measurements (See below). For crosslinked samples, disulfide formation was initiated by incubating ECs with 1 mM CSSC and 0.8 mM DTT for 15 min at 37 °C prior to mixing with the 8-mer asRNA oligo in the stopped-flow device.

To determine the off-rate of an 8mer asRNA oligo dissociating from the exit-channel RNA:RNA duplex in the EC, ECs were first reconstituted on a nucleic acid scaffold (1 μM T-5420; 1.5 μM NT-5069; 500 nM RNA-7604) in RB as described above. ECs were then incubated with 2 μM of an 8-mer asRNA oligo (RNA-6598) for 10 min at 37 °C, resulting in the formation of the exit-channel duplex and the quenching of PC

fluorescence. These quenched ECs containing the duplex were challenged with 100 μM competitor RNA (RNA-7728) in the stopped-flow instrument and an increase in fluorescent signal of PC was monitored as described above.

The resulting fluorescent traces were fit to Eq.3 and the dissociation rate of the 8-mer asRNA oligo from the duplex was calculated. The experiments were repeated in the absence of RNAP to determine the rate of dissociation of the oligo (k_{off}^{-RNAP}) from the duplex in the nucleic acid scaffold and to examine the effect of RNAP on the off-rate.

$$F_t = F_0 + A_1(1 - e^{-k_{obs}t}) \quad (3)$$

where F_0 and F_t are the initial and final fluorescence intensity of PC. A_1 is the amplitude of the fluorescent change, and k_{obs} indicates the increase in the rate constant of PC fluorescence upon addition of the competitor RNA.

To measure translocation of RNAP along the DNA, 500 nM ECs were reconstituted on a fully complementary nucleic acid scaffold where 6-MI was incorporated into either non-template DNA or template DNA (Figure 4-9A) at 37 °C for 15 min, followed by an additional incubation with ntDNA for 20 min. For samples with transcription elongation factors, 2 μM full length NusA or 1 μM RfaH-NTD was incubated with preformed ECs for an additional 10 min. ECs (20 μL) loaded in one syringe were allowed to mix rapidly with 400 μM UTP from the other syringe of the stopped-flow instrument at 37 °C (both ECs and UTP were in EB buffer supplemented with 5 mM $MgCl_2$). 6-MI fluorophore was excited at 340 nm and the fluorescence changes of 6-MI was monitored in real time in by recording emission through a 400-nm long pass filter. The resulting time traces were fit to the double exponential equation (Eq.4) to obtain the observed rate constant of the combination of the UTP addition and translocation.

$$F_t = F_0 + A_1(1 - e^{-k^1_{obs}t}) + A_2(1 - e^{-k^2_{obs}t}) \quad (4)$$

where F_0 and F_t are the initial and final fluorescence intensity of 6-MI. A_1 and A_2 represent the amplitude of the fluorescent change for the fast and slower phase, respectively. The fast and slower rate constants are denoted by k^1_{obs} and k^2_{obs}, respectively.

Steady-state fluorescence experiments

All study-state fluorescence measurements of PC fluorescence intensity were conducted on a PTI spectrofluorometer (Model QM-4/2003, Photon Technology International, Lawrenceville, NJ) using 10-mm path length and 50-μL quartz cuvettes (StarnaCells Inc, Atascadero, CA). Excitation spectra were obtained by recording fluorescence at 450 nm (8-nm bandwidth) while scanning excitation wavelengths between 300 and 400 nm (4-nm bandwidth), and emission spectra were obtained by monitoring fluorescence between 400 and 600 nm (8-nm bandwidth) and exciting at 337 nm (4-nm bandwidth) as described previously (Tinsley and Walter, 2006). Experiments were conducted with five variations in RB buffer at 25 °C: (*i*) with no oligonucleotide (RB solution); (*ii*) with 1 μM ssRNA, which contains PC (RNA-7604); or (*iii*) with dsRNA (to produce dsRNA, 1 μM of PC containing ssRNA, RNA-7604, was annealed to 2 μM of 8mer asRNA oligo, RNA-6598, for 10 min at 25 °C); (iv) with 2 μM noncomplementary RNA (1 μM of ssRNA was incubated with 2 μM noncomplementary RNA, RNA-6753, for 10 min at 25 °C); or (v) 1 μM nonfluorescent RNA (RNA-7418; Table 4-4). Equilibrium binding and dissociation assays were performed by manually mixing 250 nM ssRNA (in 100 μL RB) with 1 μM 8mer asRNA oligo (RNA-6598) or with 1 μM noncomplementary 8mer RNA oligo

(RNA-6753; as a control) in RB. After the fluorescence signal reached the equilibrium, 8mer competitor RNA oligo (RNA-7728) was added to a final concentration of 2 μM.

Rapid-quenched flow experiments

To obtain the rapid single-round UMP incorporation rate (Figure 4-8), ECs were reconstituted at 200 nM with scaffold lacking ntDNA (Figure 4-8A; yields EC_{G17}) and 600 nM core RNAP in EB buffer. After incubating ECs with 800 nM ntDNA, ECs were diluted to 100 nM in EB. Where indicated, 5 μM NusA and 250 nM RfaH-NTD were incubated with ECs for an additional 10 min prior to incubating ECs with 2 μM [α-^{32}P]-CTP for 1 min at 37 °C in EB buffer. The preformed ECs with 5′-^{32}P-labeled RNA from syringe A was rapidly mixed with 2 mM UTP (prepared in EB supplemented with 7 mM $MgCl_2$) from syringe B to initiate UMP addition. Samples were quenched with 2 M HCl using a KinTek rapid quench-flow instrument, neutralized to pH 8.0 by addition of 3 M Tris, phenol-extracted, ethanol-precipitated, resuspended in 1X stop buffer, and then electrophoresed through a denaturing 20% polycrylamide gels as described previously (Kyzer et al., 2007; Nayak et al., 2013). Gels were exposed to phosphor image screens, scanned with the Typhoon PhosphorImager, and quantified using ImageQuant software (GE Healthcare). The relative fractions of C_{18}/(>C_{18}) RNAs as a function of time were quantified by single fits (Supplementary Figures 4-8C and D).

