Trigger loop dynamics aid intrinsic transcription termination by *Escherichia coli* RNA polymerase

By

Ananya Ray-Soni

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 04/05/2017

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

Robert C. Landick, Professor, Biochemistry and Bacteriology
James L. Keck, Professor, Biomolecular Chemistry
James C. Weisshaar, Professor, Chemistry
Patricia J. Kiley, Professor, Biomolecular Chemistry
Samuel E. Butcher, Professor, Biochemistry

To my parents, Dr. Jharna Ray and Dr. Kunal Ray:

For teaching me the value of education through your words and your lives,
for your endless encouragement, and for always believing in me.

To my husband, Dr. Mufaddal Soni:

For your unwavering love and support, your patience,
and your constant faith in my abilities.

Ma, Baba, Muffi –
you are my light.
I dedicate this work to you.

ACKNOWLEDGEMENTS

I have been blessed to be born into a family that exalts education above all else, a family that to this day continues the academic pursuit with a passion and a dedication that I can only hope to one day match. It is because of this nurturing environment that I am in academia today and am approaching the completion of my doctorate. However, no amount of nature or nurture can prepare you for the trials of academic science! I would like to take this moment to thank those of you who have helped me through my journey through graduate school and made this a deeply enriching experience. It is with all of your support and encouragement that I excitedly charge ahead to pursue my future in the world of academia!

I would first like to thank my advisor, Dr. Bob Landick, for his continued support through graduate school. I have rarely come across someone with your expert ability to conceptualize a problem from the perspective of diverse disciplines, and at so many different levels. Being in your lab has taught me what it truly means to address a problem rigorously and how to communicate my work effectively. I will carry these lessons through the rest of my career. It has been extremely enriching to train under you, and to glean from you whatever wisdom I could. I hope that I can continue to emulate your critical and comprehensive thought process as I build my own academic group.

I would like to thank the members of my thesis committee –Dr. Jim Keck, Dr. Jim Weisshaar, Dr. Sam Butcher , and Dr. Tricia Kiley – for generously providing their time through the years. Your insightful questions and thoughtful suggestions to my project have certainly played an important role in shaping this work.

To my lab – what would I do without such a warm and helpful group like you to work with? What we have in our lab – the collegial attitude, the willingness to help and teach, the persistent interest in each others' projects and lives, our friendships – I think it is a rare gift. I have learned so much from each of you (Landick lab alumni included), and I am very grateful for the environment you have created and the

warmth you have showed. It goes without saying that I will miss you all, and I wish each of you the very best that life has to offer for both your careers and personal endeavors.

I could not think of leaving UW without thanking Kate Ryan of the Biochemistry Department for smilingly taking care of all of the organizational stress for us, and making it seem effortless at that! You've been simply a joy to work with, Kate. Laura Vanderploeg has been very helpful in constructing a figure for my manuscript, and was also wonderful to work with – our few sessions working together have never failed to entertain! I would also like to thank Dean Maly and Gary Pine in the Microbial Sciences Building workshop for patiently repairing all of the broken apparatus I've brought down over the years, and being so friendly all the while!

My friends here in Madison – you have all been such a caring and welcoming support-network, and provided so many wonderful experiences …and a welcome escape when project life has been difficult! My time in Madison would have been completely different if it weren't for you. Though we have dispersed now, I look forward to seeing how all of our lives unfold – there are many more memories to be made together! I must also make a specific shout out to Rika and Ranga: you both have become part of our family away from home over the years. You have both independently been an incredibly valuable scientific and emotional presence in my time here, and you continue to be so though you are no longer in Madison. Thanks for always being willing to lend a critical or understanding ear (whatever was needed at the time!). I know we have a long future of science and laughs – and life! – to share ahead of us.

Mummy, Papa and Musti – my recently acquired family, I am so thankful for your warm understanding, encouragement, and continual interest in my career. It means so much to me that you support me in my efforts and take such an interest in my every goal, in my every success, no matter how small. Thank you for always being a positive and comforting support to me. I am so excited to share each new chapter of my life with you and continue to discover each other as the years progress.

To my best friend and husband – Muffi. Without you this PhD would have been inconceivable. You make me happy every day, even when challenges are thrown my way. Your dedicated work ethic and infallibly empathetic and understanding nature inspire me to push myself to be the best version of me – in my career or otherwise. I am propelled to keep moving forward by your persistent words of encouragement, your enduring patience, and your unfaltering faith in me. Thank you for supporting me in everything I do.

And finally, my parents. Words cannot express my gratitude to you. You are the reason I am here, the reason I have the courage to pursue this dream. Baba, you have been my lifelong mentor and friend. I am so fortunate to always have your experienced and willing advice on science and this career, and I have so much more to learn from you still. I can only hope to one day achieve half of the success and respect that you have achieved already. I started my research career in your lab, and I look forward to continuing the tradition of working together. Ma, though you don't realize, you are truly an inspiration. Never have I met a person with your untiring dedication and sincerity. Your achievements as well as those of your department are a testament to your vision and determination. I am so very blessed to have a teacher, a friend and a mother in you. Ma and Baba, every day you inspire me to follow my dreams, no matter how challenging they may be. It is with the humility, perseverance and attention to detail that you both have instilled in me, and with your constant advice and encouragement that I complete my doctoral work. Thank you for always believing in me and showing me the way.

Trigger loop dynamics aid intrinsic transcription termination by *Escherichia coli* RNA polymerase

Ananya Ray-Soni

In the laboratory of Professor Robert C. Landick

University of Wisconsin – Madison

Madison, WI

ABSTRACT

RNA polymerase (RNAP) is the essential enzyme responsible for the timely and accurate expression of genes, a process crucial for cell survival and adaptability. To reliably synthesize long transcripts with diverse sequences, RNAP must form many non-specific but stabilizing interactions with the RNA transcript and the DNA chromosome as part of an elongating transcription complex (EC). At the ends of genes, however, the feat of transcription termination is to reverse these highly stabilizing contacts and disassemble the EC into its components. Termination is essential to ensure that only desired genes are expressed, and for the release of RNAP after transcription of a gene to recycle the enzyme for the next round of transcription. In bacteria, the formation of an RNA hairpin structure (T_{hp}) in the nascent transcript immediately upstream of a weak RNA–DNA hybrid in the main channel of RNAP is sufficient to dissociate the EC in a process called intrinsic termination. Remarkably, this process occurs over a precise 2-3 nt sequence window, after the EC has stably traversed up to kilobases of DNA. Whereas the roles of the sequence elements of the intrinsic terminator have been extensively characterized (described in **Chapter 1**), the contributions of structural rearrangements within RNAP that enable the termination mechanism remain almost completely unknown. The main objective of

this work was to setup an experimental system by which the contributions of various RNAP modules could be assayed, and to subsequently use the system to assess the role of one such module in termination.

As termination is a kinetically controlled process with multiple points along the termination pathway at which ECs must "decide" between termination and elongation, termination efficiency (TE) can be determined by effects on elongation or termination. Therefore, to confidently assess the effects of RNAP mutations on the termination mechanism, it is essential to be able to distinguish effects on termination from elongation. In **Chapter 2**, I describe the assay platform I developed, which enables separate measurements of elongation and termination rates at a termination site, and also measures the rates of all intervening pause steps in the termination mechanism. Importantly, ECs are only required to elongate 3 nt to either terminate or bypass the terminator in this system, enabling the study of RNAP mutants that are severely defective for elongation. To our knowledge, this approach provides the first available method to isolate the rates for each step of termination, and determine the effects of RNAP mutations at each decision point on the termination pathway.

Using the termination assay I developed, I was able to address the role of a polymorphous active site-proximal module of RNAP – the trigger loop (TL), presented in **Chapter 2**. The use of this assay was particularly important to study the TL because the TL catalyzes nucleotide addition 10^4-fold, and TL mutations can thus have profound impact on elongation. Studying the TL was of interest because it has been suggested to stimulate termination previously, by folding into the trigger helix (TH) conformation and contacting a T_{hp} that has invaded the main channel of RNAP (TH–T_{hp} model). We find that the TL significantly accelerates the rate-limiting step of termination – functional inactivation of the EC, indicating a role in stimulating the termination mechanism. However, neither the TH conformation nor an unfolded TL conformation are favorable for termination, in contradiction to the TH–T_{hp} model.

We suggest that the TL aids termination through its conformational flexibility rather than a discrete conformation, acting to expand the conformational diversity of ECs and allowing access to low energy pathways to termination. An assessment of the decisions points on the termination pathway preceding EC inactivation reveals that the TL also increases the flux of ECs into the termination pathway, thus increasing the overall TE.

Additionally, I investigated the mechanism by which a transcription factor, NusA, enhances TE. NusA is thought to be associated with all active ECs *in vivo*, and has been shown to aid both intrinsic termination and hairpin-stabilized pausing by aiding RNA duplex formation in the exit channel of RNAP. However, previous work from our lab indicated that NusA may have additional roles in termination required for its full effects in enhancing TE. In order to determine some of these roles, I tested the effects of NusA using my termination assay under conditions in which T_{hp} formation is not rate-determining. However, I found that this system is unable to recapitulate the large effects of NusA-mediated TE-enhancement that I observed from a promoter-initiated termination assay. Possible explanations for the decreased effects of NusA in my termination assay include: *(i)* RNA upstream of the RNA duplex, which would normally be present for transcription that had initiated from the upstream promoter, is required for NusA effects; *(ii)* The major effect on TE arises from NusA-stimulated RNA duplex formation, which was intentionally minimized under the tested assay conditions. However, neither the presence of upstream RNA, with or without competing RNA structures, nor conditions with slower rates of T_{hp} formation could recreate the extent of NusA-mediated TE enhancement observed in the promoter-initiated assay. We hypothesize that the presence of initiation factor sigma in the promoter-initiated experiments may slow termination, and that NusA may compete for the sigma binding site on the EC, and thereby stimulate termination.

This work provides new insight into the role of the TL in termination and highlights the importance of protein conformational dynamics in modulating termination rate. The

termination assay developed here also presents an easily adaptable platform to determine the roles of various other RNAP modules and extrinsic factors in each step of the kinetically complex intrinsic termination pathway.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Appendix A

Appendix B

Appendix D

LIST OF TABLES

Chapter 2

Chapter 3

Appendix A

Appendix B

Appendix D

My thesis work,
explained for the non-scientist

Prepared for:

The Wisconsin Initiative for Science Literacy (WISL) Thesis Award Program at the University of Wisconsin–Madison (http://scifun.chem.wisc.edu/news/thesis_awards.htm).

Chapter contributions:

A.R.S. wrote all text and constructed all Figures; editors at the WISL copyedited the text.

Have you ever considered how we get different organs and body parts if they *all contain the same DNA*?

In forensic science (think to your arsenal of knowledge from crime shows like CSI and NCIS), crime scenes are always combed for hairs or dry skin particles that may have been left by the perpetrator. From these hair or skin particles, forensic scientists can isolate DNA, which encodes the unique genetic information that makes a person who they are. In essence, it provides a DNA "fingerprint" of anyone that was at the crime scene. Once the police have a suspect, they can then take a cheek swab, isolate the DNA from this cheek swab, and compare it to the DNA from the hair/skin particles they found on the scene. If the two DNA samples match, it is likely that their suspect was in fact at the scene of the crime.

…But wait! Why would the DNA from such different parts of the body with completely different appearances, textures, and functions be comparable? It turns out they are not just comparable; they are in fact the *same*! Think about this for a second – we know that our DNA **genome** encodes the genes required for each biological "machine" to function. So if all of the different parts of a human body that are so different from each other contain the *same DNA*, how do they look and behave so differently? This is because even if all parts (and the **cells** that make up those parts) have the same genes and DNA, only some of the genes are "turned on" in each part. In other words, although all of the information needed for us to grow and survive is encoded in our DNA, our cells can only utilize this information if the necessary gene is turned on and is used to produce the biological "machine" that it encodes. And which machines are being produced in a particular cell determines its **cell-type**. By controlling which sets of genes are "on" or "off" in each cell, the same DNA can be used to produce all of the diverse cell-types that make up a human body.

Sidebar: What are **cells** and **cell-types**?

Cells (often called the "building blocks of life") are the smallest structural and functional units of all living organisms that are *self-sustaining* and have the potential to self-replicate. What do I mean by self-sustaining? All cells have the ability to produce the energy and biological resources to be able to live and grow without support from other biological structures. In fact the simplest organisms are single-cellular organisms (*e.g.* **bacteria**, see sidebar below), which can grow and replicate completely independently. Cells may have various different characteristics depending on the genes encoded in their DNA, as well as their **gene expression patterns** (*i.e.* which genes are "on" or "off").

In multi-cellular organisms (sometimes called "higher organisms," *e.g.* humans) all cells within the organism share the same DNA. However, these cells vary in their gene expression patterns, resulting in cells with distinct characteristics. We call cells with the same characteristics the same **cell-type**. Cells of the same or related cell-type often adhere together to form **tissues** (*e.g.* muscle tissue is constituted of muscle cells). And tissues of different types that work together to carry out one larger biological function (*e.g.* pumping blood to the body) form organs (*e.g.* a heart).

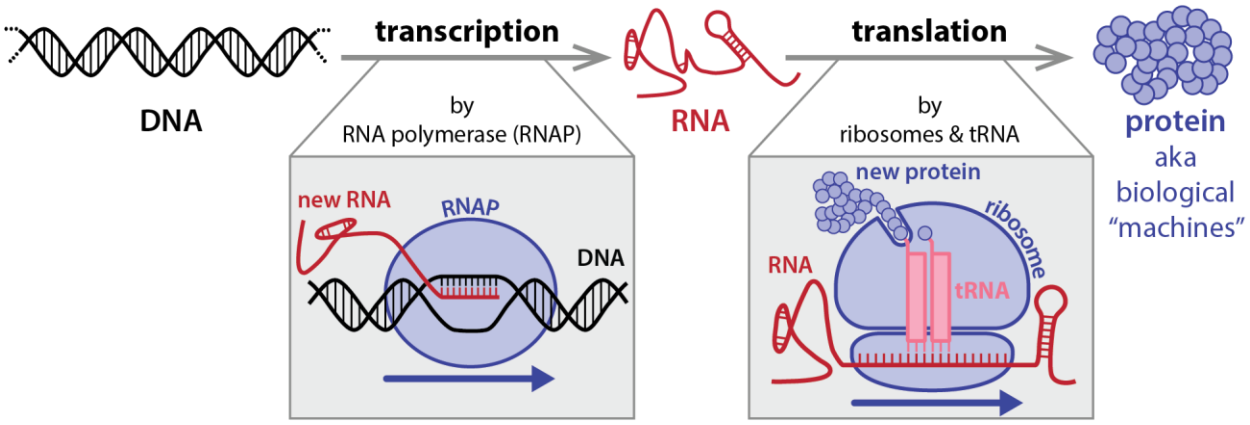
cells (*e.g.* nerve cells) → tissues (*e.g.* nerve tissue) → organs (*e.g.* brain)

How are some genes specifically "turned on", while others are "off"?

To think about how genes are turned on and off, we must first understand how the corresponding biological machines are made. In biology, the **central dogma** describes the process by which a gene is turned on and used to create the biological machine it encodes (**Figure 1**). In a process called **transcription**, the biological machine called **RNA polymerase (RNAP)** "reads" the information in the DNA gene and converts it to RNA. The RNA acts as a sort of temporary message, from which protein (*a.k.a.* the biological machines

we've been talking about) can be made in a process called **translation**. Once these proteins are made, they can carry out their functions in the cell.

Since the RNA needs to first be produced for the protein to be made from it, the process of transcription is the first step at which a gene can be turned on or off. Different environmental factors and cellular signals help ensure that RNAP transcribes the correct genes for each cell-type and condition. *I.e.*, we say that the process is **regulated** by different factors. In this work, I was studying how RNAP facilitates the process of transcription.



<u>Figure 1:</u> The central dogma of gene expression: how RNA and proteins are made from DNA.

RNA is produced from DNA by RNA polymerase (RNAP) in a process called transcription. Proteins are then produced from the RNA transcript by the ribosome and tRNAs in a process called translation. Both RNAP and ribosomes are also proteins, while tRNA is actually a functional RNA molecule (see also **Figure 3**)!

Why study transcription?

As I alluded to in the previous section, transcription is the first step at which gene expression (we say a gene is expressed when the gene is "on") can be regulated; translation of RNAs to proteins can act as an on/off switch too! However, both transcription and translation use valuable cellular resources to produce RNAs and proteins, and transcription of a gene *must* occur before translation. So if a gene needs to be off in certain conditions or cell-types, transcription is the less resource-"expensive" step to act as an on/off switch, since it is the first step of gene expression.

Transcription can largely be divided into 3 stages (**Figure 2**):

(1) **initiation** of transcription: in which RNAP binds to the DNA at the start of a gene that is turned on and begins **transcribing** the RNA **transcript** from the gene;

(2) **elongation** of the new RNA transcript: in which RNAP decodes the entire message that is encoded in the DNA gene and converts it to RNA;

(3) **termination** of transcription: which causes the transcribing **elongation complex (EC)** to be dissociated into its components – the new RNA transcript, the DNA genome, and RNAP.

In this work I was interested in understanding how RNAP carries out the last step of transcription – termination.

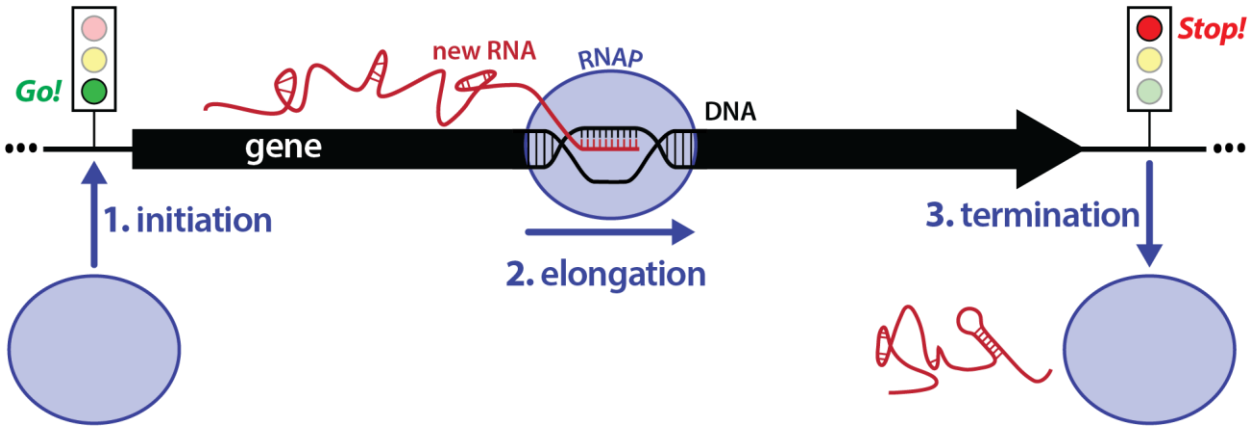


Figure 2: **The 3 stages of transcription.**

So initiation turns the genes on; why does it matter whether you terminate the process or not?

While initiation is very important for turning genes *on*, termination is essential for turning them *off* (**Figure 3**). It is also important for recycling RNAP for new initiation events, releasing the RNA, and it prevents collisions between RNAP and other proteins that travel along the DNA, as this could damage the genome.