TABLE 4-4. Strains, plasmids, and oligonucleotides

Stock #	Name	Description	Source or Note
Strains			
1620	BLR λDE3	*E. coli* B F^- λ (T7 gene 1) *ompT gal [dcm] [lon] hsdS$_B$(r_B^- m_B^-) recA$^-$* λDE3	Novagen
Plasmids			
2995	pRM795	Expresses wild-type *E. coli* RNAP(α2ββ′ω) with His_{10} and HMK tags on the β′ C-terminus	(Nayak etal., 2013)
5143	pRM843	Expresses wild-type *E. coli* RNAP(α2ββ′ω) with His_{10} and HMK tags on the β N-terminus and HMK-Strep on the β′ C-terminus	This work
5190	pRM890	Expresses β′ 258iC, β 843C *E. coli* RNAP(α2ββ′ω) with His_{10} and HMK tags on the β N-terminus and HMK-Strep on the β′ C-terminus	This work
5218	pRM918	Expresses β′ 258iC, β 1044C *E. coli* RNAP(α_2ββ′ω) with His_{10} and HMK tags on the β N-terminus and HMK-Strep on the β′ C-terminus	This work
5233	pRM933	Expresses βΔ(890–914)ΩGly_3 *E. coli* RNAP(α_2ββ′ω) with His_{10} and HMK tags on the β N-terminus and HMK-Strep on the β′ C-terminus	This work reference Kesha paper
5242	pRM942	Expresses (βΔ(890–914)ΩLeuValLeuGlu RNAP *E. coli* RNAP(α_2ββ′ω) with His_{10} and HMK tags on the β N-terminus and HMK-Strep on the β′ C-terminus	This work, reference Kuzdenelov
5174	pRM874	Expresses T563I *E. coli* RNAP(α_2ββ′ω) with His_{10} and HMK tags on the β N-terminus and HMK-Strep on the β′ C-terminus	This work
5165	pRM865	Expresses F773V *E. coli* RNAP(α_2ββ′ω) with His_{10} and HMK tags on the β′ C-terminus	This work
Oligonucleotides (5′→3′)			
5069	Non-template DNA	GGTCAGTACGTCCATTCGATCTTCGGAAGAGATTCAGAG	
5420	Template DNA	CTCTGAATCTCTTCCAGCACACATCAGGACGTACTGACC	
8111	*nt*DNA of *ds*6-MI scaffold	GGTCAGTACGTCCTGATGTGTGCTG**X**CAGAGATTCAGAG	
8103	*t*DNA for *ds*6-MI scaffold	CTCTGAATCTCTGCCAGCACACATCAGGACGTACTGACC	
8386	*nt*DNA of longer *ds*6-MI scaffold	GGTCAGTACGTCCTGATGTGTGCTG**X**AAGAGATTCAGAGTCTTCCAGTGGTGCATGAACG	
8387	*t*DNA of longer *ds*6-MI scaffold	CGTTCATGCACCACTGGAAGACTCTGAATCTCTTCCAGCACACATCAGGACGTACTGACC	
8245	*t*DNA for *us*6-MI scaffold	CTCTGAATCTCTTCCAGCACACT**X**GAGGACGTACTGACC	
8244	*nt*DNA for *us*6-MI scaffold	GGTCAGTACGTCCTccagtgtgctGGAAGAGATTCAGAG	

6598	RNA8 antisense RNA oligo	CCGGAUGA
6753	RNA8 non-complementary RNA oligo	UCGAGAGG
7728	RNA Competitor RNA	UCAUCCGG
6593	RNA17 (pause assay)	UCAUCCGGCGAUGUGUG
8107	RNA18 (*ds*6-MI)	UCAUCCGGCGAUGUGUGC
8246	RNA18 (*us*6-MI)	UCAUCCGGCCCAGUGUGC
7418	RNA19	UGCUUCAUCCGGCGAUGUGUGCU
7604	RNA19 (PC RNA)	UCAUZCGGCGAUGUGUGCU

X = 6-MI; Z = PC

References

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

Belogurov, G.A., Vassylyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassylyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. Molecular cell *26*, 117-129.

Berry, D.A., Jung, K.Y., Wise, D.S., Sercel, A.D., Pearson, W.H., Mackie, H., Randolph, J.B., and Somers, R.L. (2004). Pyrrolo-dC and pyrrolo-C: Fluorescent analogs of cytidine and 2¢-deoxycytidine for the study of oligonucleotides. Tetrahedron Lett *45*, 2457-2461.

Braunlin, W.H., and Bloomfield, V.A. (1991). 1H NMR study of the base-pairing reactions of d(GGAATTCC): salt effects on the equilibria and kinetics of strand association. Biochemistry *30*, 754-758.

Brehm, S.L., and Cech, T.R. (1983). Fate of an intervening sequence ribonucleic acid: excision and cyclization of the Tetrahymena ribosomal ribonucleic acid intervening sequence in vivo. Biochemistry *22*, 2390-2397.

Chan, C.L., Wang, D., and Landick, R. (1997). Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. Journal of molecular biology *268*, 54-68.

Cisse, II, Kim, H., and Ha, T. (2012). A rule of seven in Watson-Crick base-pairing of mismatched sequences. Nat Struct Mol Biol *19*, 623-627.

Dangkulwanich, M., Ishibashi, T., Liu, S., Kireeva, M.L., Lubkowska, L., Kashlev, M., and Bustamante, C.J. (2013). Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. eLife *2*, e00971.

Dash, C., Rausch, J.W., and Le Grice, S.F. (2004). Using pyrrolo-deoxycytosine to probe RNA/DNA hybrids containing the human immunodeficiency virus type-1 3' polypurine tract. Nucleic acids research *32*, 1539-1547.

Gusarov, I., and Nudler, E. (2001). Control of intrinsic transcription termination by N and NusA: the basic mechanisms. Cell *107*, 437-449.

Ha, K.S., Toulokhonov, I., Vassylyev, D.G., and Landick, R. (2010). The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. Journal of molecular biology *401*, 708-725.

Hawkins, M.E. (2008). Fluorescent pteridine probes for nucleic acid analysis. Methods in enzymology *450*, 201-231.

Hein, P.P., Palangat, M., and Landick, R. (2011). RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. Biochemistry *50*, 7002-7014.

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.

Herbert, K.M., Zhou, J., Mooney, R.A., Porta, A.L., Landick, R., and Block, S.M. (2010). E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. Journal of molecular biology *399*, 17-30.

Johnson, N.P., Baase, W.A., and von Hippel, P.H. (2005). Investigating local conformations of double-stranded DNA by low-energy circular dichroism of pyrrolo-cytosine. Proc Natl Acad Sci U S A *102*, 7169-7173.

Kinjo, M., and Rigler, R. (1995). Ultrasensitive hybridization analysis using fluorescence correlation spectroscopy. Nucleic acids research *23*, 1795-1799.

Kolb, K., Hein, P., and Landick, R. (2013). Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit-channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. J Biol Chem *in press*.

Kuznedelov, K.D., Komissarova, N.V., and Severinov, K.V. (2006). The role of the bacterial RNA polymerase beta subunit flexible flap domain in transcription termination. Dokl Biochem Biophys *410*, 263-266.

Kyzer, S., Ha, K.S., Landick, R., and Palangat, M. (2007). Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. J Biol Chem *282*, 19020-19028.

Lai, D., Proctor, J.R., and Meyer, I.M. (2013). On the importance of cotranscriptional RNA structure formation. RNA *19*, 1461-1473.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.

Landick, R., Stewart, J., and Lee, D.N. (1990). Amino acid changes in conserved regions of the beta-subunit of Escherichia coli RNA polymerase alter transcription pausing and termination. Genes Dev *4*, 1623-1636.

Lewicki, B.T., Margus, T., Remme, J., and Nierhaus, K.H. (1993). Coupling of rRNA transcription and ribosomal assembly in vivo. Formation of active ribosomal subunits in Escherichia coli requires transcription of rRNA genes by host RNA polymerase which cannot be replaced by bacteriophage T7 RNA polymerase. Journal of molecular biology *231*, 581-593.

Liu, C., and Martin, C.T. (2001). Fluorescence characterization of the transcription bubble in elongation complexes of T7 RNA polymerase. Journal of molecular biology *308*, 465-475.

Liu, C., and Martin, C.T. (2002). Promoter clearance by T7 RNA polymerase. Initial bubble collapse and transcript dissociation monitored by base analog fluorescence. J Biol Chem *277*, 2725-2731.

Malinen, A.M., Turtola, M., Parthiban, M., Vainonen, L., Johnson, M.S., and Belogurov, G.A. (2012). Active site opening and closure control translocation of multisubunit RNA polymerase. Nucleic acids research *40*, 7442-7451.

Matysiak, M., Wrzesinski, J., and Ciesiolka, J. (1999). Sequential folding of the genomic ribozyme of the hepatitis delta virus: structural analysis of RNA transcription intermediates. Journal of molecular biology *291*, 283-294.

Mosley, A.L., Hunter, G.O., Sardiu, M.E., Smolle, M., Workman, J.L., Florens, L., and Washburn, M.P. (2013). Quantitative proteomics demonstrates that the RNA polymerase II subunits Rpb4 and Rpb7 dissociate during transcriptional elongation. Molecular & cellular proteomics : MCP *12*, 1530-1538.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. Molecular cell *50*, 882-893.

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

Nudler, E. (2012). RNA polymerase backtracking in gene regulation and genome instability. Cell *149*, 1438-1445.

Nussinov, R., and Tinoco, I., Jr. (1981). Sequential folding of a messenger RNA molecule. Journal of molecular biology *151*, 519-533.

Pan, T., Artsimovitch, I., Fang, X.W., Landick, R., and Sosnick, T.R. (1999). Folding of a large ribozyme during transcription and the effect of the elongation factor NusA. Proc Natl Acad Sci U S A *96*, 9545-9550.

Pan, T., and Sosnick, T. (2006). RNA folding during transcription. Annu Rev Biophys Biomol Struct *35*, 161-175.

Perdrizet, G.A., 2nd, Artsimovitch, I., Furman, R., Sosnick, T.R., and Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. Proc Natl Acad Sci U S A *109*, 3323-3328.

Rist, M.J., and Marino, J.P. (2002). Fluorescent Nucleotide Base Analogs as Probes of Nucleic Acid Structure, Dynamics and Interactions. Current Organic Chemistry *6*, 775-793.

Sevostyanova, A., Belogurov, G.A., Mooney, R.A., Landick, R., and Artsimovitch, I. (2011). The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Molecular cell *43*, 253-262.