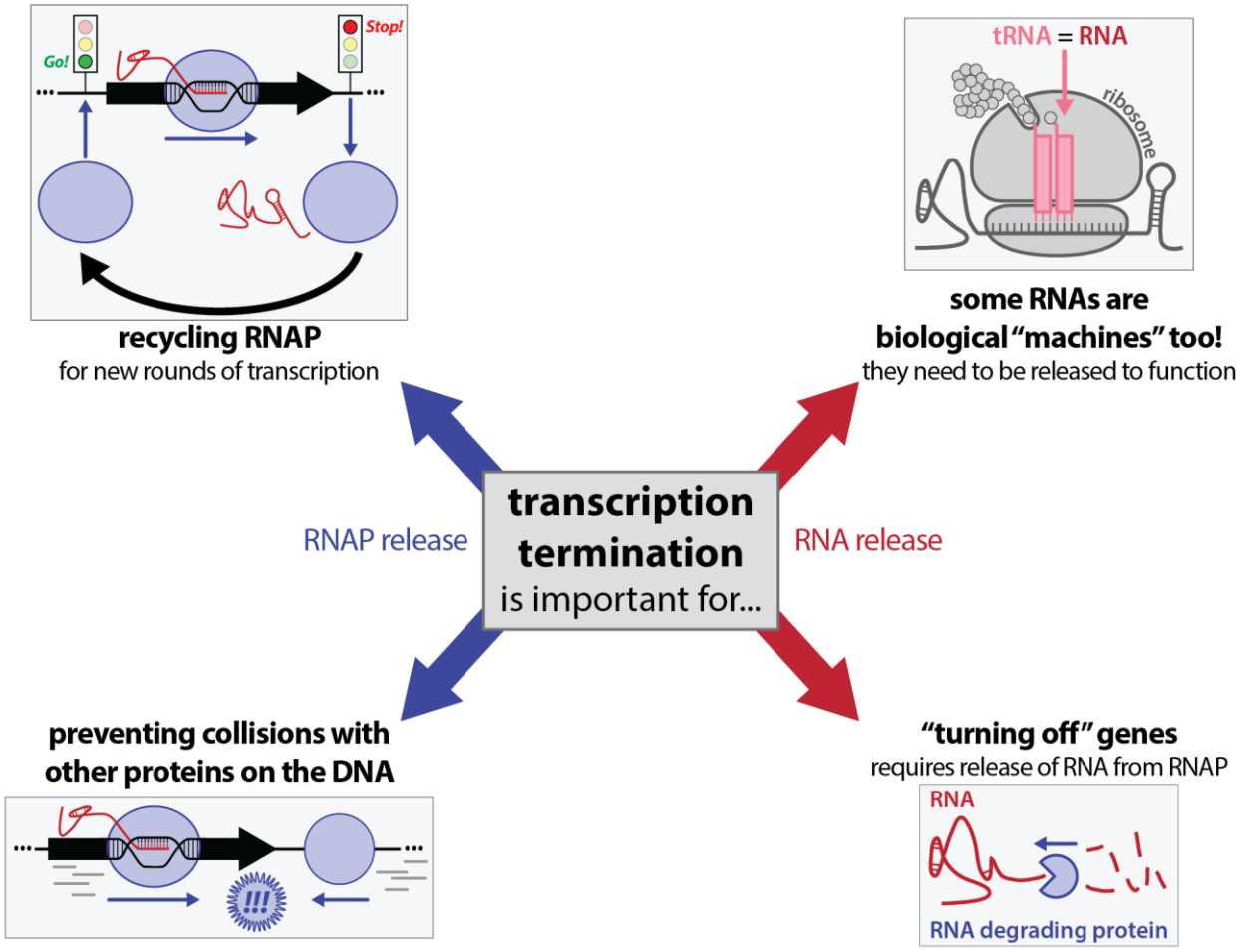


Figure 3: The importance of transcription termination is manifold.

Studying termination in bacteria

I am studying transcription termination in **bacteria**. Bacteria are diverse organisms present in and around everything we encounter, and understanding how they function has great potential for many different applications (see the sidebar below for discussion on the significance of bacteria)! In addition, the efficient minimal biological systems used by bacteria often provide a great way to understand how more complex versions of the analogous system may function in humans and other higher organisms.

Sidebar: Why are **bacteria** so important?

Bacteria are microscopic single-cell organisms that pre-date human existence on Earth. This means that they are able to carry out all of the functions needed to live and grow in just one cell! The elegant genetic makeup of these diverse organisms also allow them to exist in various temperate and extreme environments including the soil, hot springs, extreme cold or acidic environments.

We often think of bacteria as harmful, pathogenic organisms that make us sick. But this isn't true of *most* bacteria. While there are some bacteria that are parasitic, and are harmful to whatever host organisms they infect, the vast majority of bacteria is harmless and coexists with us peacefully. Did you know that every surface that we come in contact with is covered with bacteria? On average, one teaspoon of water contains ~5 million bacteria! In fact, the human body may be home to more bacterial cells than human cells! But don't be alarmed – most of these bacteria are completely harmless, and some of them actually *help us survive*. For instance, most bacterial cells associated with the human body inhabit our gut and help us unlock vital nutrients from our food.

Scientists in many industries have also found ways to utilize the unique abilities of different bacteria, *e.g.* for producing cheese and yogurt, cleaning up oil spills, mining metals, and producing fuel from food waste. And this is just the tip of the iceberg! As these examples hopefully illustrate, there is a lot of value and utility to understanding bacterial function, and likely a great deal of untapped potential to be uncovered from continued bacterial research!

What do we know about bacterial termination so far?

In bacteria, **intrinsic termination** is one of the 2 major ways by which transcription is terminated, and it is signaled by the formation of a structure in the new RNA transcript that is being made by RNAP. Interestingly, *not all ECs that encounter the RNA termination signal actually terminate*. The EC must make a *decision* between continuing elongation, and terminating. This decision involves a choice between 2 paths that require more or less energy to cross. Think of the EC as yourself riding a bike to either a friend's house or home after school. Assume your only motivation to go home or to your friend's house is based on the difficulty of crossing the path to get there. Under this assumption, if the hill to your friend's place is much steeper than the hill on the road home, you need more energy to go to your friend's place, and you are more likely to choose to go home (**Figure 4**).

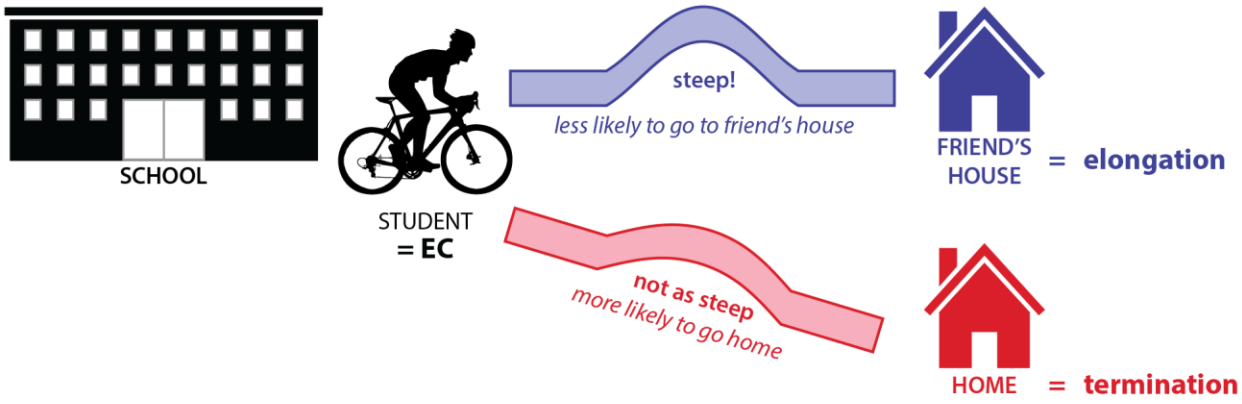
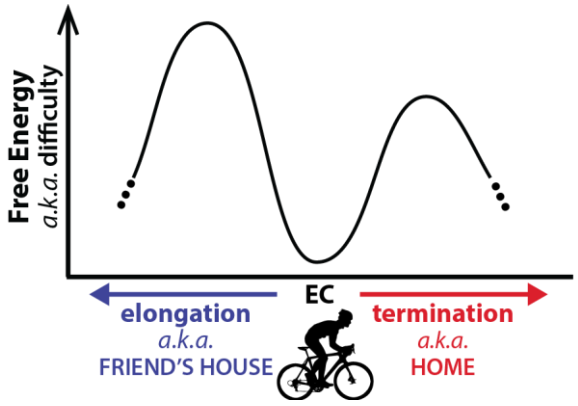


Figure 4: The difficulty of crossing hills determines which path you will take.

In this analogy, the student riding a bike represents the transcription elongation complex (EC). Here we assume that the probability that the student will go HOME (*i.e.* terminate) versus to his/her FRIEND'S HOUSE (*i.e.* elongate) is determined by the difficulty of crossing the hill on the path to either destination. *Biking cartoon adapted from http://www.veloscience.org/?p=4403.*

Similarly, the relative probability of carrying out 2 different **reactions** (*i.e.* crossing 2 different hills in our analogy) can be depicted as an **energy diagram** (**Figure 5**)**.** When an EC encounters 2 different **energy barriers** (hills) to elongation (your friend's house) versus termination (home), a higher energy barrier is harder to cross, making it the less probable choice (**Figure 5**).



<u>Figure 5:</u> An energy diagram depicts the relative energy barriers to the competing reactions of termination (right) versus elongation (left).

The height of the peak for either reaction indicates the amount of energy required to complete the reaction, where a higher peak requires more energy, and is therefore less probable. At termination sites, the energy barrier to termination is lower than the barrier to elongation, resulting in appreciable amounts of termination.

To complicate matters slightly, termination is a process with many steps, so the EC has to make the decision between elongation and termination many times before it finally successfully terminates. By our analogy, you must cross a few hills on your way home, and before each hill, you have the choice to cross a different hill and go to your friend's house instead (**Figure 6A**). Right now, the way termination is studied, researchers most commonly look at what percent of ECs terminated by the end of the experiment (called **termination efficiency, TE; Figure 6B, top**), *i.e.* they determine how many kids ultimately went home at the end of

the day. But with so many steps and decision points between school and home, a lot of information is lost about each of those decisions. I wanted to devise a method to "see" how many ECs choose to go over the hill towards home or the one towards their friend's house, *at every decision point*.

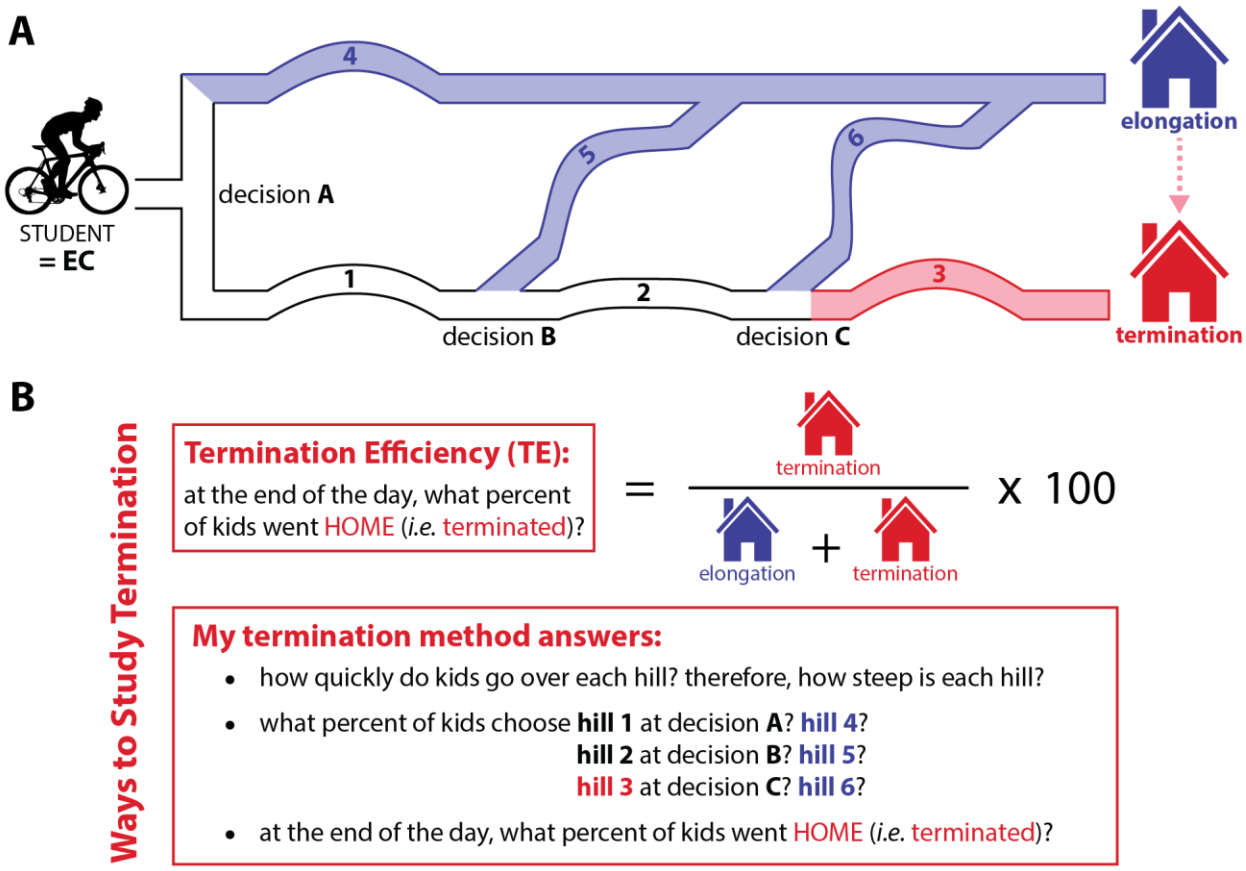


Figure 6: There are many steps and decision points on the path to termination.

(A) The multiple steps to termination of varying difficulty are depicted as numbered hills. Higher hills indicate greater difficulty, and thus lower probability of crossing that hill. The EC must cross **hills 1, 2 and 3** to successfully terminate. Paths that result in elongation and termination are *shaded blue* and *red,* respectively. The *pink dashed arrow* indicates that while not all ECs will terminate at the given termination site, they will eventually terminate at a later site on the DNA. **(B)** Termination efficiency (top) is the most common and easy-to-use method of studying termination. This method has been used to determine a great deal of information about the RNA termination signal. The method I developed in my graduate work (bottom), however, yields a lot more information about each of the steps of termination, and is thus a powerful new way to study the termination process. See explanation in the section below.

Problem solving time! How do I watch ALL the decision points?

When you are riding your bike over a hill, the height and steepness of the hill will not only affect your motivation to take that path, but also the speed at which you cross this hill. If the hill is higher, it will take you longer to cross the hill. Analogously, a population of ECs cross a higher energy barrier at a slower speed (or **rate**) than a lower energy barrier. Why? A higher energy barrier means that the reaction requires more energy to overcome the barrier, and is therefore less likely; a reaction that requires less energy is more likely. Another way of saying this is: if the reaction requires less energy, more ECs are likely to obtain the necessary energy to overcome the barrier. This results in a faster reaction rate, which we can measure in our experiments. Then, using the information about the rate at which ECs cross different energy barriers, I can figure out how high the barrier was. BUT, as we saw in **Figure 6A**, the termination process is complicated, with many hills and decision points.

When we encounter a complex problem in arithmetic, we break it down in to simpler parts, right? For *e.g.*, if I ask you to calculate: 22 + (8 x 4), PEMDAS (the order of operations) tells you to do each function one at a time. You would first compute the part in parentheses (8 x 4 = 32), simplifying the problem to: 22 + 32. Then the problem is simply addition, and you can easily determine that the solution is 54.

Similarly, to simplify the problem of too many different paths home and to the friend's house, I first measured the reaction rates with the road leading home (to termination) blocked. How did I do this? Recall how I told you that there was an RNA signal that causes termination. Well I changed the signal just enough, so it doesn't cause termination anymore. In our metaphor, I effectively made one of the hills leading home (**hill 2** in **Figure 6A**) so high that it was impossible for anyone to go over them. With **hill 2** blocked, no ECs could cross **hills 2, 3 or 6**. In this way I could simplify the problem and determine the reaction rates for crossing only **hills 1, 4 and 5** first.

With half of the reaction rates determined, I next wanted to find out how quickly ECs could cross the remaining hills on the path home. To do this, I gave the ECs the full termination signal, which made **hill 2** "crossable", and followed how quickly ECs terminated (went home) or elongated (went to a friend's house). With half of the information already in place – the reaction rates for crossing **hills 1, 4 and 5** – I only had to figure out the other half of the information from this experiment. This was MUCH more doable than trying to extract all the information from one experiment!

Using this method, I could glean a lot more information about the termination process, and each of the steps along the way, including the overall termination efficiency (**Figure 6B, bottom**). I was able to determine how quickly ECs crossed all the hills to either the friend's place or home (*i.e.* the reaction rates for crossing each of the energy barriers). By extension I could also calculate the relative heights of each of those hills (or energy barriers). And since we know that an EC is twice as likely to go over a hill that is half the height of another, I could also figure out how likely it was that an EC would take the path towards home at each decision point. Thus, the method I developed helped me identify which steps in the termination process are faster and slower, and which decision points are the most critical in determining if an EC ultimately terminates.

Cool, so I have a new method! …what else can I use it for?

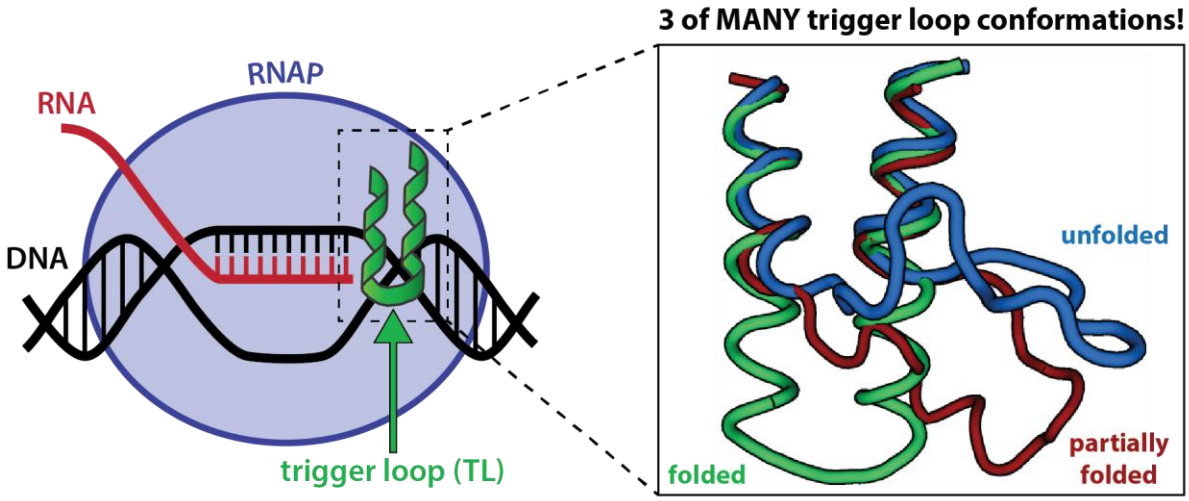
As I briefly described above, intrinsic termination is instigated by a sequence signature in the new RNA transcript, which forms an RNA structure within the EC. Through many carefully designed experiments, scientists before me have figured a lot out about how specific elements of this RNA sequence help signal termination of the EC, and favor the termination decision. However, the way that RNAP responds to this RNA termination signal and enables termination remained almost completely unknown. I wanted to use my method to start to understand how movements of RNAP facilitate termination.