Shankar, S., Hatoum, A., and Roberts, J.W. (2007). A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. Molecular cell *27*, 914-927.

Svetlov, V., Belogurov, G.A., Shabrova, E., Vassylyev, D.G., and Artsimovitch, I. (2007). Allosteric control of the RNA polymerase by the elongation factor RfaH. Nucleic acids research *35*, 5694-5705.

Tinsley, R.A., and Walter, N.G. (2006). Pyrrolo-C as a fluorescent probe for monitoring RNA secondary structure formation. RNA *12*, 522-529.

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

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

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. Molecular cell *27*, 406-419.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., and Artsimovitch, I. (2007a). 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. (2007b). Structural basis for substrate loading in bacterial RNA polymerase. Nature *448*, 163-168.

Weixlbaumer, A., Leon, K., Landick, R., and Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. Cell *152*, 431-441.

Wetmur, J.G. (1991). DNA probes: applications of the principles of nucleic acid hybridization. Crit Rev Biochem Mol Biol *26*, 227-259.

Wickiser, J.K., Winkler, W.C., Breaker, R.R., and Crothers, D.M. (2005). The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. Molecular cell *18*, 49-60.

Wong, T.N., and Pan, T. (2009). RNA folding during transcription: protocols and studies. Methods in enzymology *468*, 167-193.

Yakhnin, A.V., and Babitzke, P. (2010). Mechanism of NusG-stimulated pausing, hairpin-dependent pause site selection and intrinsic termination at overlapping pause and termination sites in the Bacillus subtilis trp leader. Mol Microbiol *76*, 690-705.

Zamft, B., Bintu, L., Ishibashi, T., and Bustamante, C. (2012). Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. Proc Natl Acad Sci U S A *109*, 8948-8953.

Zhang, J., Palangat, M., and Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. Nat Struct Mol Biol *17*, 99-104.

Zhou, J., Ha, K.S., La Porta, A., Landick, R., and Block, S.M. (2011). Applied force provides insight into transcriptional pausing and its modulation by transcription factor NusA. Molecular cell *44*, 635-646.

Chapter Five

Conclusions and future directions

Conclusions and significance of this thesis work

The RNA polymerase (RNAP) structure/function studies presented in the preceding chapters of this thesis have advanced our understanding of gene regulation, with impacts for prokaryotic and eukaryotic molecular biology and biotechnology. These results help define the central roles of RNA polymerase translocation and transcriptional pausing in the regulation of transcription elongation. Many processes contribute to the remarkable control of elongation complex (EC) activity, but at the heart of these events lay conformational changes in EC structure that inhibit pausing and promote efficient elongation. By elucidating the detailed mechanisms that govern catalysis, translocation, and transcriptional pausing, these studies provide a foundation for the broad understanding of gene regulation and knowledge that will advance the mechanistic understanding of central processes involving transcription elongation, pausing, and RNAP translocation.

There is always a growing need for new medicines effective against drug-resistant pathogens. Our work significantly impacts the quest for new and innovative drugs. The mechanisms elucidated in this dissertation will aid in identifying and characterizing inhibitors and contribute to drug designs that target RNAP or other enzymes (*e.g.,* HIV reverse transcriptase) that share similar mechanisms.

We first determined the details of translocation, which contributes to overall catalysis. We provided evidence that the sequence at the 3′ end of RNA governs forward translocation of *E. coli* RNAP (chapter two). Prior to this study, this central step in the nucleotide addition cycle was poorly understood.

The central role of transcriptional pausing in both bacterial and eukaryotic gene regulation has long been appreciated. We demonstrated that the pause prolonging effect of the *his* pause hairpin can be mimicked by annealing a complementary oligonucleotide (oligo) to the nascent RNA. We also reported that RNA exit-channel specificity of RNAP requires at least 8-bp RNA:RNA duplexes for maximum enhancement of pausing and that RNA:RNA duplexes were more effective than DNA:RNA (chapter three). Our work elucidated the mechanistic contributions of elongation regulators such as NusA and RfaH to pausing. In the presence of NusA, the optimal exit-channel duplex length increased to 10 bp and the ability of DNA:RNA to mimic RNA:RNA was significantly reduced. RfaH, which binds to the RNAP clamp and suppresses oligo-stabilization of pausing, exhibited a concentration-dependent competition with oligonucleotides for effect, suggesting that RfaH and exit-channel duplexes affect RNAP clamp conformation in opposite directions (chapter three).

Finally, we discovered regulatory communication between the active site and the nascent RNA exit channel, and determined the effect of RNAP and RfaH on RNA structure formation (chapter four). We presented how the configuration of the active site affects structure formation, and further demonstrated that the formation of RNA secondary structures inhibits translocation at the *his* pause site (chapter four). We showed that elongation regulators such as NusA and RfaH, affect transcriptional pausing by modulating the translocation of RNAP at a pause site in addition to modulating RNA structure formation and the folding dynamics of the trigger loop (chapter four).

Future directions

Our studies presented in the preceding chapters lay a foundation for the further investigation of RNA polymerase structure/function, specifically more detailed investigation of the conformational changes in RNAP domains that occur during transcription. Future studies may address the following important questions:

1. How can we obtain direct biochemical measurements of the dynamics of clamp movement? How do difference sequences in the nucleic acids or interactions of regulatory signals with RNAP affect the clamp movements?

Movement of the RNAP clamp has been proposed to occur during pausing and stabilize the pause state. In one proposed model the clamp moves as the RNAP enters the paused state (Weixlbaumer et al., 2013). Crystal structures of bacterial ECs reveal the existence of two states (open and closed clamp conformation) which ECs can adopt during transcription elongation (Tagami et al., 2010; Vassylyev et al., 2007; Weixlbaumer et al., 2013). A single-molecule fluorescence resonance energy transfer (FRET) approach that detects the RNAP clamp movements in different transcription complexes by measuring FRET efficiency between fluorophores in RNAP's pincers supports the idea that the clamp can adopt different conformational states (Chakraborty et al., 2012). An alternative approach to single-molecule FRET is to measure competitive disulfide formation between strategically placed cysteine (Cys) residues that would be favored in either the open or closed conformation of RNAP (chapter four). The formation of disulfide bonds is a well-validated approach to detect protein conformations (Darby and Creighton, 1997), and has been proven effective in probing the folding dynamics of

the trigger loop in *E.coli* RNAP due to the minimal perturbation created by introduction of Cys side chains (Nayak et al., 2013). We designed a triple Cys reporter (TCR) system with one Cys in the β′ lid domain (Cys is inserted between two Gly residues: 258 and 259 of the *E.coli* β′ subunit; β′ 258i) that can crosslink to either a second Cys-substitution in the β flap region at β P1044C or a third Cys at β T843C position. Monitoring the crosslinking of β1044/β′258i (favored in the closed-clamp) or of β843/β′258i (favored in the open-clamp) reports the clamp conformation and thus allows determination of relative clamp conformation (Figure 5-1A). Hereafter, I refer to closed and open Cys-crosslinks as C (β1044/β′258i) and O (β843/β′258i), respectively.

We have successfully purified *E.coli* RNAP containing this TCR using our RNAP overexpression/purification system (Artsimovitch et al., 2003; Toulokhonov et al., 2007). Crosslinking of C (β1044/β′258i) and O (β843/β′258i) in the purified RNAP with TCR system has been detected successfully by SDS-PAGE (Figure 5-1B; β-β′ disulfides; Nayak et al., 2013). Importantly, O (β843/β′258i) crosslinked species migrates slower than C (β1044/β′258i) crosslinked species, so this system should allow the straight-forward monitoring of the clamp positions in different ECs.

We will next determine the competition between C (β1044/β′258i) and O (β843/β′258i) for disulfide formation in ECs, ECs bound to elongation regulators, ECs formed with mutant RNAPs described in preceding chapters, hairpin-stabilized pause elongation complexes (PECs), and backtracked PECs to learn how the clamp conformations differ among ECs and PECs. We showed that the pause hairpin effect can be mimicked by annealing of the 8-mer antisense RNA oligo to the exiting RNA of ECs (chapter three). It is possible that the clamp would become more open in the presence of

the antisense RNA oligo because of the formation of the RNA:RNA duplex in the RNA exit channel of RNAP; we can test that with this system.

The view that the clamp undergoes conformational changes during transcriptional pausing is well supported by crystallographic studies and our biochemical evidence described in chapter four (Weixlbaumer et al., 2013). However, it is unclear how changes in the nucleic acid sequences that contacts RNAP affect clamp position. For instance, sequences in downstream DNA have been reported to play an important role at the *his* pause site through contacting RNAP clamp and jaw domains (Ederth et al., 2006; Lee et al., 1990; Palangat et al., 2004). It is possible that alterations in downstream DNA sequences that are predicted to decrease pausing could potentially favor the closed-clamp positions. We will test this prediction by systematically changing the downstream DNA sequences and probing the clamp positions with our TCR system. Similarly, we can ask how sequences in the RNA:DNA hybrid of the EC influence the clamp conformations. We can also ask if translocation state of ECs affects the clamp dynamics because the clamp is proposed to loosen up some of its contacts to the nucleic acids and transiently open during translocation step (Weixlbaumer et al., 2013). To test whether the clamp movement is affected by translocation states of ECs, we will form ECs that favor either pre-translocated state or post-translocated state by changing the sequence at the 3′ end of RNA (based on our findings reported in chapter two) and probe the clamp position. The findings from these experiments will unambiguously test the idea that the RNAP clamp adopts different conformations during transcription elongation and that clamp positions respond to alterations in RNAP and in nucleic acid scaffold, RNAP pausing, or regulators' binding.

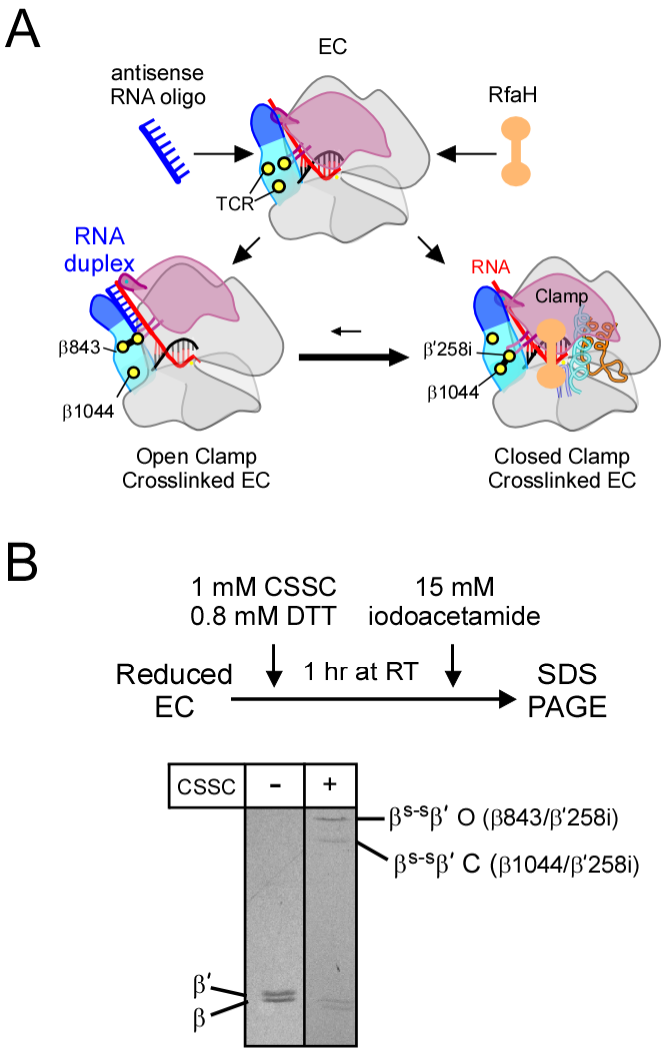


Figure 5-1. A triple-cysteine reporter system (TCR) to monitor the clamp dynamics