Proteins are highly mobile structures that change shape (or **conformation**) to carry out their functions as biological machines. You can imagine that any machine has many moving parts (like levers and gears), and it's important to understand how these moving parts facilitate the main overall function(s) of the machine. Similarly the specific movements of different elements (or **modules**) of the protein structure – "cogs" in the machine – together facilitate the function of the protein. I was interested in determining the function and conformational changes of one such module of RNAP: the trigger loop (TL; **Figure 7**). The TL is known to be highly structurally dynamic, and requires this flexibility to enable RNAP to carry out many of its functions during transcription. I wanted to know whether the TL is important for any of the steps and decision points in the termination process.

How do we find out if parts of a molecular machine are important? In biochemistry, we either remove the module entirely (called deleting it), or change it in a specific, predictable way (called mutating it). We made RNAPs with such a TL deletion or TL mutations that behaved in ways that we had already determined by other experiments. I assembled ECs using these TL-deleted or TL-mutated RNAPs, and tested their abilities to terminate and elongate using my method. With this information I could determine:

(1) how the TL impacts the ability of the EC to cross each hill on the paths to termination and elongation, and

(2) how likely these ECs were to move towards termination at each decision point.

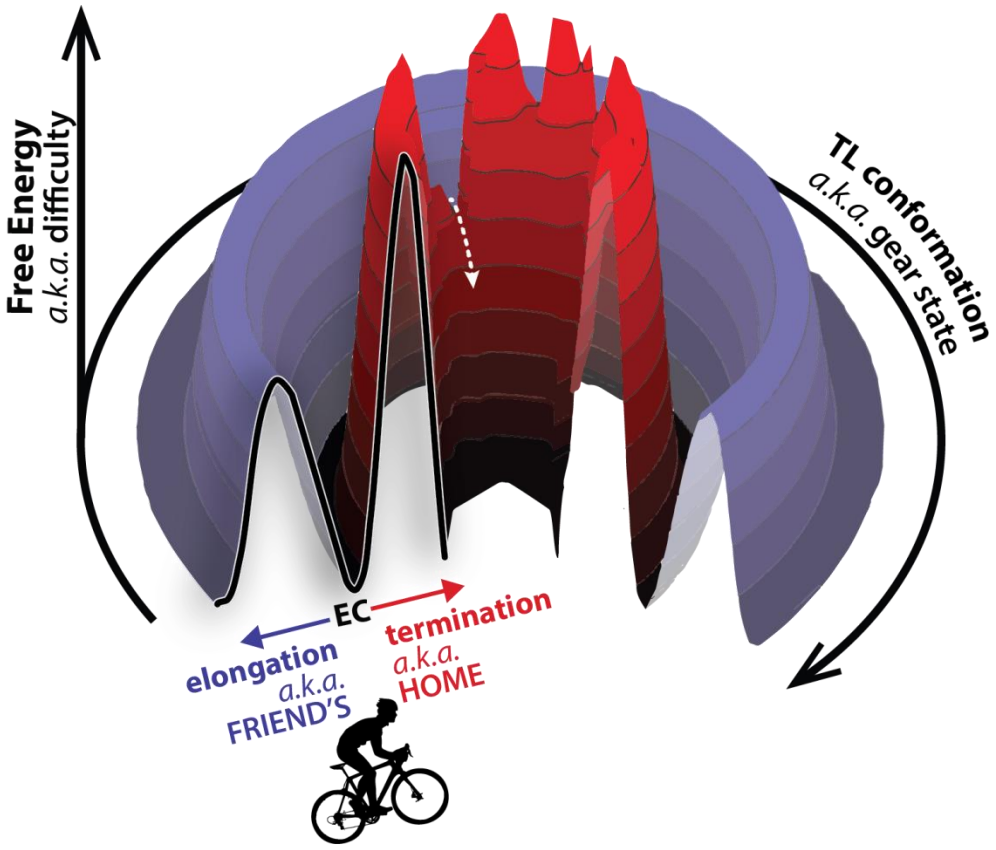


<u>Figure 7:</u> The structurally dynamic trigger loop (TL) module of RNA polymerase (RNAP) must adopt diverse conformations to enable the different functions of RNAP during transcription.

What did I find out? Was the TL involved in termination? Did it help send ECs down the termination path?

Spoiler alert: yes it did! From our experiments, we found out that the TL was in fact very important for a number of the decision points on the path to termination. I found that in the absence of the TL (*i.e.* with TL-deleted RNAP), the ability of the ECs to complete the final step of termination (**hill 3** in **Figure 6A**) was *severely decreased*! Strikingly, this impairment of TL-deleted ECs in termination was not detectable from measurements of termination efficiency. The termination method I developed was thus instrumental in identifying this important role of the TL in termination.

My results also indicated that mobility of the TL*, i.e.* its ability to take on many shapes or conformations, was very important for its function in termination. In our analogy, think of the bike you are riding as a multi-gear bike. The same hill may seem more or less daunting depending on the "conformation" of the gears, *i.e.* the gear-setting. When the bike is set to the ideal gear-setting, it seems easier to cross the hill. Conversely, in other gear-settings (too high or too low settings alike) it may be *extremely* tiring to cross the hill, making it seem like an insurmountable obstacle! In this scenario, the TL is the gear-shifter, allowing you to change the gear (or conformation) of the bike (EC). Now if the gear-shifter is absent from your bike, and you aren't able to switch gears, you are stuck to one gear-setting. Crossing the hill will always seem to be of the same level of difficulty. By having the gear-shifter, you have the option to try different gears, and find the setting that makes crossing the hill the least difficult. Analogously, the TL (gear-shifter) thus enables the EC (you, riding a bike) to sample various conformations (gear-settings) and find conformations that have the lowest energy barrier on the path to termination (gear- settings that make crossing the hill seem the least difficult; **Figure 8, white dashed arrow**).



<u>Figure 8:</u> Flexibility of the TL allows it to explore conformations for which the energy barrier to termination is relatively low.

A hypothetical 3D energy diagram (related to the energy diagram in **Figure 5**) depicts the relative energy barriers (vertical axis) to termination (*red*) versus elongation (*blue*) for different TL conformations (circular axis). Our results suggest that when the TL is flexible, it enables the EC to adopt states for which the energy barrier to termination is relatively low (depicted by the *white dashed arrow*). However, when TL dynamics are limited (in the TL-deleted or TL-mutated ECs), the EC is likely trapped in a state for which the energy barrier to termination is much higher, making it much less feasible. NOTE: The energy barrier to elongation is shown as a single height in this depiction; however, it likely also changes as a function of TL conformation.

Can this method be used by anyone else? What can they use it for?

Absolutely! In science, we publish our work in peer-reviewed journals that are publicly available. This means that we write up all of our results in a story, explaining exactly how we did each experiment, and submit the story to a journal that we think will get an interested audience. Then other experts studying similar problems will read our work and make sure all of the logic and experiment design checks out. If they approve it, our work gets published in the journal, and it's available for anyone to read!

The way we designed our assay, it's easily adaptable for asking other questions about the different steps in termination. Just like the TL, other cogs in the RNAP machine undoubtedly play a role in the process of termination. The key is to design smart and thoughtful experiments and RNAP mutants to address those questions. Someone could designs mutations in other parts of RNAP, or the termination signal itself, and determine the effects at each decision point. With more and more of these kinds of experiments, we can have a much clearer picture of how all the cogs in the RNAP machine must work together to facilitate a complete successful journey home.

Are there any human health-related implications of my work?

While the bacteria provide a model system for understanding the fundamentals of human biology, there are some key differences between bacterial RNAPs and human RNAPs that can help us target bacterial RNAPs with antibiotics. As you may be aware, antimicrobial resistance is a growing and serious problem worldwide. The Centers for Disease Control and the World Health Organization (WHO) have called for bacteriologists to identify more new antibiotics to help combat this problem.

Specifically, the WHO recently identified 12 bacterial organisms that are extremely dangerous and resistant to all or most available drugs. Some of these 12 pathogens contain a specific element within their TL domains that is absent in human RNAPs. My work as well as others' indicates that this element is very important for the activities of the TL, and RNAP, and could therefore serve as a potent and specific drug target. Moreover, the way to kill pathogens like these is to target their essential processes – processes that they need to have functioning for survival. Transcription termination is just one such process that is essential to bacteria; there are many other processes that need to be better understood so drugs can be designed against them. There are many, many more problems to be solved. Will you join the workforce?

Chapter 1:
Introduction

Adapted from:

Ananya Ray-Soni, Michael J. Bellecourt, and Robert Landick. (2016) Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review of Biochemistry*. **85:** 319-347.

Chapter contributions:

A.R.S. wrote all text and constructed Figures 1.4 and 1.6; M.J.B. constructed Figures 1.1, 1.2, 1.3 and 1.5 for the *Annu. Rev. Biochem.* manuscript; R.L. oversaw and edited text in sections 1.1-1.6 for the *Annu. Rev. Biochem.* manuscript.

1.1 Bacterial Transcription Termination: A Bird's Eye View

Understanding the mechanism and sequence determinants of transcription termination by bacterial RNA polymerases (RNAPs) is a long-standing goal in the study of gene regulation. Termination (*a*) prevents the inappropriate expression of downstream genes and interference from antisense transcripts, (*b*) defines RNA 3′ ends, important for precise RNA structures (*e.g.*, small noncoding RNAs) and temporal regulation, (*c*) recycles RNAP for efficient gene expression, and (*d*) minimizes collisions with replication complexes that cause damaging double-strand breaks in the chromosome. Genetically programmed transcription termination in bacteria occurs by one of two pathways: intrinsic termination, requiring only RNAP, RNA, and DNA, or factor-dependent termination, involving proteins that dissociate the elongation complex (EC), such as Rho. Factor-dependent termination can also occur by other means, such as with Mfd, which targets damaged ECs independent of a genetically encoded termination signal.

Intrinsic termination occurs when an RNA secondary structure called a terminator hairpin (T_{hp}; **Figure 1.1**) forms in the exit channel of an RNAP that is positioned on an unstable RNA–DNA hybrid, causing EC dissociation. Rho-dependent termination relies on the adenosine triphosphate (ATP)-dependent RNA–DNA helicase Rho, which binds to and translocates the nascent RNA until it encounters a paused EC and causes the release of RNAP. Rho requires a less-specific sequence at which to dissociate than intrinsic termination, which is important for its essential role in suppressing R-loops, transcription of horizontally acquired genes, and antisense transcription (1-4). Unlike intrinsic or Rho-dependent termination, Mfd-dependent termination is not programmed by specific nucleic acid (NA) sequences. Instead, Mfd recognizes stalled ECs and removes them from the DNA to prevent potentially deleterious collisions with the DNA replication machinery (5). Upon binding the stalled RNAP, Mfd utilizes its ATP-dependent translocase activity to induce forward translocation of RNAP without the

concomitant nucleotide addition (6). This shortens the RNA–DNA hybrid, destabilizing the EC and inducing dissociation of RNAP and the nascent transcript. Mfd then recruits the DNA repair factors UvrA and UvrB to the DNA lesion (7), thus initiating the transcription-coupled repair pathway.

In this work I focus on the mechanism of intrinsic termination as one of the major genetically encoded pathways that affect gene regulation, for which an advanced understanding of RNAP structure developed over the past several years has raised important questions for future research.

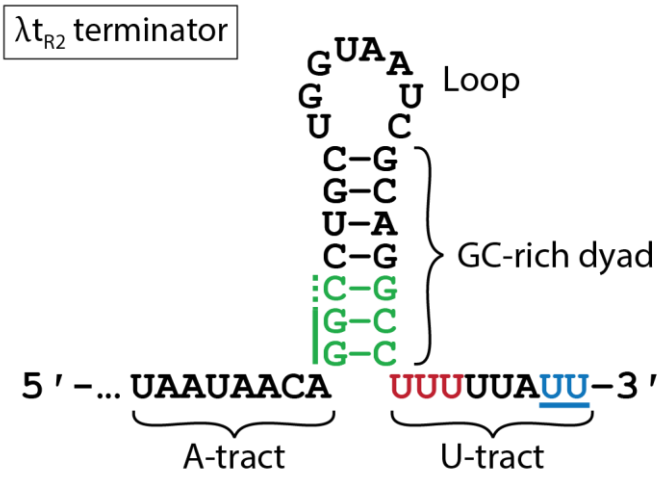


Figure 1.1: A canonical intrinsic terminator.

The model λtR2 terminator is shown with key elements indicated. The *green* bps indicate the 2–3 bps of the T_{hp} stem that pair during the second step of the two-step T_{hp} folding process, T_{hp} extension; *red* U's indicate highly conserved, upstream hybrid U's that are melted upon T_{hp} extension; the blue underlined U's mark the sites of termination. The A-tract is not an essential element of a canonical terminator.

1.2 Elongation Complex Stability and its Destabilization by Termination

1.2.1 Elongation Complex Stability

To understand the dissociation of ECs at terminators, one must first understand the determinants of EC stability. RNAP has evolved to form a remarkably stable, yet dynamic, network of contacts with the NA scaffold in the EC, enabling it to transcribe more than 104 nucleotides (nt) without dissociating from the template. The enzyme protects a 12–14 base pair (bp) transcription bubble, a 9–10 bp RNA–DNA hybrid, approximately 5 nt of newly transcribed single-stranded RNA (ssRNA), and approximately 18 bps of downstream duplex DNA (8-10) **(Figure 1.2)**. The major sources of EC stability are polar and van der Waals contacts between RNAP and the RNA–DNA hybrid backbone in the main channel of the enzyme, augmented by H-bonds to the ssRNA in the exit channel and long-range electrostatic and van der Waals contacts to the downstream DNA (10). Comparison of the structures of the EC versus free RNAP reveals that these contacts are established upon closure of the clamp module and a small rotation of the shelf module relative to the core of RNAP (10-14) **(Figure 1.2A)**. These changes are detectable by disulfide-bond formation, using cysteine pairs strategically engineered into *Escherichia coli* or *Thermus thermophilus* RNAP (15, 16). The closed-clamp state appears to be stabilized, in part, by interactions between downstream DNA and the RNAP clamp and jaw, and, most significantly, by extensive contacts among the folded forms of three switches and the RNA–DNA hybrid and downstream DNA (9, 17-19). The switches (Sw1–5) are five highly conserved polymorphous elements at the base of the clamp that rearrange to enable the clamp to swing open 30◦ relative to the RNAP core, widening the main cleft (17). Compaction of the switches conversely stabilizes contacts between the closed clamp and the NAs in the EC. Specifically, Sw1 makes stabilizing contacts to the template strand in the RNA–DNA hybrid;

highly conserved Sw2 K334 (*E. coli* RNAP numbering) marks the downstream edge of the RNA–DNA hybrid by stabilizing the 90◦ kink in the template strand between +1 and +2; and Sw3 forms a hydrophobic binding pocket for the first ssRNA base upstream of the RNA–DNA hybrid (10, 17, 20) (*i.e,* the –10 RNA base in the post-translocated state; **Figure 1.2C)**. The upstream edge of the hybrid is further stabilized by stacking interactions with the lid (10) **(Figure 1.2C)**. This network of nonspecific interactions creates an EC that is stable up to 70 ◦C and >0.5 M salt (21-23) but still allows for efficient translocation of DNA and RNA through RNAP.

1.2.2 Destabilization of Elongation Complexes by Termination and the Role of Transcriptional Pausing

The challenge of transcription termination is to sufficiently destabilize the highly stable EC at specific genome locations to enable RNAP dissociation from the DNA and RNA. However, termination can occur at a given position only if RNAP enters the termination pathway faster than nucleotide addition moves the DNA template to the next position **(Figure 1.3, step ① *vs.* elongation)**. Consequently, efficient termination is aided by factors that increase the kinetic window at the termination branch point, and it is impeded by factors that increase the elongation rate. This idea can also be explained by a well-established thermodynamic model for the elongation versus termination decision (21, 24, 25). Termination could theoretically occur at any sequence position, where the relative energy barriers to elongation versus termination determine the proportion of complexes that elongate or terminate. Higher energy barriers would result in slower rates of reaction because the ECs must achieve a higher energy state to overcome the barrier and complete the reaction. At non-terminator positions, the energy barrier to termination is thought to be prohibitively high, making elongation the significantly preferred pathway **(Figure 1.4A)**. However, at termination sites, sequence features of the terminator itself destabilize the EC (see below), effectively decreasing the energy barrier to termination and

allowing appreciable levels of termination to occur **(Figure 1.4B)**. Therefore, factors that increase the kinetic window to termination act by increasing the energy barrier to elongation, effectively decreasing elongation rate and creating more time for ECs to enter the termination pathway. Transcriptional pausing is one such mechanism that causes a proportion of ECs to enter a slow-elongating state due to an increased barrier to elongation, reversibly halting ECs along the template for durations of seconds to minutes (26, 27). Pausing at termination sites therefore increases the kinetic window for entry into the termination pathway **(Figure 1.4B)**, and is considered a necessary first step for efficient termination (22, 28) **(Figure 1.3, step ①)**.

Transcriptional pausing is defined as a kinetic branching from the main elongation pathway where the EC enters a distinct conformational state that inhibits nucleotide addition. Pause escape occurs when the active site rearranges back into an active configuration and resumes nucleotide addition. An EC may isomerize into an off-pathway, paused state at any position on a template (29), but these excursions from the elongation pathway, into so-called elemental pauses, are preferentially triggered and stabilized by specific sequences in the RNA–DNA hybrid, the non-template strand (NT-DNA), downstream DNA, and active site nucleotides (20, 30, 31). An analysis of the crystal structure of RNAP bound to elemental pause sequences revealed a widened main cleft, with the switch regions swung away from the NAs, breaking some contacts with the RNA–DNA hybrid (20). This rearrangement partially reverses changes seen during the transition from free RNAP to the EC, as observed in previously reported crystal structures (10, 12, 17). Because complete clamp opening does not appear to occur in the elemental paused EC when probed by disulfide bond formation (16) (P.H. Hein, M.J. Bellecourt, D. Nayak, R.A. Mooney & R. Landick, *unpublished findings*), the sequences of the elemental pause likely loosen the clamp contacts sufficiently to allow transient clamp opening, which appears to become stabilized in the open position by crystal packing forces (16). The rearranged structure also features a kink in the active site-proximal bridge helix (BH) **(Figure 1.2B),**

which obstructs entry of the next template base and cognate nucleoside triphosphate (NTP) into the active site, preventing nucleotide addition.

The lifetimes of elemental pauses can be increased by two known mechanisms: (*i*) by the formation of RNA secondary structures in the RNA exit channel (called pause hairpins) in hairpin-stabilized pausing, or (*ii*) by backward translocation of the EC on the DNA template in backtrack pausing. Hairpin-stabilized pausing occurs when complementary sequences in the nascent transcript pair to form a pause hairpin in the RNA exit channel 11 or 12 nt from the RNA 3′ end and, thus, stabilize the open clamp state (15, 16). Formation of the hairpin is likely enabled by the loosening of the clamp in the elemental paused EC (20). Hairpin-stabilized pausing also requires interactions between the pause hairpin and the flap domain (32, 33) **(Figure 1.2A)**, and it can be aided by binding of NusA (34-36) (see section **1.2.3 Impact of Transcription Factors on Termination**), which contacts both the flap tip and the pause hairpin (37-39).