(A) Diagrammatic drawings representing TCR show closed-clamp crosslinked (C (β1044/β′258i); bottom left) and open-clamp crosslinked (O (β843/β′258i); bottom right) ECs. Antisense oligo binding and duplex formation stabilizes open-clamp conformation whereas RfaH binding favors closed-clamp position.

(B) Crosslink formation of C (β1044/β′258i) and O (β843/β′258i) in ECs detected by non-reducing SDS-PAGE.

2. What is the consequence of trapping the clamp in the open- or closed-conformation permanently? Do these changes affect the rapid elongation rate, translocation, and transcription factors binding?

In chapter four, I showed that crosslinking of C (β1044/β′258i) decreases oligo-mediated pausing and hairpin-mimic oligo binding whereas crosslinking of O (β843/β′258i) enhances pausing and results in a long-lived PEC even in the absence of pause hairpin-mimic oligo. Therefore, it would be interesting to ask if locking the clamp either open or closed would allow translocation, elongation, or binding of regulators. We will use the same approach described in chapter four to test the effect of blocking the clamp opening by crosslinking the lid to the flap to stabilize either the closed-clamp or open-clamp conformations.

(1) To test whether the clamp opening is required for translocation, we will measure translocation rates of crosslinked ECs containing either C (β1044/β′258i) or O (β843/β′258i) RNAPs by using the fluorescent-based translocation assay described in chapter four (Malinen et al., 2012). For these experiments, we will incorporate a fluorescent base analog, 6-methylisoxanthopterin (6-MI), either in the nontemplate DNA (ntDNA) strand at +2 position (Figure 5-2A) or in the template DNA (tDNA) strand at -8 position to detect structural changes (Figure 4-9). The 6-MI analog is unlikely to perturb RNA structure and translocation bias due to its small size and structural similarity to guanine (Hawkins, 2008; Rist and Marino, 2002) and hence it is suitable for monitoring changes in nucleic acids upon RNAP translocation through DNA strands. Initially, the fluorescence of 6MI is quenched through base-stacking interactions with surrounding nucleotides especially with a neighboring Guanine nucleotide (GMP), which acts as a

quencher. The fluorescent signal of 6MI increases upon addition of UTP to the 3′ end of RNA and subsequent translocation of RNAP, separating the quencher GMP from the 6MI (Figure 4-9A and B).

(2) We will next investigate transcription elongation of crosslinked C (β1044/β′258i) or O (β843/β′258i) RNAPs and compare their rates of elongation to that of WT RNAP. Additionally, we can test the effects of restricting clamp movement on binding and activity of regulators by repeating experiments in the presence of different concentrations of regulators such as NusA and RfaH. Similarly, we can reconstitute ECs *in vitro* using either C or O RNAPs on the *his* pause nucleic acid scaffold and assay whether the clamp opening is required for NusA stimulation of pausing or NusA binding. We anticipate that NusA will stabilize the formation of the RNA structures, but this has never been established experimentally. If NusA has the expected effect, we can extend the experiments to look at alterations of NusA that may affect its interaction with nascent RNA structures.

3. Is opening of the clamp required for formation of other stabilized pauses (e.g., backtrack-stabilized pauses)?

The view that the clamp domain undergoes dramatic conformational changes during hairpin-stabilized pausing is well supported by crystallographic snapshots and our biochemical evidence. However, it is unclear whether clamp opening is required for other stabilized pauses such as backtrack pauses or non-backtrack pauses stabilized by active-site rearrangements including misalignment of the NTP substrate in the active site (Artsimovitch and Landick, 2000; Kireeva and Kashlev, 2009; Landick, 2006, 2009).

Locking the clamp closed may suppress the tendency for RNAP to backtrack and thus inhibit backtrack pauses. On the other hand, stabilization of the open-clamp conformation may favor backtracking and pausing because RNAP-nucleic acid contacts are loosened or broken upon opening of the clamp based on the crystal structure of the elemental pause elongation complex (Weixlbaumer et al., 2013). Careful testing of predictions based on these ideas is now necessary to advance understanding of RNA polymerase structure and function.

To ask these questions, we will assemble ECs *in vitro* on a nucleic acid scaffold containing a pause site with either closed- or open-clamp crosslinked RNAPs. After crosslinking these RNAPs with different oxidizing agents, RNAPs will be allowed to transcribe through the pause site encoded in the nucleic acid sequence. The pause efficiency (fraction of RNAPs that recognize the pause site) and the pause escape rate of these RNAPs will be obtained and compared to those values observed for WT. These experiments will help define the central roles of the clamp dynamics in transcriptional pausing and provide mechanistic understanding of pausing stabilized by backtracking of RNAP and active-site rearrangements.

4. Do changes in RNAP conformation control termination in ECs? Does locking the clamp movements and the RNA exit channel affect intrinsic or ρ-dependent termination?

RNAP conformational changes including some clamp movement have been proposed to occur during transcription termination (Peters et al., 2011). In an intrinsic termination model, terminator hairpin formed in the RNA exit channel may affect the clamp position similar to that observed in hairpin-stabilized transcriptional pausing. The

involvement of clamp movement in termination is also supported by a reduction in termination efficiency when part of the clamp region is deleted or sequence changes in the downstream DNA where the clamp contacts (Epshtein et al., 2007). Furthermore, allosteric models of termination postulate that the clamp opening participates in Rho-dependent termination as well. In these models, the terminator hairpin or Rho progressively open the RNA exit and main channels through the clamp movements, rather than extract RNA through the exit channel (Epshtein et al., 2007; Epshtein et al.; Peters et al., 2011). We can test the allosteric models by measuring the rates of Rho or antisense oligo-driven EC dissociation using these RNAPs in which the clamp can be crosslinked shut or open. Similarly to the experiments I performed using an oligo to mimic the pause hairpin (chapter three), we can anneal an antisense oligo to the 5′ end of the nascent RNA of ECs halted at a termination site, previously reported to allow termination (Artsimovitch and Landick, 1998; Shankar et al., 2007). We predict that locking the RNA exit-channel by crosslinking could inhibit intrinsic termination either by inhibiting terminator RNA:RNA duplex formation or by blocking a conformational change required for termination (*e.g.,* clamp movement). We will use the same assays described in chapter four to test whether the terminator duplex can still form when the clamp is closed shut. We will also perform oligo-mediated termination assays with both closed and open crosslinked RNAPs to elucidate the mechanism of termination and conformational changes in RNAP necessary for termination. If termination is not affected, the allosteric models will be excluded. Finally, we will use crosslinked RNAPs to resolve whether or not Rho-dependent termination requires major RNAP

rearrangement involving the clamp movements, opening of the RNA exit channel, or opening of the RNAP main channel.

5. Can we measure the rate of duplex formation accurately using single-molecule methods to test whether the exit-channel duplex can still form in closed-clamp crosslinked ECs?

In chapter four, we show that locking the clamp shut has a dramatic effect on oligo-mediated pausing. One explanation for these results is that the paused RNA:RNA duplex formation is affected by crosslinking of the clamp in the closed conformation. In agreement with the idea that an oligo would bind weakly to the nascent RNA in crosslinked ECs, crosslinking C (β1044/β′258i) RNAPs reduced the rate of duplex formation (chapter four). To rigorously test our hypothesis that a paused hairpin-mimic oligo can anneal to the nascent RNA of closed-clamp ECs with a slow rate and to deduce this slow rate accurately, we will use single molecule methods. Single molecule methods such as colocalization single molecule spectroscopy (CoSMoS) offers a great advantage over typical bulk biochemical assays because a large number of different states are usually present in a molecular ensemble (Friedman et al., 2006; Friedman and Gelles, 2012). CoSMoS allows direct observation of both RNAP dynamics and binding of antisense RNA oligo to ECs in real time (Figure 5-2). In this method, a biotin-tagged template DNA will be first immobilized on the surface through biotin-streptavidin linkage. ECs will be reconstituted *in vitro* with WT, mutant, and crosslinked RNAPs labeled with Alexa fluor 488 and a nucleic acid scaffold containing the TAMRA-labeled nascent RNA (Figure 5-2). The fluorescence of Alexa 488 will report the presence of

RNAP in EC whereas individual RNA molecules in ECs can be detected through the fluorescence of TAMRA.

Next, we will determine binding of an antisense RNA oligo to ECs by studying the FRET between two fluorophores. TAMRA dye on the nascent RNA can be served as a donor molecule. We will incorporate Cy5 dye, an acceptor molecule, at the 3′ end of the antisense oligo. We will then measure the FRET efficiencies as the duplex RNA forms. If this duplex RNA formation brings these two fluorophores closer to each other, energy transfer between two fluorophores will be observed. Alternatively, we will place TAMRA dye on the antisense RNA oligo. When it binds to the nascent RNA labeled also with TAMRA in ECs, fluorescence of TAMRA present in both the nascent RNA and the antisense RNA oligo will be quenched (Figure 5-3). We can extend this study to test regulatory factors, which suppress or enhance pausing. Therefore, single molecule observations in conjunction with fluorescent biochemical studies will allow us to investigate whether hairpin formation or binding affinities of short oligos to the exiting RNA is affected by restricting the clamp opening, regulatory binding, or alterations in RNAPs.