Backtracking is a mechanism by which RNA and DNA reverse translocate through RNAP, such that the 3′-nascent RNA extrudes into the secondary channel of RNAP **(Figure 1.2B)**, consequently removing the RNA 3′ end from the active site. It can occur as a proofreading mechanism in response to misincorporated nucleotides or when the EC is positioned over a weak RNA–DNA hybrid and can gain stability by backtracking to a position where the RNA–DNA hybrid is stronger (40-42). Backtracked pauses can be resolved by forward translocation to restore the 3′ end to the active site, by the intrinsic transcript cleavage activity of RNAP to create a new 3′-OH, or by GreA- or GreB-stimulated cleavage, in which the Gre transcription factor binds in the secondary channel and aids cleavage of the backtracked transcript in the RNAP active site.

Intrinsic terminators encode a T_{hp} that is structurally similar to a pause hairpin **(Figure 1.1)**, but whether this T_{hp} causes a hairpin-stabilized pause is still an open question (see section

1.3.2 Terminator Hairpin Nucleation). Intrinsic terminators also encode a uracil (U)-rich RNA–DNA hybrid that is related to elemental pause sequences (30) and induces pausing (22). The weak rU–dA bp (43) of the U-rich hybrid could cause backtracking in the absence of the T_{hp}. However, backtracking is an unlikely predecessor to termination, as it would inhibit T_{hp} formation. In fact, one role of the T_{hp} is thought to be the prevention of backtracking in order to lock the EC on the weak hybrid that is favorable for dissociation (23, 44). Although the precise nature of the pre-termination pause is still uncertain, the U-rich hybrid at intrinsic terminators may both induce an elemental pause and instigate the initial RNAP conformational change necessary for termination by loosening protein–NA contacts in the EC and making the complexes susceptible to subsequent inactivation and dissociation. A better description of these early events in the termination pathway is vital to understanding the role of pausing in intrinsic termination.

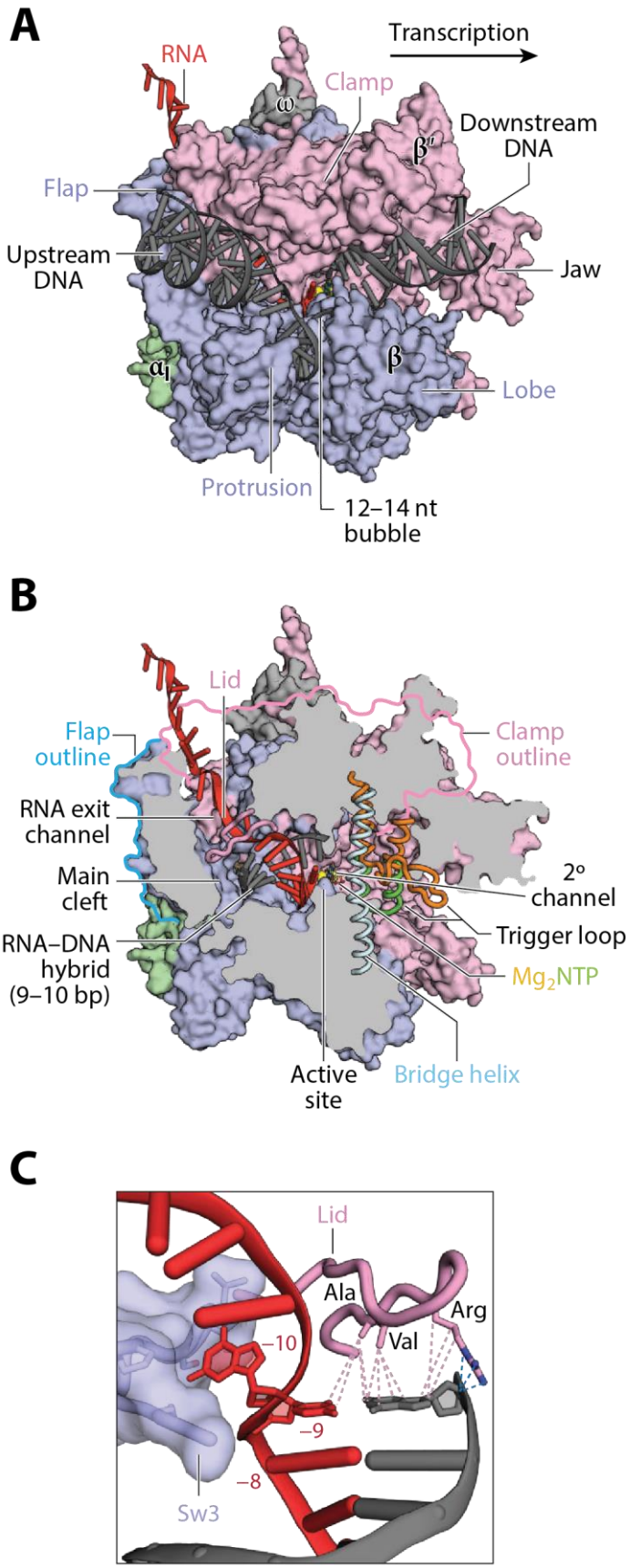
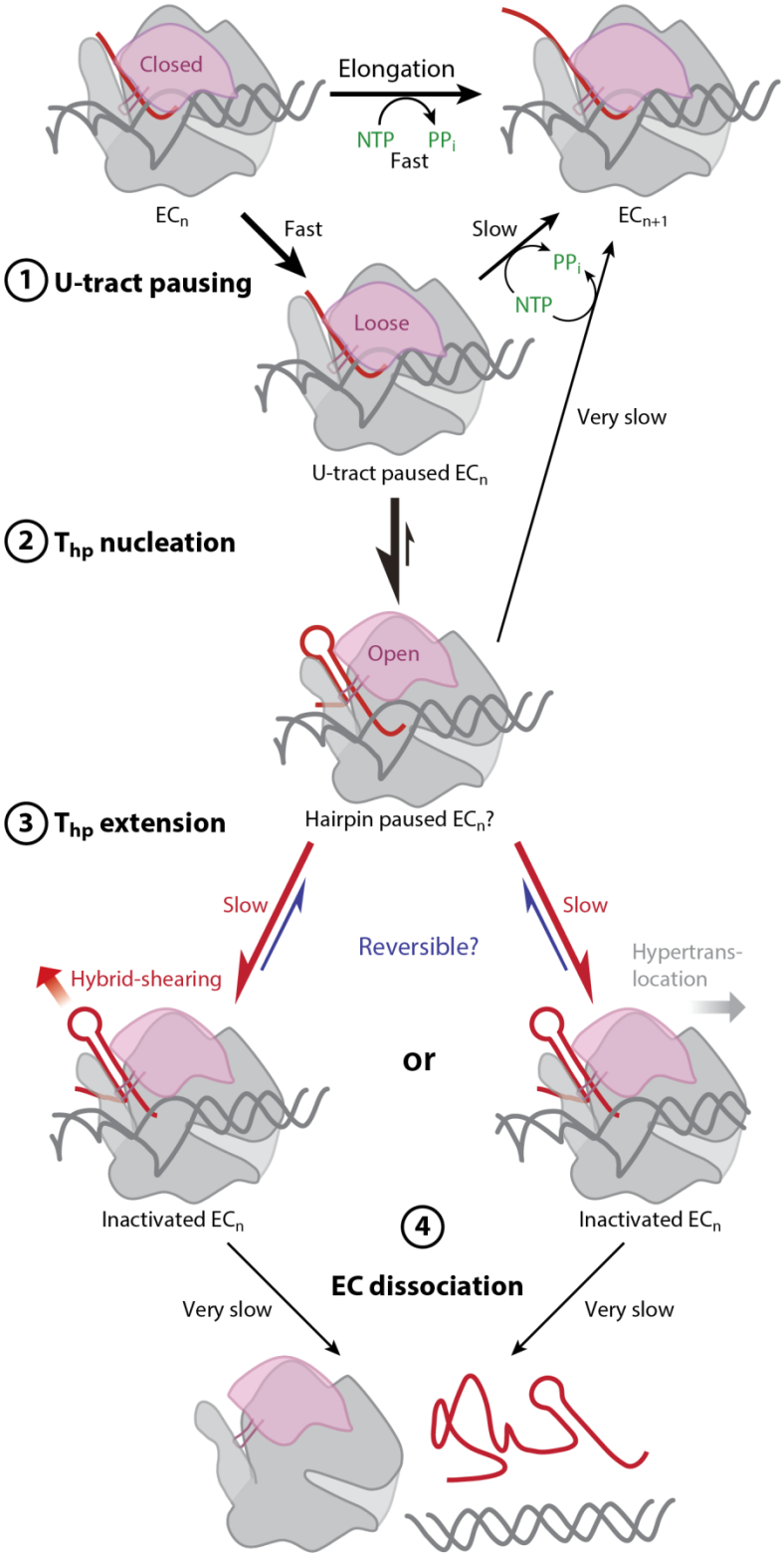

A
Transcription
RNA
Clamp
ω
β′
Downstream DNA
Flap
Upstream DNA
Jaw
α1
β
Lobe
Protrusion
12–14 nt bubble
B
Lid
Flap outline
Clamp outline
RNA exit channel
Main cleft
2° channel
Trigger loop
RNA–DNA hybrid (9–10 bp)
Mg2NTP
Active site
Bridge helix
C
Lid
Ala
Arg
Val
–10
–9
–8
Sw3

Figure 1.2: Structural features of the EC.

(A) Model of the EC structure based on a crystal structure of the *Thermus thermophilus* EC (PDB 2O5J); subunits and features of interest are indicated. **(B)** Sliced view of the EC, revealing key features of the RNA exit channel, main cleft, secondary (2◦) channel, and active site. Both folded (*green*; PDB 1IW7) and unfolded (*orange*; PDB 2O5J) conformations of the trigger loop (TL) are depicted. **(C)** The interface between the RNA exit channel and main cleft, with RNA numbering representing an EC in the post-translocated register (PDB 2O5I). Dashed *pink* and *blue* lines indicate, respectively, the stacking and ionic interactions between the lid and the –9 RNA (10) and DNA (*gray*) bp that mark the upstream edge of the RNA–DNA hybrid. The blue surface depicts Sw3 interacting via its hydrophobic binding pocket with the first single-stranded RNA base (–10).


Closed
Elongation
NTP
PPi
Fast
ECn
ECn+1
Fast
Slow
PPi
NTP
1 U-tract pausing
Loose
U-tract paused ECn
Very slow
2 Thp nucleation
Open
Hairpin paused ECn?
3 Thp extension
Slow
Slow
Reversible?
Hypertrans-
location
Hybrid-shearing
or
Inactivated ECn
Inactivated ECn
4
EC dissociation
Very slow
Very slow

<u>Figure 1.3:</u> The mechanism of intrinsic termination.

A reaction scheme is shown with all known steps in the termination pathway, as well as competing steps that result in terminator read-through. The EC is depicted schematically in an orientation similar to that in **Figure 1.2A**, with the RNAP lid and mobile clamp module (*pink*), the secondary channel (*light gray*), RNA (10), and DNA (*dark gray*). The mechanism is shown in four steps: ① pausing at the 3′ end of the U-tract, ② T_{hp} nucleation, ③T_{hp} extension and possible EC inactivation, and ④ EC dissociation. Two alternative mechanisms for T_{hp} extension and EC inactivation are depicted. *Red* arrows depict the proposed EC inactivation step, which appears to be irreversible *in vitro*; *blue* arrows depict possible EC reactivation, which may occur *in vivo*, possibly facilitated by transcription factors. The EC subscripts n and n + 1 denote the transcript lengths.

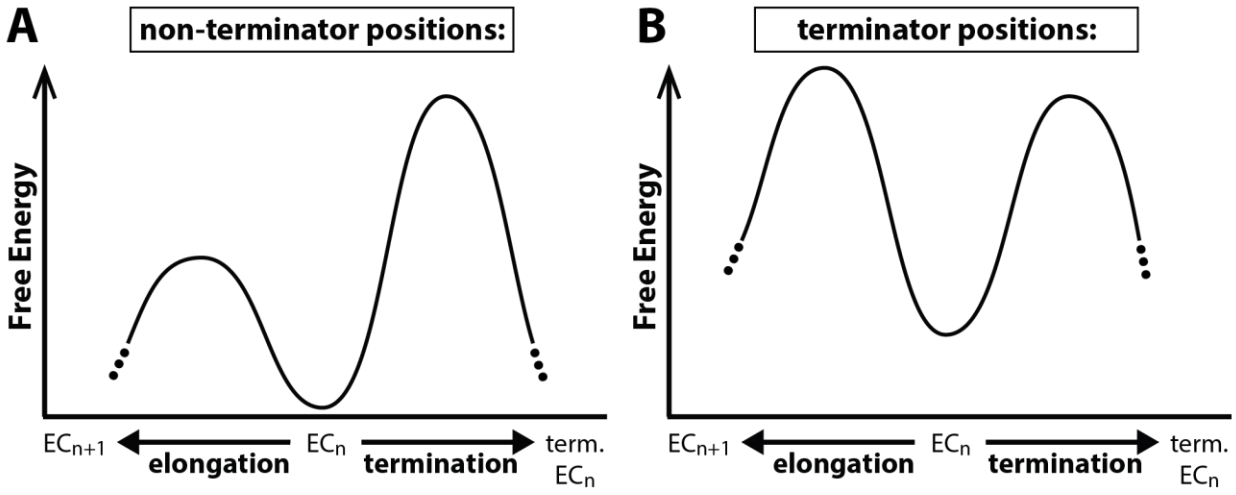


Figure 1.4: The thermodynamic model for termination (21, 24, 25).

The relative energy barriers to elongation (left) versus termination (right) determine the probability of either event. At non-terminator positions **(A)** the barrier to elongation is significantly lower than the barrier to termination, whereas at terminator positions **(B)**, the weak bps of the U-tract destabilize the RNA–DNA hybrid, and therefore the EC, effectively decreasing the energy barrier to termination. This reduced energy barrier allows appreciable termination efficiencies.

1.3 Impact of Transcription Factors on Intrinsic Termination

Most mechanistic studies have been performed in cell-free environments to leverage the manipulable nature of *in vitro* experiments. Although this focus enables the investigation of the termination mechanism in the absence of extrinsic factors, it may miss more complex layers of regulation that occur *in vivo*. Multiple factors in the cell may affect termination, including transcriptional factors that can bind ECs and modulate their stability at various steps along the termination pathway.

Transcriptional factors that bind the EC can perturb termination efficiency (TE) by modulating elongation rate, pause entry or pause lifetime, nascent RNA structure formation, or EC stability. For example, the universally conserved NusG protein has been shown to promote forward translocation — and, hence, elongation — of *E. coli* ECs by binding the clamp near the upstream fork junction and likely helping upstream DNA reanneal (45). In other bacterial strains, however, NusG has been shown to stimulate hairpin-stabilized pausing (in *Bacillus subtilis*) (46, 47) or to enhance intrinsic TE (in *Mycobacterium bovis*) (48), possibly by inhibiting rather than stimulating translocation as observed in *E. coli* (49). The NusG paralog, RfaH, inhibits termination through a distinct mechanism; RfaH stabilizes the closed clamp by binding the clamp helices and a loop in the lobe domain, thereby impeding the clamp opening required for both pausing and termination (39, 50-52). Finally, the exit channel binding factor, NusA, can stimulate hairpin-stabilized pausing as well as termination at weak intrinsic terminators by stimulating hairpin formation and stabilizing the hairpin's interaction with the flap domain, effectively decreasing the rate of pause escape (32, 35, 36, 38, 39, 53, 54). However, NusA can paradoxically also act as an anti-termination factor that inhibits intrinsic termination when complexed with other transcription and translation factors (55). Therefore, gaining a complete picture of termination *in vivo* will require an understanding of the abundance of

various transcriptional factors in the cell and how they impact the termination mechanisms identified in purified ECs.

1.4 Mechanism of Intrinsic Termination

At intrinsic terminators, DNA and RNA elements at the site of termination direct a series of steps that, after RNAP transcribes kilobases of gene information, cause the enzyme to (*a*) terminate transcript elongation at a discrete location spanning 2–3 nt and (*b*) release from the chromosome and transcript. The canonical intrinsic terminator sequence **(Figure 1.1)** consists of two essential elements: a guanine and cytosine (GC)-rich dyad that forms the T_{hp} 7–8 nt from the transcript's 3′ end, followed immediately by a 7–8 nt U-rich tract, of which the first three U's are the most highly conserved (56). In the next section, we describe the steps by which this termination signal directs EC destabilization: ① pausing at the 3′ end of the U-tract, ② T_{hp} nucleation, ③T_{hp} extension and possible EC inactivation, and ④ EC dissociation **(Figure 1.3)**.

1.4.1 Pausing at the U-tract

As RNAP transcribes the final nucleotides of the terminator U-tract, it pauses **(Figure 1.3, step ①)**, favoring the termination pathway in the kinetic competition between elongation and termination (22, 28) **(Figure 1.4B)**. Decreased elongation rates have been shown to favor TE (57), as pausing creates a window of time in which the T_{hp} can form in the exit channel and initiate the cascade of events that cause termination (22). Mutation of the 3′-terminal U in the U-tract of the λt_{R2} terminator, which has been shown to abrogate pausing at the termination site, also eliminated termination (22). Therefore, pausing at the U-tract must be crucial for the

formation of the T_{hp}. In addition to the U-tract, NA sequences in the active site, downstream DNA channel, and the RNA exit channel must have a role in establishing the pause and determining pause efficiency and duration (26, 34, 58, 59). These sequences thus contribute to the maximal TE by determining the proportion of complexes that enter the elemental U-tract pause and the fraction of those complexes that are able to form the T_{hp} prior to pause escape **(Figure 1.3)**.

1.4.2 Terminator Hairpin Nucleation

Pausing of the EC at the end of the 7–8 nt U-tract stalls RNAP for a sufficient time to allow the T_{hp} to form in the exit channel **(Figure 1.3, step ②)**. Two separate studies have shown that a 7 bp T_{hp} does not nucleate until the EC transcribes 7 nt downstream from the T_{hp}-encoding sequence (22, 60). This is consistent with evidence that RNAP protects 14 nt of RNA (approximately 9 nt in the hybrid and 5 nt in the RNA exit channel) (8), with a preference for a $T_{hp} \geq 7$ bps at canonical terminators (56, 61). Termination can also occur at low efficiency 8 nt downstream from a 6 bp T_{hp} (35), with hairpin nucleation likely occurring 8 nt downstream from the T_{hp} sequence. Conversely, a T_{hp} with a longer stem or with a stabilized loop may nucleate before the EC reaches this position so that programmed pausing at the end of the U-tract becomes dispensable for T_{hp} nucleation (62), although this relationship has not been studied systematically. However, in the absence of a U-tract, hairpin formation does not result in efficient termination (62), verifying that both the T_{hp} and U-tract are essential for the termination mechanism.

Upon T_{hp} nucleation, all but the bottom 2–3 bps of the T_{hp} rapidly pair (60), aided by the flap domain tip **(Figure 1.2)** in the case of a weak T_{hp} (32). This first step in T_{hp} formation produces a structure remarkably similar to a pause hairpin, and likely stabilizes the pre-termination pause prior to complete extension of the T_{hp} (59, 60).Whether the partially formed

T_{hp} produces a hairpin-stabilized pause was previously unresolved by direct experimental evidence. However, our findings strongly support the existence of such a hairpin-stabilized intermediate on the termination pathway in response to the partially formed T_{hp} (see section **2.4.4 The TL promotes both elemental and hairpin-stabilized pausing at an intrinsic terminator**).

Historically, some studies have described effects of sequences far removed from a terminator on TE, and they have considered various explanations for the phenomenon, including the induction of persistent alternative conformations of the EC that resist termination (63, 64).However, competing RNA structures formed with the upstream transcript can inhibit T_{hp} formation by sequestering the upstream arm of the T_{hp} prior to hairpin nucleation (60, 65). These more recent demonstrations of the effects of alternative RNA folding highlight the importance of ensuring that comparisons of terminators are made in the context a constant sequence context — preferably in the context of a set of multiple defined sequences — to avoid the potentially complicating effects of RNA folding.

1.4.3 Terminator Hairpin Extension

1.4.3a Consequences and energetic constraints

The structure of the EC suggests that the pairing of the bottom 2–3 bps of the T_{hp} stem would require displacement of the –10 RNA base from its Sw3 binding pocket and the unstacking of the lid from the –9 RNA–DNA bp **(Figure 1.2C** and **Figure 1.5)**. Consistent with this idea, Lubkowska et al. (10) found that T_{hp} formation occurs in two steps for the λt_{R2} T_{hp} (60) **(Figure 1.1)**. In the first step, all but the bottom 2 bps formed upon T_{hp} nucleation 7 nt downstream from the T_{hp}-forming sequences. In the second step, the bottom 2 bps of the T_{hp} paired once the EC transcribed the eighth nucleotide downstream from the T_{hp} sequence. A single-molecule force-clamp study consistently observed two separate steps for T_{hp} nucleation and complete

extension of the T_{hp} (30). The two-step process of T_{hp} folding indicates that, consistent with our structural prediction, interactions that limit complete T_{hp} extension create an energetic barrier for this step. That is, the need to disrupt the interactions between the lid and hybrid and between Sw3 and the –10 RNA base likely create all or part of the energy barrier to T_{hp} extension **(Figure 1.3, step ③)**.

Termination is frequently observed 7 nt downstream from the T_{hp} on endogenous terminators (such as in **Figure 1.1**) and requires that T_{hp} nucleation and extension occur at the same template position (22, 66), which is inconsistent with the findings of the T_{hp}-folding study described above (60). It is possible that Lubkowska et al.'s use of stabilized hybrids (60) to enable the EC to walk downstream from the T_{hp} sequence precluded completion of the T_{hp} until the eighth nucleotide was added (22, 23). Upstream melting of the stabilized hybrid would be disfavored in this case due to the lack of U's in the hybrid (see next paragraph), thereby making the energetic barrier to T_{hp} extension even higher and preventing completion of the T_{hp} at this position. However, the two-step folding pathway of T_{hp} nucleation, followed by energetically expensive T_{hp} extension, likely exists for all cases— independent of terminator structure and the exact position for each step—due to the energetic barrier created by the interactions between the lid and hybrid and Sw3 and –10 RNA base.

T_{hp} extension causes 3–4 bps of the upstream RNA–DNA hybrid to melt (22, 23) **(Figure 1.1, red U's** and **Figure 1.5)**, thus contributing to the destabilization of the EC. This hybrid melting step is essential for T_{hp} extension and termination, and is consistent with the high degree of sequence conservation for U's at the first three positions in intrinsic terminator U-tracts (22, 23, 56). Thus, from our earlier observations, we infer that the kinetics and position of T_{hp} nucleation are dictated by the length of the T_{hp} stem and the structure of the T_{hp} loop; they determine the position at which the loop and the GC-rich dyad emerge from the RNA exit channel, and the length of solution-exposed RNA needed to nucleate T_{hp} formation. However,

the template position at which T_{hp} extension can occur is likely determined by the relative strengths of the bottom 2–3 bps of the T_{hp} compared with the upstream 3–4 bps of the hybrid **(Figure 1.1, green bps** *vs.* **red U's)**. These relative strengths dictate the energetic cost of melting the upstream hybrid and breaking the interactions between the lid and hybrid and Sw3 and –10 RNA base, compared with the energetic gain from pairing of the bottom of the T_{hp}.

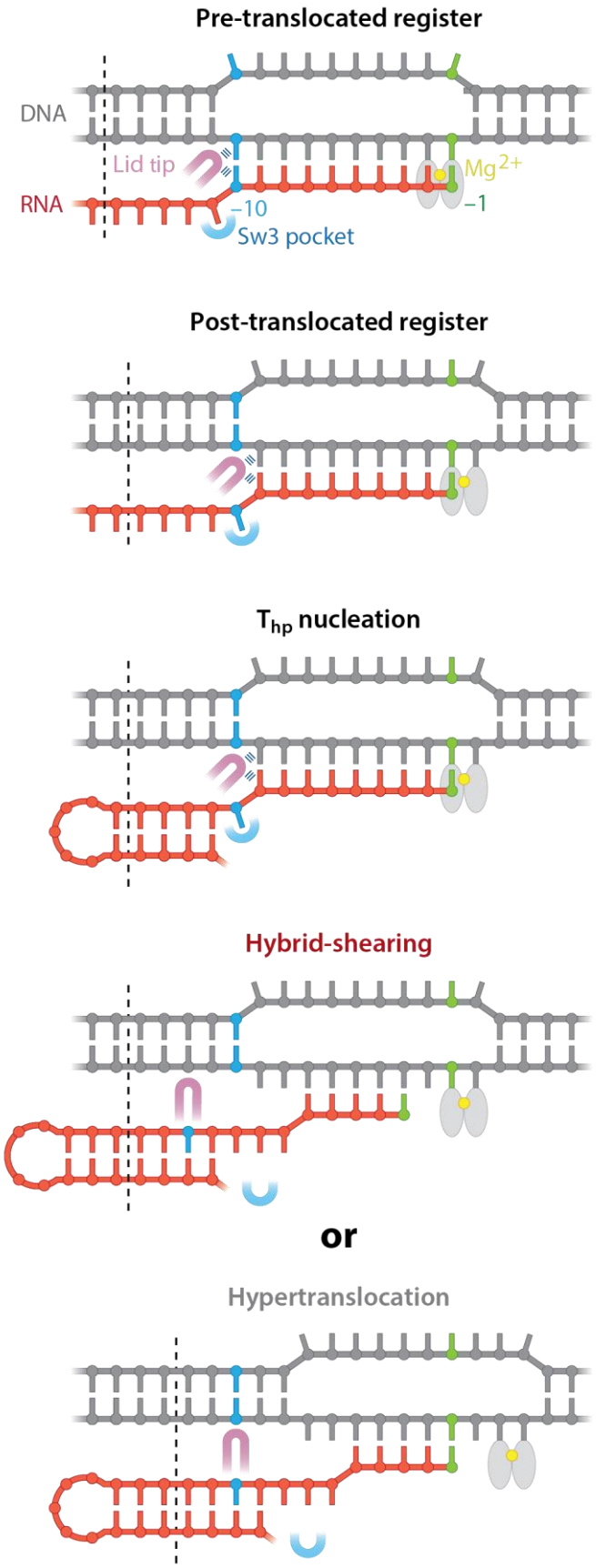
Upstream hybrid melting results from an incompatibility of a complete T_{hp} and a 9–10 bp RNA–DNA hybrid, but the structural basis for this incompatibility is unclear from EC crystal structures. It would appear entirely feasible that in an EC with a fully open clamp, the exit channel and main cleft could accommodate a relatively continuous (albeit modestly kinked) bipartite duplex of the extended T_{hp} stem and RNA–DNA hybrid, perhaps stabilized by a stacking interaction between the lid, the hybrid, and the T_{hp} (67).Nonetheless, the demonstrated hybrid melting effect indicates that steric constraints of some sort must limit the possible NA conformations, and suggests that currently undefined RNAP structural rearrangements are likely to accompany this step.

1.4.3b Proposed models

Several different models have been proposed to describe how the completed T_{hp} affects the EC. These models involve (*a*) hypertranslocation of the EC 2–4 bps downstream from the termination point, without concomitant nucleotide addition (28, 65, 68), or (*b*) shearing or slippage of the RNA from the RNA–DNA hybrid by a lifting or rotational wrenching of the T_{hp} from the exit channel (23, 65) **(Figure 1.3** and **Figure 1.5)**. A single-molecule force-clamp experiment revealed that both mechanisms may be possible, depending on the sequence of the terminator (65). Larson et al. (65) found that a terminator with an imperfect U-tract (a U-tract interrupted by one or more GC bps), t_{500}, was rendered more efficient by RNA pulling and impeded by a hindering force on DNA translocation (consistent with hypertranslocation),

whereas two terminators with near-perfect U-tracts (7 or more tandem rU-dA bps and at most one rA-dT bp), t_{his} and λt_{R2}, were also aided by RNA pulling, but unaffected by force on the DNA (consistent with hybrid-shearing). Taken together, these data suggest that either mechanism can occur, with hypertranslocation favored on imperfect U-tracts and hybrid-shearing favored on weaker near-perfect U-tracts. Consistently, Peters et al. (69) have found that in a set of 100 *E. coli* terminators, the downstream DNA is enriched for adenine and thymine (AT)-rich sequences at positions +10 to +12 after imperfect U-tracts, but not after near-perfect U-tracts, indicating a need for the melting of downstream DNA in hypertranslocating terminators, but not in hybrid-shearing terminators. These results support a composite hybrid-shearing–hypertranslocation model in which the mechanism employed is determined by the relative energetic cost of shearing the U-tract hybrid versus melting downstream DNA (65, 70) **(Figure 1.3, step ③)**. Extensive characterization of various terminator structures *in vivo* has found that the strongest terminators have both near-perfect U-tracts and AT-rich downstream DNA (61). This finding also supports the idea that both hybrid-shearing and hypertranslocation can occur and that TE is highest if both mechanisms are energetically favorable. We note that while these two mechanisms for T_{hp} extension have been elucidated, there likely exist many more as-of-yet undiscovered routes for this step, which may involve these or other NA rearrangements, but almost certainly involve RNAP conformational changes. We address one RNAP module likely involved in restructuring ECs during this step on the termination pathway in **Chapter 2** (see also section **1.7.1 The Trigger Loop (TL) and Questions of Interest**).

T_{hp} extension thus appears to be the first step that distinguishes paused complexes from termination-competent intermediates. This step results in NA rearrangements in the EC, likely with concerted rearrangements of RNAP crucial for efficient termination. Some evidence suggests that these rearrangements may irreversibly inactivate nucleotide addition within the EC and, consequently, commit the EC to the termination pathway.


Pre-translocated register
DNA
Lid tip
RNA
Mg2+
–10
–1
Sw3 pocket
Post-translocated register
Thp nucleation
Hybrid-shearing
or
Hypertranslocation

<u>Figure 1.5:</u> Nucleic acid scaffold conformations at different stages of transcript elongation and termination.

Depicted states include: pre-translocated EC, post-translocated EC, after T_{hp} nucleation, and after EC inactivation and T_{hp} extension by the hybrid-shearing or hypertranslocation mechanisms. The active site is denoted by *two gray ovals*; the tip of the lid by a *purple loop*; and Sw3 by a *cyan semicircle*. Stacking interactions between the lid and the upstream hybrid are depicted by *blue dashed lines*. The position of the upstream end of RNAP relative to RNA is represented by a *black dashed line*. *Green* and *blue* nucleotides represent, respectively, the –1 and –10 bps.

1.4.4 Elongation Complex Inactivation and Commitment to Termination

1.4.4a Evidence for an inactivation step

Although several studies have modeled TE as being determined by direct kinetic competition between elongation/pause escape and EC dissociation (21, 24, 25), there is some evidence — as well as uncertainty — that an irreversible EC inactivation step occurs in the termination pathway prior to EC dissociation. In a single-molecule termination assay, Yin et al. (71) found that a terminal dwell occurred at the t_{his} terminator prior to transcript dissociation only for complexes that terminated but not for complexes that read through the terminator. This result suggests that nucleotide addition is irreversibly halted for terminating ECs in a step distinct from dissociation. In contrast, although a 1 s terminal dwell was measured for the t_{500} terminators in a force-clamp experiment, no terminal dwell was observed for the t_{his}, or λt_{R2} terminators (65). It is possible to explain this discrepancy if the assisting force used in the force-clamp study increased the rate of EC dissociation, masking a terminal dwell on t_{his} and λt_{R2} prior to dissociation. The terminal dwell in the Yin et al. study (71), then, implies that ECs at the termination site can exist in two distinct states: one that is able to recover from the pre-termination pause and continue elongation past the terminator, and another, inactivated state that is incapable of nucleotide addition but remains bound to the hybrid at the termination site until EC dissociation occurs at a slower rate. Moreover, fits to the force-clamp experiments described above (65) did require the inclusion of a parameter that describes the distance moved by the DNA or RNA prior to dissociation — without concomitant nucleotide addition — suggesting that ECs are functionally inactivated prior to dissociation. This step, requiring inactivation of the RNAP active site, thus represents the commitment of ECs to the termination pathway.

Notably, Nudler and coworkers (22) also reported the existence of an irreversibly "trapped", or inactivated, termination complex at the λt_{R2} terminator prior to EC dissociation,

although this interpretation has been disputed (72). The rate of transcript release from ECs stalled at a termination site *in vitro*, after T_{hp} extension and upstream hybrid melting, occurs on the order of many minutes (23), too slow to compete with pause escape, which can occur on the order of seconds even when stabilized by a pause hairpin (27).Given this slow rate of EC dissociation, an irreversible inactivation step seems kinetically required to prevent pause escape and the resumption of elongation. As we show in **Appendix A**, however, artificially stalled complexes may not reproduce dissociation rates that are highly representative of physiological conditions (see section **A.3.1 Termination efficiency and stalled EC dissociation rate measurements are not adequate to study the termination mechanism**). The existence of such a step therefore still requires experimental validation with the help of an assay that can rigorously measure the rates of EC inactivation (*i.e.* the rate at which ECs become resistant to elongation) and EC dissociation in the absence of assisting or hindering force, and determine if they are distinguishable.

1.4.4b Proposed mechanisms

EC inactivation may occur through the restructuring of the active site. All proposed models for commitment posit T_{hp} extension as the driving force for downstream hybrid rearrangements that result in EC inactivation (22, 23, 65). In one model, T_{hp} formation is proposed to be sufficient to inactivate ECs due to the substantial destabilizing effect of upstream hybrid melting (22). Here, the RNA 3′ end remains in the active site until dissociation occurs, and allosteric inhibition is proposed to be responsible for EC inactivation via invasion of the T_{hp} into the main channel of RNAP (thus termed the “hairpin-invasion model”) (73). However, the low-salt conditions used in these studies have been shown to allow dissociation and subsequent rebinding of the RNA in the main cleft, allowing for alternative interpretations of these results (72). Another model for EC inactivation is suggested by the composite hybrid-shearing–

hypertranslocation model for T_{hp} accommodation; both hybrid-shearing and hypertranslocation would directly inactivate the EC via removal of the RNA 3′ end from the RNAP active site (65) **(Figure 1.5)**. Further study is required to understand the mechanism of EC inactivation and whether the NA rearrangements described by any of these models are (*a*) sufficient to prevent nucleotide addition and (*b*) irreversible.

1.4.4c Reconciling observed effects of A-tracts on termination efficiency

Mechanisms that invoke an irreversible commitment step to termination require that all regulation of TE occurs at a step prior to the inactivation of the EC; by definition all committed ECs should terminate, so events occurring after irreversible inactivation cannot affect the proportion of ECs that terminate. However, a recent study suggested that events occurring after full T_{hp} formation can affect TE *in vivo* (61). Voigt and co-workers (61) found that A-tracts immediately upstream of the T_{hp} **(Figure 1.1)** could increase TE, likely by pairing with the U-tract because complementarity of the A-tract and U-tract was required to observe the effect. A–U-tract pairing would extend the T_{hp} duplex and completely eliminate the RNA–DNA hybrid. Therefore, this pairing must occur subsequent to T_{hp} extension and the presumed EC inactivation step. This logic implies that, *in vivo*, T_{hp} extension must be reversible and that A–U-tract pairing can compete with EC reactivation **(Figure 1.3, reverse of step ③)** by accelerating the dissociation of RNA from the EC **(Figure 1.3, step ④)**.

Interestingly, this A-tract-dependent enhancement of TE has not been observed *in vitro* (35). One interpretation is that in purified ECs, the rate of EC reactivation after T_{hp} formation is so slow relative to the dissociation rate that it is functionally meaningless, rendering any downstream A–U-tract pairing effect irrelevant to TE. However, cellular elongation factors, such as NusA and NusG, that interact with the EC may be capable of binding the inactivated intermediate and aiding EC reactivation. This would allow competition between dissociation

and EC reactivation **(Figure 1.3, step ④ *vs.* the reverse of step ③)**, thus allowing the acceleration of dissociation by A–U-tract pairing to affect TE. Potential mechanisms of transcription-factor-mediated EC reactivation include (*a*) NusA-mediated stabilization of the flap–T_{hp} interaction in the inactivated complex (32) or (*b*) NusG-mediated stabilization of the clamp–DNA interaction. Either interaction could facilitate EC reactivation by affecting NA pairing patterns in the T_{hp}, the upstream hybrid, and the upstream DNA fork junction. We have attempted to interrogate whether NusA facilitates an effect of A-tracts on TE *in vitro* (see **Chapter 3** and section **1.7.3 Investigating the Role of NusA in the Intrinsic Termination Pathway**). However, preliminary tests suggest that NusA may not be responsible for this effect. Further *in vitro* testing of the effect of A-tracts on TE and EC dissociation in the presence of NusA, NusG, and other EC-stabilizing factors is needed to determine whether these factors sufficiently stabilize termination intermediates to make an otherwise irreversible EC inactivation step reversible.

1.4.5 Elongation Complex Dissociation

The final step in the termination pathway is the dissociation of RNAP from both the chromosome and nascent transcript **(Figure 1.3, step ④)**. Several factors likely contribute to dissociation. The RNA–DNA hybrid is significantly weakened by the rU–dA bps of the terminator U-tract (43), and it is destabilized further by the shortening of the hybrid after upstream hybrid melting (22, 23). Komissarova and co-workers (23) demonstrated that a 5 bp hybrid, representative of the residual hybrid length retained by the terminator after the upstream hybrid melts **(Figure 1.5)**, almost completely recapitulates the dissociation rate of a termination complex. T_{hp} extension must disrupt the interactions of Sw3 and the –10 RNA base and the lid and –9 hybrid (described above) **(Figure 1.2C)**, and it likely also disrupts the H-bonds between ssRNA and the exit channel (10). It is unclear whether any of these contacts

could be re-established upon hypertranslocation or hybrid-shearing **(Figure 1.3** and **Figure 1.5)** and the consequent repositioning of the RNA with respect to RNAP.