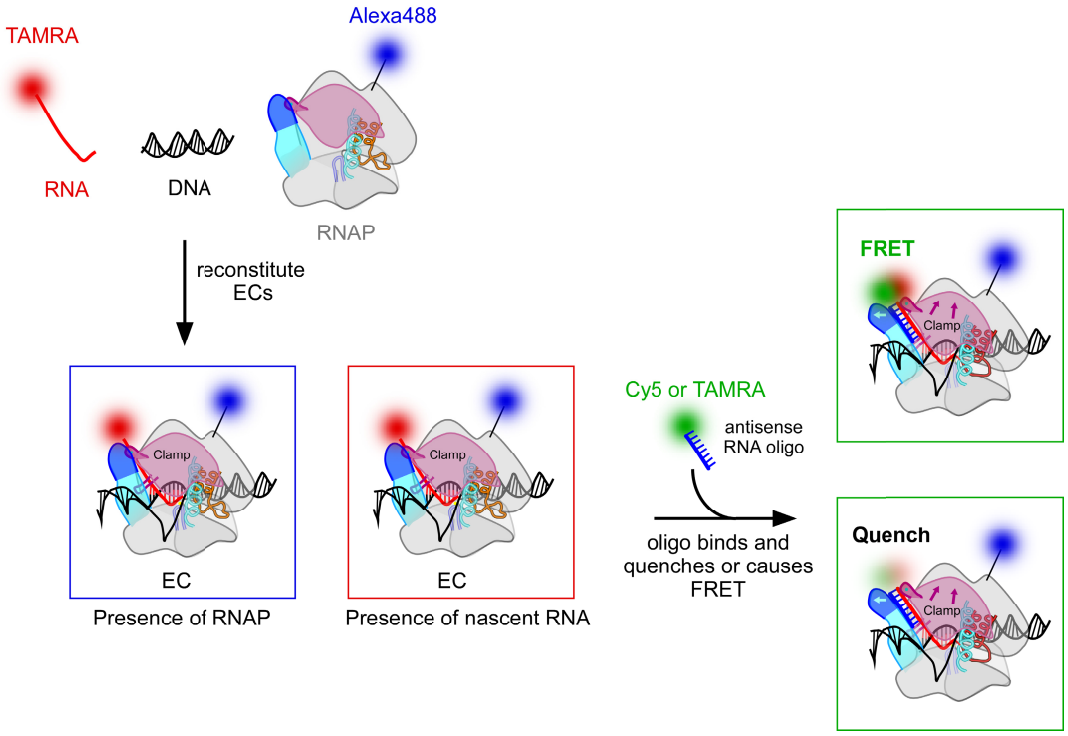


Figure 5-2. Colocalization single-molecule spectroscopy for monitoring the oligonucleotide binding to reconstituted ECs

Figure 5-3 legend.

First, a biotin-tagged template DNA will be immobilized on the streptavidin-conjugated surface of a flow chamber. Reconstitution of ECs will then be performed. RNAP will be labeled with Alexa Fluor 488 (blue dot) and the presence of RNAP (highlighted in blue box) can be detected through the fluorescence of an attached Alexa488 using a total internal reflection fluorescence (TIRF) microscope. Similarly, carboxytetramethylrhodamine (TAMRA, red dot) fluorescence present in the nascent corresponding to individual nascent RNA will report existence of nascent RNA in ECs (highlighted in a red box). Antisense RNA oligo (blue) labeled with either Cy5 (for FRET analysis) or TAMRA (for quenching) will then be introduced. The fluorescence of TAMRA dye that is place on the nascent will be quenched upon binding to TAMRA labeled antisence RNA oligo (green box: bottom panel). If the Cy5-labeled the antisense oligo is used, FRET will occur between TAMRA that is placed on the nascent RNA and Cy5 that is on the antisense oligo (green box: Top panel).

References

Artsimovitch, I., and Landick, R. (1998). Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. Genes Dev *12*, 3110-3122.

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

Artsimovitch, I., Svetlov, V., Murakami, K.S., and Landick, R. (2003). Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. J Biol Chem *278*, 12344-12355.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R.*, et al.* (2012). Opening and closing of the bacterial RNA polymerase clamp. Science *337*, 591-595.

Darby, N., and Creighton, T.E. (1997). Probing protein folding and stability using disulfide bonds. Mol Biotechnol *7*, 57-77.

Ederth, J., Mooney, R.A., Isaksson, L.A., and Landick, R. (2006). Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. J Mol Biol *356*, 1163-1179.

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). An allosteric path to transcription termination. Mol Cell *28*, 991-1001.

Epshtein, V., Dutta, D., Wade, J., and Nudler, E. An allosteric mechanism of Rho-dependent transcription termination. Nature *463*, 245-249.

Friedman, L.J., Chung, J., and Gelles, J. (2006). Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. Biophys J *91*, 1023-1031.

Friedman, L.J., and Gelles, J. (2012). Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. Cell *148*, 679-689.

Hawkins, M.E. (2008). Fluorescent pteridine probes for nucleic acid analysis. Methods Enzymol *450*, 201-231.

Kireeva, M.L., and Kashlev, M. (2009). Mechanism of sequence-specific pausing of bacterial RNA polymerase. Proc Natl Acad Sci U S A *106*, 8900-8905.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.

Landick, R. (2009). Transcriptional pausing without backtracking. Proc Natl Acad Sci U S A *106*, 8797-8798.

Lee, D.N., Phung, L., Stewart, J., and Landick, R. (1990). Transcription pausing by Escherichia coli RNA polymerase is modulated by downstream DNA sequences. J Biol Chem *265*, 15145-15153.

Malinen, A.M., Turtola, M., Parthiban, M., Vainonen, L., Johnson, M.S., and Belogurov, G.A. (2012). Active site opening and closure control translocation of multisubunit RNA polymerase. Nucleic Acids Res *40*, 7442-7451.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. Mol Cell *50*, 882-893.

Palangat, M., Hittinger, C.T., and Landick, R. (2004). Downstream DNA selectively affects a paused conformation of human RNA polymerase II. J Mol Biol *341*, 429-442.

Peters, J.M., Vangeloff, A.D., and Landick, R. (2011). Bacterial transcription terminators: the RNA 3'-end chronicles. J Mol Biol *412*, 793-813.

Rist, M.J., and Marino, J.P. (2002). Fluorescent Nucleotide Base Analogs as Probes of Nucleic Acid Structure, Dynamics and Interactions. Current Organic Chemistry *6*, 775-793.

Shankar, S., Hatoum, A., and Roberts, J.W. (2007). A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. Mol Cell *27*, 914-927.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., and Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. Nature *468*, 978-982.

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.

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.

Weixlbaumer, A., Leon, K., Landick, R., and Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. Cell *152*, 431-441.