An open-clamp conformation with loosened contacts between the RNA–DNA hybrid and switch regions also has been proposed to favor EC dissociation (13, 14, 17, 20). Consistent with this idea, results of studies using disulfide-crosslinking to probe the clamp state indicate that an open-clamp/rotated-shelf state is favored by duplex formation in the exit channel (15, 16), suggesting that T_{hp} formation may stabilize the open clamp and weaken contacts between the NA scaffold and RNAP.

Other RNAP rearrangements have been proposed to aid EC dissociation, based on the reduced stabilities of ECs deleted for the lid, rudder, or flap modules (32, 36). However, the conformational rearrangements of RNAP required to dissociate a termination complex remain poorly characterized. Further, the order of RNA and DNA release remains unknown. We have attempted to characterize the order of RNA versus DNA release from stalled ECs in **Appendix A** and observed faster release of RNA than DNA. However, as described above, dissociation rates measured from stalled complexes will need to be verified by other means. Moreover, it has yet to be determined whether this order of release is obligate, stochastic or affected by terminator sequences.

1.5 The Role of RNA Polymerase Conformational Changes in Intrinsic Termination

A surprisingly understudied aspect of the termination mechanism is the contribution of conformational changes of RNAP to the energetics of each step in the termination pathway. Are the steps of termination governed principally by NA folding and translocation dynamics, or are conformational changes of RNAP also necessary to inactivate and dissociate the enzyme? Toulokhonov et al. (32) found that the tip of the flap module, which comprises the outer wall of the RNA exit channel **(Figure 1.2)**, aids nucleation of weak T_{hp}s; the zinc-finger was similarly suggested to be involved in aiding hairpin formation (73). Restructuring of the clamp (13-16) and lid (67) have also been proposed to be involved in EC dissociation, although the contribution of these domains to termination has yet to be shown directly. Do other modules also have a role in the pathway? What inactivates the EC active site at terminators? Is the removal of the RNA 3′ end from the active site (as expected by the hypertranslocation and hybrid-shearing models; **Figure 1.5**) irreversible on the timescale of dissociation, or is active site restructuring at the protein level also essential for efficient termination?

As discussed above, the switch regions are thought to couple clamp closure to NA scaffold binding, and, conversely, to promote efficient termination by destabilizing contacts to the scaffold upon clamp opening (17, 20). Switch rearrangements are therefore likely to be involved in EC dissociation. Mutational analyses of RNAP have identified several mutations in the switch regions that result in altered termination efficiencies (74, 75), consistent with a role for the switch regions in the termination mechanism — presumably by mediating clamp opening. This effect on TE *in vitro* implies that switch remodeling occurs upstream of EC inactivation **(Figure 1.3, step ③)**, perhaps by clamp loosening that aids both the U-tract elemental pause and T_{hp} formation (20) **(Figure 1.3, steps ① and ②)**.

In the hairpin-invasion model described above (73), Nudler and co-workers suggest that folding of the active site–proximal trigger loop (TL; **Figure 1.2B**) is also required to form a termination-competent complex, possibly by contributing to inactivation of the active site **(Figure 1.3, step ③)**. The kinking of another active site–proximal element, the BH, was been observed in the crystal structure of the elemental pause (20). A similar conformation of the BH in the inactivated termination complex is an attractive possibility, as it would obstruct templated substrate entry into the active site and elongation past the termination position (20). The role of these modules in termination has yet to be shown conclusively, and because rates of pause escape (26, 27) are fast relative to the dissociation rates observed from actively transcribing ECs by Gelles and co-workers (71), additional stabilization of the inactivated enzyme state may be required for appreciable TE.

Despite the identification of RNAP alterations that impact EC stability and TE, the state of these potentially dynamic modules in termination complexes remains almost completely uncharacterized. A major future challenge will be to isolate and characterize termination intermediates. Although existing biochemical methods and X-ray crystallography have not succeeded in this task, cryo-electron microscopy (cryo-EM) methods are improving rapidly (76), and determination of the structures and populations of these complexes in the termination pathway is likely to become possible in the near future. In this work I present an assay I have developed that may complement information obtained by cryo-EM (see **Chapter 2** and section **1.7.2 Development of a New Assay Platform to Make Isolated Measurements of Termination Rate**).

1.6 OUTLOOK

The sequence determinants and the basic steps of intrinsic termination are now relatively well understood. Answering many of the important remaining mechanistic questions will likely require the application of advanced technologies. Of particular interest are what the roles of RNAP domains and their conformational states are during each step of the termination pathway, whether interactions between Sw3 and the –10 RNA base and the lid and the –9 hybrid bp (10) are responsible for the energetic barrier to T_{hp} extension that separates T_{hp} folding into two steps (60, 65), and whether the second step in which extension of the T_{hp} causes melting of the upstream RNA–DNA hybrid (22, 23) is accompanied by essential RNAP rearrangements that commit the EC to the termination pathway (22, 65, 71, 73). The effects of pairing between A-tracts upstream of T_{hp} and the U-tract on TE *in vivo* (61) that are in apparent disagreement with the formation of termination-committed intermediates upon T_{hp} extension will also need to be reconciled. The idea of dissociable transcription factors that allow commitment to become reversible *in vivo* is an attractive possibility that needs testing.

Understanding the apparent differences in termination activities and mechanisms observed *in vitro* and *in vivo* will require biochemical experiments that better mimic the cellular environment, including the addition of appropriate transcription factors and better replication of cellular solute conditions. Most *in vitro* studies have used significant concentrations of chloride ion (Cl^-), yet Cl^- can compete for binding of DNA and RNA to proteins (77), and is not representative of the more complex cellular environment (78). The use of acetate or glutamate in place of Cl^- may create a better mimic of the bacterial cytoplasm (66, 70, 77, 79), and these anions should be used in future studies of termination mechanisms, when possible. The study of transcription mechanisms *in vivo* will also benefit from further extensions of high-throughput sequencing-enabled *in vivo* methods, including ChIP-sequencing and native elongating transcript (NET)-sequencing. Finally, addressing structural questions about termination

mechanisms is likely to depend heavily on the exploitation of recent advances in single-particle and time-resolved cryo-EM. Rapid freezing of actively terminating ECs can be used to trap intermediates in the pathway and to determine the contributions of RNAP domains, NA scaffold rearrangements, and dissociable transcription factors. Obtaining information about important pathway intermediates using these new methodologies has the potential to clarify the identity and thermodynamic stability of meta-stable intermediates, and, thus, define the energetics of each step in the termination pathway (80). Thus, these techniques together with complementary biochemical assays such as that presented in this work may be central to providing the next layer of understanding of the intrinsic termination mechanism.

1.7 Key Questions Addressed By This Work

The main goal of this work was to address the largely uncharted area of the effects of RNAP conformational changes on intrinsic termination. As a significant component of the energy barrier to termination is posed by stabilizing RNAP–NA interactions (25) that are likely altered by RNAP conformational changes, we hypothesized that the dynamic movements of RNAP may modify the energy barrier to termination. The active site-proximal trigger loop (TL) module of RNAP **(Figure 1.2B)** thus presented an attractive candidate as a highly dynamic module of RNAP that could potentially alter the energy barrier to termination. Moreover, the TL was also suggested to be involved in the intrinsic termination previously (73).

The role of dissociable transcription factors on the mechanism of intrinsic termination also remains incompletely characterized. We were interested in characterizing the role of NusA specific to intrinsic termination, as it is an essential factor in *Escherichia coli* (97), thought to be

associated with all active ECs (98). Described below are questions of interest pertaining to these topics that I have addressed in this work.

1.7.1 The Trigger Loop (TL) and Questions of Interest

The TL module is evolutionarily conserved in the multi-subunit RNA polymerases, including eukaryotic RNAPs I, II and II, as well as bacterial and archaeal RNAPs (81). The TL must adopt several distinct conformations to catalyze or enable different enzymatic activities in the active site of RNAP, and has been captured in several distinct conformations by X-ray crystallography **(Figure 1.6A)** (10, 12, 82). Folding of the TL into an α-helical hairpin (trigger helices, TH), stacked against the α-helical BH, positions cognate NTPs in the active site, thus catalyzing nucleotide addition up to 10^4-fold (27). However, TL unfolding is necessary for translocation of the DNA template and RNA, and for release of the pyrophosphate (PP_i) byproduct, in order to free the active site for the next round of nucleotide addition (83-87). The TL must also unfold when ECs reverse translocate (*i.e.* backtrack) on the template in response to misincorporated nucleotides or unstable RNA–DNA hybrids, and for factor-mediated cleavage of the RNA that is extruded through the secondary channel as a consequence of backtracking. Stabilization of the hairpin-stabilized paused EC state, however, requires a specific partially folded TL conformation (or set of conformations), swinging the catalytic residues of the TL away from the active site and temporarily disabling nucleotide addition (20, 84, 88). It remains an open question whether the TL is involved in a more long-lived inactivation of the active site during intrinsic termination.

The TL has been implicated in the modulation of TE previously. Two different point mutations in the TL that alter elongation rate and pausing were shown to decrease TE (73). However, as termination occurs in kinetic competition with elongation, TE can be influenced by factors that affect the mechanism of termination or by changes in the elongation rate or pausing; effects on elongation and pausing can alter the flux of complexes into and out of the termination

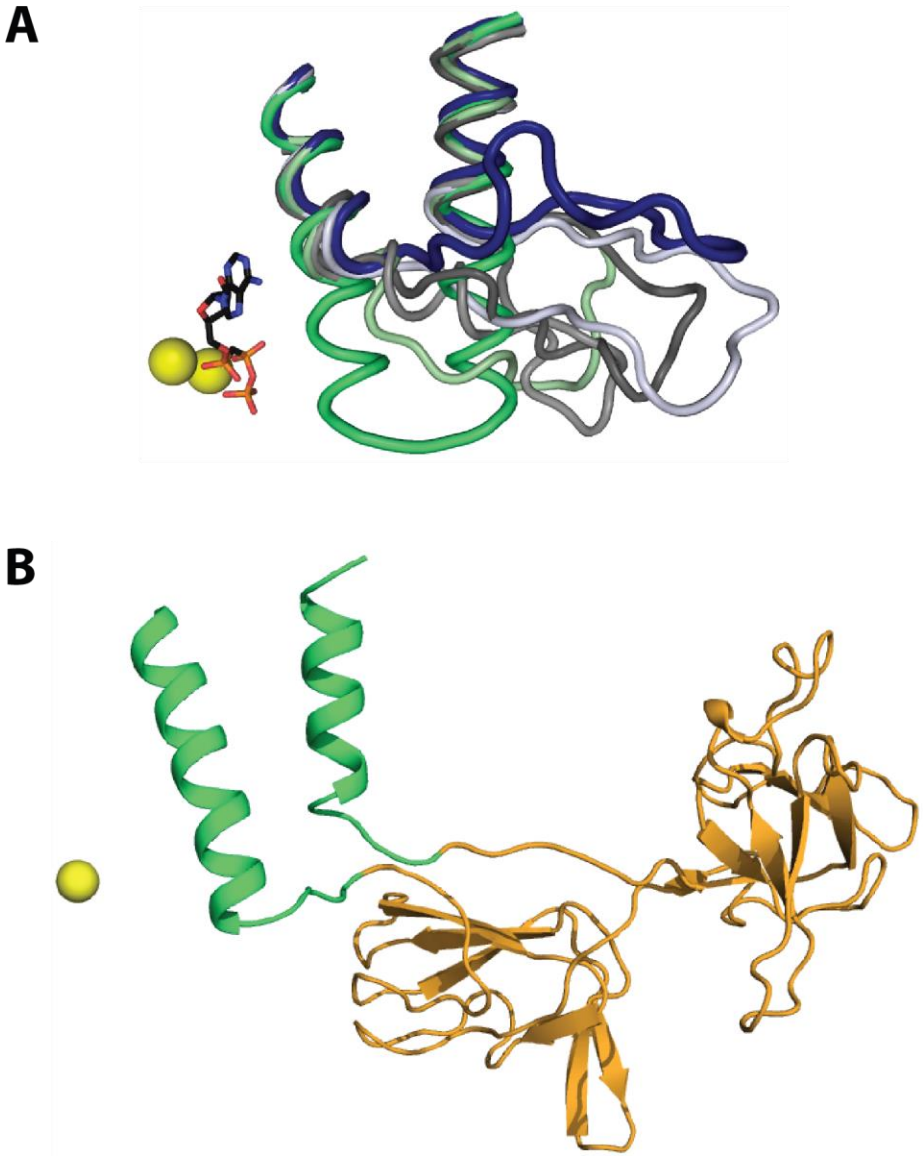
pathway by affecting the kinetic window during which termination can occur **(Figure 1.3)**. As TL mutations can have profound effects on elongation (88-90), it was unclear whether the observed decrease in TE was due to effects on termination.

To address the potential conflation of termination and elongation effects, dissociation rates from ECs stalled at a termination site were measured (73), revealing that a point mutation in the active site-distal arm of the TL decreased dissociation by a factor of ~3. It was further suggested in the same study that the T_{hp} must invade the main channel of RNAP and contact a folded TL for termination to occur (the hairpin invasion model, henceforth referred to as the TH–T_{hp} contact model), although the low salt conditions under which these experiments were conducted may have affected the results (72).

Thus, the role of TL in termination remains unclear, and there are a number of important questions that need to be addressed: Does the TL enhance termination independently of its effects on elongation and pausing? Which steps along the multi-step termination pathway are aided by TL activity? Is a folded TL necessary for TL-mediated termination enhancement? Does the T_{hp} occupy the main cleft of RNAP in a terminating EC? I answer many of these questions in **Chapter 2**.

Escherichia coli RNAP, used as a model system in this work as well as in the previous studies of TL termination effects (73, 91), contains a bulky 188-aa sequence insertion that is inserted in the (~44-aa) TL and called sequence insertion 3 (SI3; **Figure 1.6B**). SI3 is present in certain Gram-negative bacterial lineages (92, 93), and forms a mobile element extending in front of RNAP as two sandwich-barrel hybrid motif (SBHM) domains thought to interact with a third SBHM that is formed by the jaw domain of RNAP when the TL is unfolded (94), and is structurally related to similar SBHM repeat insertions present in cyanobacteria (92, 93). SI3 is essential for the correct assembly of the SI3-containing β′ subunit in the RNAP core enzyme (95), and has also been shown to affect other activities of the TL, including hairpin-stabilized

pausing and intrinsic cleavage (88, 95, 96). We were interested in determining whether the role of the TL in termination is modulated by SI3. A monoclonal antibody (mAb) thought to restrict TL motion through contact to SI3 was shown to decrease the dissociation of ECs stalled at a termination site by a factor of ~3, indicating that this domain may impact EC stability. In **Chapter 2** I determine the role of SI3 in each step of the intrinsic termination pathway.



<u>Figure 1.6:</u> The polymorphous trigger loop (TL) and sequence insertion 3 (SI3).

(A) A sampling of the many conformations of the active site-proximal polymorphous TL is shown, captured by X-ray crystallography in RNAPs from various organisms. The TL has been crystallized in various other conformations between the folded and unfolded states shown above, indicating that the TL can stably adopt a wide spectrum of conformations. PDBs shown (from folded TL, *green* to unfolded TL, *blue*) are: 205j, 3gtg, 2ppb, 1y1v, 1zyr, and 1iw7. Substrate NTP, *black stick structure*; active site and NTP-associated magnesium ions, *yellow spheres*. **(B)** A model of SI3, present in the *E. coli* RNAP TL, inserted in the folded TL (*green*) is shown in *orange*. Active site magnesium ion, *yellow sphere*.

1.7.2 Development of a New Assay Platform to Make Isolated Measurements of Termination Rate

Given the complex kinetics involved in the multi-step termination pathway, we were interested in determining which steps are specifically impacted by the TL, and whether TL dynamics impact the termination mechanism. However, the study of TL mutants poses a few technical challenges, most significantly that TL mutations are known to alter and in some cases severely impair the elongation rate, which can affect TE without altering the termination rate and also makes transcript elongation to a distant termination site impossible.

In this work (**Chapter 2**) I have developed an assay using *Escherichia coli* RNAP that enables isolated measurement of the termination rate, as well as the elongation, pause entry and pause escape rates along the termination pathway that impact TE. By allowing the assembly of active ECs 2 nt away from a termination site, I was able to study the effects of severely elongation-impaired mutants, including a complete deletion of the TL module, which slows nucleotide addition 10^4-fold relative to WT RNAP (27). We note also that the effects of TL dynamics on termination rate as observed by this termination assay are unlikely to be captured by cryo-EM; our results address the role of the TL in diversifying transition state heights during the penultimate step of termination **(Figure 1.3, step ③)**, whereas cryo-EM detects meta-stable intermediates. This assay thus provides an easily adaptable platform to investigate the roles of other RNAP structural modules such as the BH and clamp **(Figure 1.2B)** that are likely involved in the termination mechanism. Finally as I discuss in **Appendix A**, we suggest that this assay is better equipped to comment on termination mechanism than the most commonly used termination metrics of termination efficiency and the dissociation rate of artificially stalled ECs.

1.7.3 Investigating the Role of NusA in the Intrinsic Termination Pathway

The universal bacterial and archaeal transcription factor NusA is essential in *E. coli* (97), and is thought to be associated with all active ECs upon release of the initiation factor sigma (98). As described above, NusA has been shown to dock near the RNA exit channel through contacts to the tip of the flap module **(Figure 1.1)** and aid hairpin-stabilized pausing and termination of ECs from terminators with weak T_{hp}s (32, 35, 36, 38, 39, 53, 54). The role of NusA in aiding intrinsic termination and hairpin-stabilized pausing was thus proposed to be in stimulating hairpin formation in the exit channel of RNAP (38); NusA was consistently shown to increase the rate of intermolecular RNA duplex formation *in vitro* in a flap-tip dependent manner (15). Interestingly, whereas the N-terminal domain of NusA (NusA-NTD) was shown to be sufficient to recapitulate the effects of full-length NusA (FL-NusA) on hairpin-stabilized pausing and RNA duplex formation *in vitro*, even micromolar concentrations of NusA-NTD were unable to recapitulate the effect of FL-NusA on intrinsic termination (15, 38). The insufficiency of Nus-NTD for intrinsic termination raises key questions: Does NusA play a role in termination aside from nucleation of the T_{hp}? Is the effect of A-tracts on TE observed *in vivo* facilitated by NusA as proposed in section **1.4.4c Reconciling observed effects of A-tracts on termination efficiency**? Does NusA stimulate hairpin-stabilized pausing on the termination pathway? I begin to address these questions in **Chapter 3**.

1.8 REFERENCES

1. Peters JM, *et al.* (2012) Rho and NusG suppress pervasive antisense transcription in Escherichia coli. *Genes Dev* 26(23):2621-2633.

2. Peters JM, *et al.* (2009) Rho directs widespread termination of intragenic and stable RNA transcription. *Proc Natl Acad Sci U S A* 106(36):15406-15411.

3. Cardinale CJ, *et al.* (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science* 320(5878):935-938.

4. Leela JK, Syeda AH, Anupama K, & Gowrishankar J (2013) Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli. *Proc Natl Acad Sci U S A* 110(1):258-263.

5. Selby CP & Sancar A (1993) Molecular mechanism of transcription-repair coupling. *Science* 260(5104):53-58.

6. Park JS & Roberts JW (2006) Role of DNA bubble rewinding in enzymatic transcription termination. *Proc Natl Acad Sci U S A* 103(13):4870-4875.

7. Selby CP & Sancar A (1995) Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. *J Biol Chem* 270(9):4882-4889.

8. Komissarova N & Kashlev M (1998) Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proc Natl Acad Sci U S A* 95(25):14699-14704.

9. Korzheva N, *et al.* (2000) A structural model of transcription elongation. *Science* 289(5479):619-625.

10. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, & Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448(7150):157-162.

11. Darst SA, *et al.* (2002) Conformational flexibility of bacterial RNA polymerase. *Proc Natl Acad Sci U S A* 99(7):4296-4301.

12. Vassylyev DG, *et al.* (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* 417(6890):712-719.

13. Chakraborty A, *et al.* (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science* 337(6094):591-595.

14. Tagami S, *et al.* (2010) Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* 468(7326):978-982.

15. Hein PP, *et al.* (2014) RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nat Struct Mol Biol* 21(9):794-802.

16. Sekine S, Murayama Y, Svetlov V, Nudler E, & Yokoyama S (2015) Ratcheting of RNA polymerase toward structural principles of RNA polymerase operations. *Transcription* 6(3):56-60.

17. Gnatt AL, Cramer P, Fu J, Bushnell DA, & Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* 292(5523):1876-1882.

18. Ederth J, Artsimovitch I, Isaksson LA, & Landick R (2002) The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J Biol Chem* 277(40):37456-37463.

19. Cramer P*, et al.* (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288(5466):640-649.

20. Weixlbaumer A, Leon K, Landick R, & Darst SA (2013) Structural basis of transcriptional pausing in bacteria. *Cell* 152(3):431-441.

21. Wilson KS & von Hippel PH (1994) Stability of Escherichia coli transcription complexes near an intrinsic terminator. *J Mol Biol* 244(1):36-51.

22. Gusarov I & Nudler E (1999) The mechanism of intrinsic transcription termination. *Mol Cell* 3(4):495-504.

23. Komissarova N, Becker J, Solter S, Kireeva M, & Kashlev M (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Mol Cell* 10(5):1151-1162.

24. Yager TD & von Hippel PH (1991) A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. *Biochemistry* 30(4):1097-1118.

25. von Hippel PH & Yager TD (1992) The elongation-termination decision in transcription. *Science* 255(5046):809-812.

26. Lee DN, Phung L, Stewart J, & Landick R (1990) Transcription pausing by Escherichia coli RNA polymerase is modulated by downstream DNA sequences. *J Biol Chem* 265(25):15145-15153.

27. Toulokhonov I, Zhang J, Palangat M, & Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* 27(3):406-419.

28. Yarnell WS & Roberts JW (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* 284(5414):611-615.

29. Neuman KC, Abbondanzieri EA, Landick R, Gelles J, & Block SM (2003) Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* 115(4):437-447.

30. Larson MH*, et al.* (2014) A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* 344(6187):1042-1047.

31. Vvedenskaya IO*, et al.* (2014) Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science* 344(6189):1285-1289.

32. Toulokhonov I & 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(5):1125-1136.

33. Wang D, Severinov K, & 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(16):8433-8438.

34. Chan CL & 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(1):25-42.

35. Wilson KS & von Hippel PH (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A* 92(19):8793-8797.

36. Toulokhonov I, Artsimovitch I, & Landick R (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* 292(5517):730-733.

37. Yang X*, et al.* (2009) The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Rep* 10(9):997-1002.

38. Ha KS, Toulokhonov I, Vassylyev DG, & 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(5):708-725.

39. Kolb KE, Hein PP, & Landick R (2014) Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. *J Biol Chem* 289(2):1151-1163.

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

41. Guajardo R & Sousa R (1997) A model for the mechanism of polymerase translocation. *J Mol Biol* 265(1):8-19.

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

43. Martin FH & Tinoco I, Jr. (1980) DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res* 8(10):2295-2299.

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

45. Herbert KM*, et al.* (2010) E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *J Mol Biol* 399(1):17-30.

46. Yakhnin AV & 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(3):690-705.

47. Yakhnin AV, Yakhnin H, & 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(42):16131-16136.

48. Czyz A, Mooney RA, Iaconi A, & Landick R (2014) Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. *MBio* 5(2):e00931.

49. Yakhnin AV, Murakami KS, & Babitzke P (2016) NusG Is a Sequence-specific RNA Polymerase Pause Factor That Binds to the Non-template DNA within the Paused Transcription Bubble. *J Biol Chem* 291(10):5299-5308.

50. Sevostyanova A, Belogurov GA, Mooney RA, Landick R, & Artsimovitch I (2011) The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol Cell* 43(2):253-262.

51. Svetlov V, Belogurov GA, Shabrova E, Vassylyev DG, & Artsimovitch I (2007) Allosteric control of the RNA polymerase by the elongation factor RfaH. *Nucleic Acids Res* 35(17):5694-5705.

52. Belogurov GA*, et al.* (2007) Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol Cell* 26(1):117-129.

53. Schmidt MC & Chamberlin MJ (1987) nusA protein of Escherichia coli is an efficient transcription termination factor for certain terminator sites. *J Mol Biol* 195(4):809-818.

54. Mondal S, Yakhnin AV, Sebastian A, Albert I, & Babitzke P (2016) NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat Microbiol* 1:15007.

55. Santangelo TJ & Artsimovitch I (2011) Termination and antitermination: RNA polymerase runs a stop sign. *Nat Rev Microbiol* 9(5):319-329.

56. d'Aubenton Carafa Y, Brody E, & Thermes C (1990) Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* 216(4):835-858.

57. McDowell JC, Roberts JW, Jin DJ, & Gross C (1994) Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. *Science* 266(5186):822-825.

58. Kireeva ML & Kashlev M (2009) Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc Natl Acad Sci U S A* 106(22):8900-8905.

59. Chan CL, Wang D, & Landick R (1997) Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *J Mol Biol* 268(1):54-68.

60. Lubkowska L, Maharjan AS, & Komissarova N (2011) RNA folding in transcription elongation complex: implication for transcription termination. *J Biol Chem* 286(36):31576-31585.

61. Chen YJ*, et al.* (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat Methods* 10(7):659-664.

62. Penno C*, et al.* (2015) Productive mRNA stem loop-mediated transcriptional slippage: Crucial features in common with intrinsic terminators. *Proc Natl Acad Sci U S A* 112(16):E1984-1993.

63. Goliger JA, Yang XJ, Guo HC, & Roberts JW (1989) Early transcribed sequences affect termination efficiency of Escherichia coli RNA polymerase. *J Mol Biol* 205(2):331-341.

64. Telesnitsky AP & Chamberlin MJ (1989) Sequences linked to prokaryotic promoters can affect the efficiency of downstream termination sites. *J Mol Biol* 205(2):315-330.

65. Larson MH, Greenleaf WJ, Landick R, & Block SM (2008) Applied force reveals mechanistic and energetic details of transcription termination. *Cell* 132(6):971-982.

66. Reynolds R, Bermudez-Cruz RM, & Chamberlin MJ (1992) Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators. *J Mol Biol* 224(1):31-51.

67. Toulokhonov I & Landick R (2006) The role of the lid element in transcription by E. coli RNA polymerase. *J Mol Biol* 361(4):644-658.

68. Santangelo TJ & Roberts JW (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol Cell* 14(1):117-126.

69. Peters JM, Vangeloff AD, & Landick R (2011) Bacterial transcription terminators: the RNA 3'-end chronicles. *J Mol Biol* 412(5):793-813.

70. Reynolds R & Chamberlin MJ (1992) Parameters affecting transcription termination by Escherichia coli RNA. II. Construction and analysis of hybrid terminators. *J Mol Biol* 224(1):53-63.

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

72. Kashlev M & Komissarova N (2002) Transcription termination: primary intermediates and secondary adducts. *J Biol Chem* 277(17):14501-14508.

73. Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S, & Nudler E (2007) An allosteric path to transcription termination. *Mol Cell* 28(6):991-1001.

74. Weilbaecher R, Hebron C, Feng G, & Landick R (1994) Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev* 8(23):2913-2927.

75. Cheeran A*, et al.* (2005) Escherichia coli RNA polymerase mutations located near the upstream edge of an RNA:DNA hybrid and the beginning of the RNA-exit channel are defective for transcription antitermination by the N protein from lambdoid phage H-19B. *J Mol Biol* 352(1):28-43.

76. Binshtein E & Ohi MD (2015) Cryo-electron microscopy and the amazing race to atomic resolution. *Biochemistry* 54(20):3133-3141.

77. Leirmo S, Harrison C, Cayley DS, Burgess RR, & Record MT, Jr. (1987) Replacement of potassium chloride by potassium glutamate dramatically enhances protein-DNA interactions in vitro. *Biochemistry* 26(8):2095-2101.

78. Cayley S, Lewis BA, Guttman HJ, & Record MT, Jr. (1991) Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo. *J Mol Biol* 222(2):281-300.

79. Chan CL & Landick R (1997) Effects of neutral salts on RNA chain elongation and pausing by Escherichia coli RNA polymerase. *J Mol Biol* 268(1):37-53.

80. Fischer N*, et al.* (2015) Structure of the E. coli ribosome-EF-Tu complex at <3 A resolution by Cs-corrected cryo-EM. *Nature* 520(7548):567-570.

81. Jokerst RS, Weeks JR, Zehring WA, & Greenleaf AL (1989) Analysis of the gene encoding the largest subunit of RNA polymerase II in Drosophila. *Mol Gen Genet* 215(2):266-275.

82. Wang D, Bushnell DA, Westover KD, Kaplan CD, & Kornberg RD (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127(5):941-954.

83. Vassylyev DG*, et al.* (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448(7150):163-168.

84. Nayak D, Voss M, Windgassen T, Mooney RA, & Landick R (2013) Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol Cell* 50(6):882-893.

85. Silva DA*, et al.* (2014) Millisecond dynamics of RNA polymerase II translocation at atomic resolution. *Proc Natl Acad Sci U S A* 111(21):7665-7670.

86. Malinen AM*, et al.* (2012) Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Res* 40(15):7442-7451.

87. Kireeva ML*, et al.* (2012) Molecular dynamics and mutational analysis of the catalytic and translocation cycle of RNA polymerase. *BMC Biophys* 5:11.

88. Windgassen TA*, et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

89. Zhang J, Palangat M, & Landick R (2010) Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat Struct Mol Biol* 17(1):99-104.

90. Mejia YX, Nudler E, & Bustamante C (2015) Trigger loop folding determines transcription rate of Escherichia coli's RNA polymerase. *Proc Natl Acad Sci U S A* 112(3):743-748.

91. Epshtein V, Dutta D, Wade J, & Nudler E (2010) An allosteric mechanism of Rho-dependent transcription termination. *Nature* 463(7278):245-249.

92. Iyer LM, Koonin EV, & Aravind L (2004) Evolution of bacterial RNA polymerase: implications for large-scale bacterial phylogeny, domain accretion, and horizontal gene transfer. *Gene* 335:73-88.

93. Lane WJ & Darst SA (2010) Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J Mol Biol* 395(4):671-685.

94. Opalka N*, et al.* (2010) Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. *PLoS Biol* 8(9).

95. Zakharova N, Bass I, Arsenieva E, Nikiforov V, & Severinov K (1998) Mutations in and monoclonal antibody binding to evolutionary hypervariable region of Escherichia coli RNA polymerase beta' subunit inhibit transcript cleavage and transcript elongation. *J Biol Chem* 273(38):24912-24920.

96. Artsimovitch I, Svetlov V, Murakami KS, & 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(14):12344-12355.

97. Nakamura Y & Uchida H (1983) Isolation of conditionally lethal amber mutations affecting synthesis of the nusA protein of Escherichia coli. *Mol Gen Genet* 190(2):196-203.

98. Mooney RA*, et al.* (2009) Regulator trafficking on bacterial transcription units in vivo. *Mol Cell* 33(1):97-108.

Chapter 2:

Trigger loop contributions to conformational dynamics of *Escherichia coli* RNAP aid intrinsic termination

Adapted from:

Ananya Ray-Soni, Rachel A. Mooney, and Robert Landick. (2017) Trigger loop dynamics can explain intrinsic termination by bacterial RNA polymerase without terminator hairpin contact.

Manuscript in preparation for submission to:
Proceedings of the National Academy of Sciences.

Chapter contributions:

A.R.S. performed all experiments, experimental design, analysis and kinetic modeling, and prepared the manuscript; R.A.M. constructed plasmids from which RNAP variants were purified, and copyedited the text; R.L. oversaw the experimental design and manuscript preparation.

2.1 ABSTRACT

Intrinsic transcription termination is a process by which a highly stable elongating transcription complex (EC) is dissociated over a short window of 2–3 nt after synthesizing up to kilobases of gene product. Termination is signaled by the formation of a stem-loop structure in the nascent RNA transcript and weak base pairing between the RNA and DNA in the EC, leading to a series of events that ultimately results in the dissociation of the EC into its components: RNA polymerase (RNAP), DNA, and RNA. Although the contributions of the sequence elements of the terminator have been extensively characterized, the role of conformational changes of RNAP in the termination mechanism remains almost completely undescribed. The active site-proximal polymorphous trigger loop (TL) – responsible for the catalysis of nucleotide addition and the mediation of responses to certain regulatory pause signals – was an interesting candidate for a module whose structural rearrangements could impact termination, being one of the most dynamic modules of RNAP. Existing assays to study termination of actively elongating ECs are impacted by effects on elongation, which can be significant for TL variants. To address the role of the TL in intrinsic termination independent of elongation effects, we developed an assay to separately measure the kinetics of elongation and termination for each step in the complex termination pathway. Using this assay, we determined that the TL significantly stimulates the termination mechanism but is not essential, and that the TL itself – not its large sequence insertion SI3 – is responsible for this effect. However, both the TL and SI3 act to increase the flux of ECs into the termination pathway by increasing pausing at the termination sites, and as such present a regulatory access point for intrinsic termination. It was suggested in a previous study (1) that the TL must adopt the folded α-helical hairpin conformation during intrinsic termination to contact the terminator hairpin that invades the main cleft of RNAP. However, our results indicate that the folded and unfolded TL conformations are both unfavorable for termination. We propose that a dynamic TL contributes significantly to EC conformational

diversity, and thereby affords access to pathways with lower energy barriers to termination. As such, it presents a unifying model to explain conflicting proposals for the termination-favorable TL conformation in factor-dependent versus -independent termination (1, 2).

2.2 INTRODUCTION

Transcription termination is an essential process in all organisms by which the RNA transcript and the DNA template are released from RNA polymerase (RNAP). Termination prevents unwanted gene expression, recycles RNAP for new initiation events, and prevents genome-destabilizing collisions with the replication machinery. In bacteria, intrinsic termination of the elongating transcription complex (EC) can occur without accessory factors in response to a signal expressed in the newly transcribed RNA. Intrinsic termination occurs over a 2–3 nt window of EC destabilization (3) caused by the formation of a GC-rich hairpin in the RNA exit channel of RNAP, immediately upstream from a 3′-terminal, ~8 nt U-rich RNA tract **(Figure 2.1A)**. Precise termination of the highly stable EC within this small sequence window requires rapid entry into the termination pathway before continued nucleotide addition moves the EC beyond the region of destabilization (3).

A well-supported thermodynamic termination model posits that termination can occur at any sequence position in kinetic competition with elongation, where the relative energy barriers to the two processes determine the proportion of ECs that terminate versus elongate (3-5) **(Figure 2.1B)**. The EC is stabilized by a combination of polar, van der Waals, and electrostatic contacts between RNAP and *(i)* the 9–10 bp RNA–DNA hybrid formed within an ~12 bp transcription bubble; *(ii)* the ~18 bp of DNA in the downstream DNA channel; and *(iii)* the ~5 nt of unpaired nascent RNA in the RNA exit channel (6-9). The RNA–DNA hybrid itself constitutes a significant portion of EC stability (6, 10-12) **(Figure 2.1B)**. These stabilizing

interactions must be broken for dissociation of the EC in the final step of termination **(Figure 2.1C)**. At non-terminator positions, the barrier to termination is prohibitively high as a consequence of the EC stability maintained by both the RNA–DNA hybrid and RNAP–nucleic acid (NA) interactions. At termination sites, however, the hybrid is weakened by the prevalence of weak rU–dA bps in the U-tract (13). The GC-rich terminator hairpin (T_{hp}) further "locks" RNAP over the weak hybrid, preventing reverse translocation (*i.e.,* backtracking) of the EC to a more stable hybrid position. This lowers the energy barrier to termination, allowing competitive termination rates and appreciable termination efficiency (TE).

The terminator sequence elements are crucial for favoring termination in the multi-step termination pathway. In the first step of the termination pathway, the U-tract sequence causes a fraction of ECs to enter an elemental paused state (14, 15) **(Figure 2.1C)**, increasing the barrier to elongation and creating a kinetic window that allows the T_{hp} to nucleate. The partially formed T_{hp} may further slow pause escape via hairpin-stabilized pausing, although the duration of hairpin-stabilized pausing during termination is uncertain. Next, the hairpin is extended by 2–3 bp (16), causing the upstream 3 rU–dA bp in the U-tract to melt (5, 6, 14). Hairpin extension is proposed to melt the upstream third of the hybrid by alternative routes: *(i)* in hybrid-shearing, the extending hairpin pulls the RNA out of register from the DNA by shearing the U-rich hybrid (6, 17); *(ii)* in hypertranslocation, hairpin extension pushes RNAP forward on the DNA without concomitant nucleotide addition when non-rU–dA bp in the hybrid inhibit shearing (17-19). Partial melting of the already weak U-rich hybrid is postulated to destabilize the EC sufficiently for the final step of termination – EC dissociation (6). It has also been proposed that a step preceding dissociation may irreversibly inactivate nucleotide addition, functionally committing the complexes to termination (17, 20).

Although the role of NA in termination has been extensively characterized (reviewed in (21)), the RNAP–NA contacts that contribute to termination remain incompletely described.

Contacts between the RNAP clamp module **(Figure 2.1D)** and downstream DNA have been shown to affect TE (1); the flap (22) and zinc finger (1, 9) modules also affect TE, likely through contacts to the T_{hp}. However, RNAP–NA interactions constitute a significant component of the energy barrier to termination (4) **(Figure 2.1B)** as evidenced by an up to 10-fold effect of chloride ion concentration on TE at some terminators (23), and thus merit further investigation.

RNAP–NA interactions are likely modulated during termination by key rearrangements of RNAP structural modules, the contributions of which remain almost wholly uncharacterized. It has been proposed that the clamp module of RNAP, which closes around the NA in a stable EC to form numerous stabilizing contacts, may need to open for EC dissociation (1, 24-27). A point mutation in the RNAP trigger loop (TL; **Figure 2.1D**) – the module that catalyzes nucleotide addition by folding into the α-helical trigger helices (TH) – has been shown to affect dissociation of ECs stalled at a termination site (1), suggesting a role for the TL in EC stability. The TL has also been proposed to aid termination by folding into the TH conformation and contacting the T_{hp}, enabling the T_{hp} to invade the main RNAP cleft (1). The highly dynamic TL adopts various conformations to enable the multiple activities of the RNAP active site **(Figure 2.1E)** and has been captured in several different conformations by X-ray crystallography (29-31); as such, the TL poses an attractive candidate as a mobile module of RNAP that may impact termination. However, direct contacts between the TL and upstream RNA, required by the TH–T_{hp} contact model, have not been documented in any crystal structure reported to date, and the biochemical evidence for this contact in a terminating EC has been disputed (32). A testable prediction raised by the TH–T_{hp} contact model is that the folded TL stimulates termination by contacting the T_{hp} in the main cleft. Moreover, given that termination occurs by a kinetically complex multi-step termination pathway, it remained to be determined whether the TL plays a role in termination in steps preceding EC dissociation.

To understand the specific role of the TL on individual steps of termination, we developed an experimental approach that can separate effects on elongation rate and pausing – known to be significant for TL alterations (33-35) – from effects on termination rate. Whereas traditional TE measurements report an aggregate of effects on termination and elongation, isolated measurement of termination rate in our assay enabled the comparison of the effects of TL mutations specifically on the termination mechanism. Using RNAP variants that *(i)* bias the TL towards either the folded or unfolded states, *(ii)* remove the sequence insertion SI3 from the TL, or *(iii)* delete the TL completely, we were able to test the contribution of the TL to termination rate, termination efficiency, and the pausing steps leading into the termination pathway.

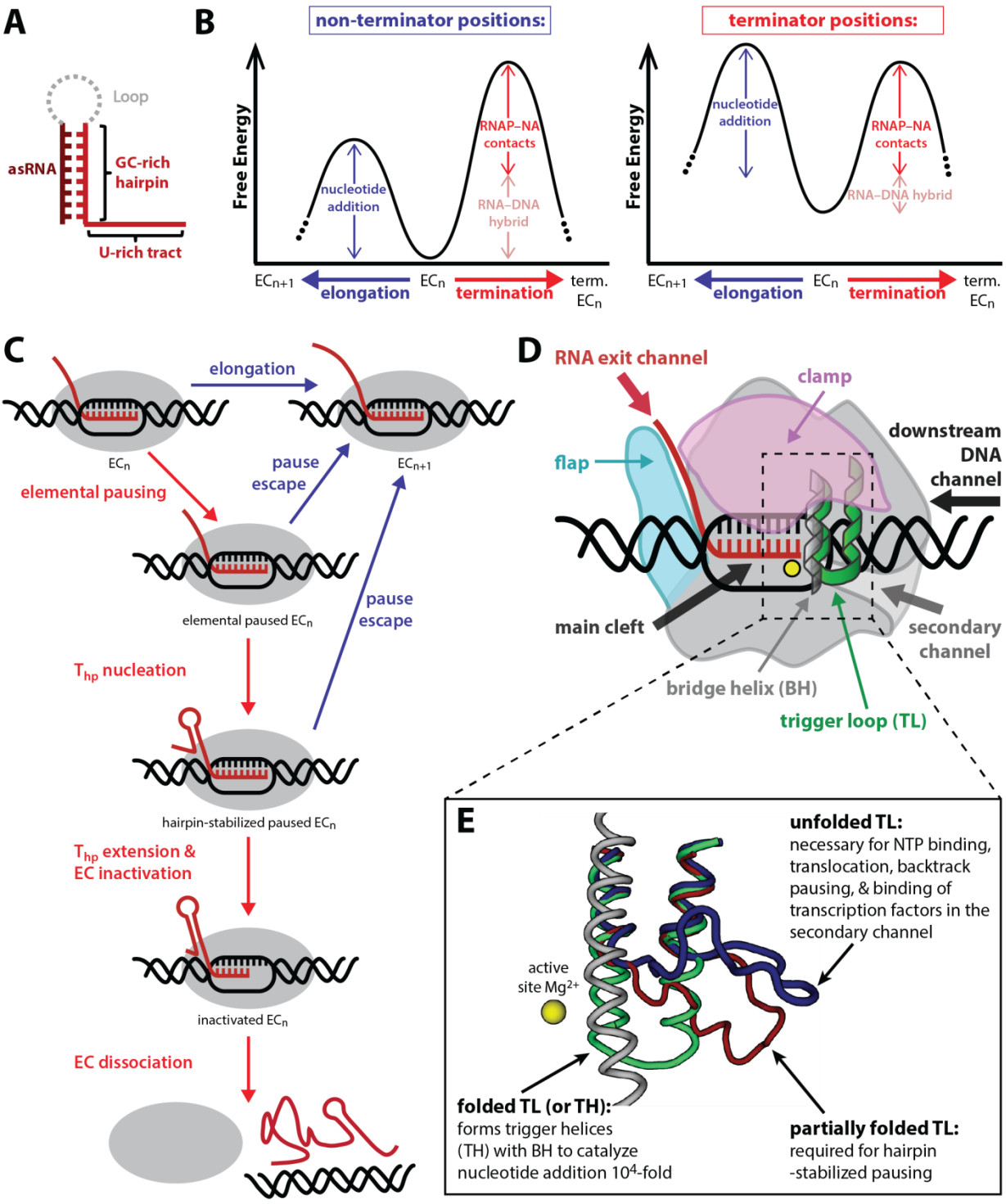


Figure 2.1: Intrinsic transcription termination and the polymorphous trigger loop (TL).

(A) The 3 main elements of a canonical intrinsic terminator are shown. The endogenous terminator structure formed in the nascent RNA contains a U-rich tract, immediately preceded by a GC-rich stem-loop (or hairpin) structure. In our system the upstream arm of the terminator is mimicked by an antisense

RNA (asRNA, *dark red*) complementary to the downstream terminator arm (*lighter red*), creating an artificial terminator hairpin (T_{hp}); the loop (*gray*) is absent in the artificial T_{hp}. **(B)** The thermodynamic model for termination (3-5), in which the relative energy barriers to elongation (*blue*) *versus* termination (*red*) determine the probability of either event. At non-terminator positions the barrier to elongation is significantly lower than the barrier to termination, where the termination barrier is created by 2 components: *(i)* the stability of the RNA–DNA hybrid (*pink*), and *(ii)* interactions between RNAP and the nucleic acids (NA; *red*). At terminator positions, the weak base pairs of the U-tract destabilize the RNA–DNA hybrid and therefore the EC, effectively decreasing the RNA–DNA hybrid component of the energy barrier to termination. The reduced energy barrier allows appreciable termination efficiencies. **(C)** The currently defined steps of intrinsic termination as identified by the contributions of the nucleic acid rearrangements. Steps towards the termination pathway are indicated by red arrows; blue arrows indicate steps that result in terminator read-through. RNA, *red*; DNA, *black*; RNAP, *gray*. **(D)** Cartoon depiction of an EC and relevant elements. The active site is marked by a magnesium ion (*yellow*). RNA, *red*; DNA, *black*; RNAP, *gray, purple & blue*. **(E)** Three conformations of the polymorphous TL captured in crystal structures and some of their roles are shown. The TL has been crystallized in various other conformations between the folded and unfolded states shown above, indicating that the TL can stably adopt a wide spectrum of conformations. Folded TL (or, TH), *green* (PDB 205j); partially folded TL, *red* (PDB 2nvq); unfolded TL, *blue* (PDB 1iw7), bridge-helix, *gray* (PDB 205j); magnesium, *yellow*.

2.3 RESULTS

2.3.1 Measuring termination from elongation-compromised RNAP mutants

To determine the effect of the TL on the intrinsic termination pathway, we needed an assay with two key features: (*i*) a way to measure the rate of termination itself at a given EC position and not just the TE that results from competition between the elongation and termination pathways; and (*ii*) a way that catalytically compromised ECs (*e.g.*, ΔTL ECs) could transcribe into the window of termination on the template from a nearby upstream position so that termination of TL mutants could be studied. We developed an assay with these two key features by modifying the hairpin-stabilized *his* pause sequence to resemble an intrinsic terminator (t_{his2}; **Figure 2.2A**). The hybrid was weakened by the addition of rU-dA bps to create a terminator U-tract stable enough to allow efficient EC assembly 2 nt upstream of the termination sites, and the upstream half of the T_{hp} was eliminated to allow the measurement of elongation and pause kinetics in the absence of a complete termination signal. An antisense oligonucleotide (oligo) that pairs to the exiting RNA 8–10 nt from the RNA 3′ end could then be added to create an artificial T_{hp} that enabled measurement of the termination rate (15). The sequence design also allowed us to avoid three distinct problems observed on other sequences we tested: (*i*) spontaneous EC dissociation due to an unstable RNA–DNA hybrid; (*ii*) backtracking of ECs and intrinsic cleavage of the nascent RNA prior to addition of the antisense oligo; and (*iii*) inefficient or slow termination (see **Appendix A**, section **A.2.6 Challenges with previously tested scaffolds)**. The sequences and buffer conditions were also selected to slow termination sufficiently to enable termination and elongation rate measurements.

Another strategy used previously to study termination without the potentially confounding effects of termination has been to monitor dissociation from ECs stalled at a termination site (1, 6). However we found that these experiments produce dissociation rates that

are not likely to be highly representative of physiological termination rates, and are relatively insensitive to RNAP mutation (see **Appendix A**).

Thus, using our assay design, we first tested whether both wild-type (WT; fast elongating), and ΔTL (slowly elongating) ECs could be efficiently assembled and elongated past the termination sites on the t_{his2} scaffold (C18 & U19; **Figure 2.2A**) in the absence of the T_{hp}. WT and ΔTL ECs were assembled at position G17. The RNAs were radiolabeled by addition of [α^{32}-P]-CTP and then elongated through U19 by incubation with UTP & ATP (100 μM UTP and 10 μM ATP for WT; 10 mM UTP and ATP for ΔTL) in the absence of an antisense oligo **(Figure 2.2B)**. Under these conditions, we observed nearly complete elongation past the termination site upon NTP addition for both WT and ΔTL ECs in the absence of an antisense oligo.

To test whether antisense oligos complementary to the exiting RNA could stimulate termination and dissociation of ECs on this scaffold, WT G17 ECs were reconstituted with His-tagged RNAPs and tethered to paramagnetic Co^{2+}-beads to allow separation of RNAP-bound (pellet) and released (supernatant; S) RNAs. The ECs were radiolabeled by addition of [α^{32}-P]-CTP and then elongated through U19 with 100 μM UTP and 10 μM ATP in the presence and absence of complementary (asDNA or asRNA) or non-complementary (nasDNA or nasDNA) oligos **(Figure 2.2C)**. Importantly, the complementary but not the non-complementary oligos stimulated release of RNA into the supernatant, indicating that termination occurred only in response to duplex formation in the RNAP exit channel. We found that the asRNAs stimulated more termination than asDNAs **(Figure 2.2C)**. Given that asRNAs also more closely replicate natural terminators, we used asRNAs for all subsequent experiments.

We used two different asRNAs for termination rate measurements, which we termed −8asRNA and −10asRNA because they paired to the −8 or −10 positions of the nascent RNA, respectively (where the 3′ RNA nt at the U19 termination site is −1; **Figure 2.2A**). The

−8asRNA mimics a canonical terminator that directly disrupts the −8 RNA–DNA bp during the T_{hp} extension step by formation of an RNA–RNA bp in the T_{hp} stem **(Figure 2.1C)**, and causes melting of the upstream 3 bp of the remaining RNA–DNA hybrid (6, 14). The −10asRNA forms a weaker terminator that does not directly compete for base pairing in the hybrid upon extension of the T_{hp} but still destabilizes the EC (3, 15). The two asRNA variants were needed to enable termination rate measurements from the various TL mutants (discussed further in sections below).

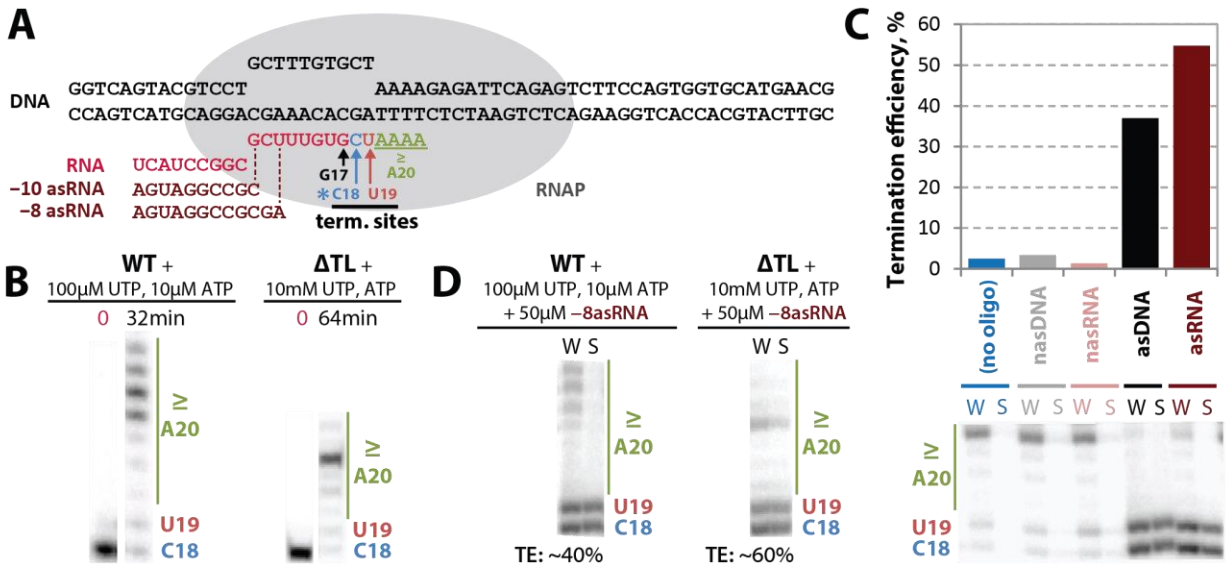


Figure 2.2: Measuring termination from elongation-compromised RNAP mutants.

(A) The t_{His2} terminator sequence and asRNAs used in this study are shown. The −8asRNA pairs at its 5′ end to the −8 RNA base (where the 3′ RNA nt at the U19 termination site is −1), and the −10asRNA pairs to the −10 RNA base, forming a weaker terminator by moving the base of the hairpin away from the active site. The asterisk denotes the position of a radiolabel in the nascent RNA under our assay setup. Termination is observed from 2 sites in the scaffold, C18 & U19. **(B)** Complete elongation past the termination sites was observed for WT ECs with 100 μM UTP & 10 μM ATP by 32 min, and for ΔTL ECs with 10 mM UTP/ATP by 64 min. **(C)** Robust termination and release of C18 & U19 RNAs occurs with oligonucleotides complementary to the exiting RNA, but not with non-complementary oligonucleotides. ECs were assembled at G17, tethered to paramagnetic Co^{2+}-beads using His-tagged RNAP, and the RNAs radiolabeled at C18 by the addition of [α^{32}P]-CTP. Elongation and termination were initiated by adding 100 μM UTP & 10 μM ATP, and 50 μM 12 nt RNA or DNA oligonucleotides complementary (as) or non-complementary (nas) to the RNA in the RNAP exit channel. The complementary asRNA is equivalent to the −8asRNA in **(A)**. Termination efficiency (TE) was calculated as the proportion of complexes remaining at C18 & U19 after 32 min. Whole (W) = pellet (RNAP-bound RNA) + supernatant (S; released RNA). **(D)** Both WT and ΔTL ECs terminate efficiently at C18 & U19. WT and ΔTL ECs were assembled and radiolabeled at C18 as in **(C)**. Elongation and termination were initiated for WT ECs as in **(C)**; ΔTL ECs were elongated from C18 with 10 mM UTP & ATP, and 50 μM −8asRNA was added after 280 s to allow some U19 accumulation. Termination efficiency (TE) is calculated as the percent of RNAs released in the supernatant. Whole (W) = pellet (RNAP-bound RNA) + supernatant (S; released RNA).

2.3.2 Deletion of the TL increases overall termination efficiency

Given the report that TL–T_{hp} interactions are required for efficient termination, we first tested whether deleting the TL altered TE. To answer this question, WT and ΔTL G17 ECs were assembled with His-tagged RNAP on the t_{his2} scaffold and bound to paramagnetic Co^{2+}-beads. RNAs were radiolabeled at C18, and then elongated through U19 with UTP and ATP in the absence or presence of –8asRNA **(Figure 2.2D)**. The –8asRNA was added to ΔTL ECs 280 s after beginning the elongation reaction to allow U19 ECs to accumulate before termination was stimulated. Samples of both the supernatant (S) and whole (W; W = S + pellet) fractions were taken after 32 min or 128 min for WT and ΔTL ECs, respectively. TE was determined as the percentage of C18 and U19 RNAs released from RNAP relative to the total RNA in the reaction. The combined TE (as well as individual TEs) at C18 and U19 was unexpectedly greater for ΔTL ECs (~60%) than for WT ECs (~40%; see also **Table 2.1A** for comparison of TEs with –10asRNA) even at the elevated NTP concentration used for ΔTL ECs, which would be predicted to increase pause escape rate and decrease TE. The simplest interpretation of these data is that the TL is dispensable for termination. This is in contrast to the prediction of the TL–T_{hp} contact model (1). However, the incomplete termination (i.e. less than 100% TE) of ΔTL ECs suggests that TL deletion may slow termination rate. Because TE is a product of competing termination and elongation rates and TL deletion impairs elongation rate 10^4-fold (8, 36), ΔTL ECs would be expected to terminate at 100% TE if there was no effect of TL deletion on termination rate. However, the less than 100% TE for ΔTL ECs observed here suggests that termination rate is likely also decreased by deletion of the TL. Such an effect of the TL on termination rate would further indicate a role of the TL in the termination mechanism, independent of its effects on elongation rate.

2.3.3 The TL strongly stimulates the rate of intrinsic termination

To determine how the TL affects the termination mechanism, we next measured the termination rates of WT and ΔTL ECs using our termination scaffold, t_{his2}, to isolate TL effects on termination versus elongation. Rapid kinetic measurements were required to follow the elongation and termination kinetics for WT ECs at single nucleotide resolution. We used a quench-flow apparatus rather than manual pipetting to measure the fast elongation and termination rates for WT RNAP. Elongation and pause kinetics over the termination sequence were measured in the absence of the T_{hp}, and then factored into the overall kinetics when termination was stimulated by addition of an antisense oligo to recapitulate the T_{hp} (15) **(Figure 2.1A** and **Figure 2.2A)**. Termination rate could then be determined by modeling competing rates for elongation and termination, where the rate at which complexes become resistant to elongation represents the rate at which they are functionally inactivated, *i.e.* terminated.

WT and ΔTL ECs were assembled at position G17 on the t_{his2} scaffold, radiolabeled at C18, and elongated through U19 with UTP and ATP **(Figure 2.3A)** in the absence **(Figure 2.3B-D)** or presence (**Figure 3E-F** and **Figure 2.S1A-B**) of –8asRNA. Addition of asRNA was delayed for ΔTL EC reactions to allow U19 accumulation, whereas for WT, asRNA was added with the NTPs. Although some WT ECs read through the terminator, ~87% of the ΔTL ECs positioned at C18 and U19 at the time of –8asRNA addition terminated **(Table 2.1A)**, as evidenced by the immediate plateau formed by the U19 population upon addition of –8asRNA to ΔTL ECs **(Figure 2.3F)**. With minimal escape of ΔTL ECs from the termination site, only a lower limit could be estimated for the termination rate with –8asRNA (>175 x 10^{-4} s^{-1} for ΔTL *vs.* [995 ± 300] x 10^{-4} s^{-1} for WT; **Table 2.1A**).

To reduce TE and enable measurement of termination rate, shorter asRNAs that pair to the RNA further from the active site were tested. Shortening the asRNA by 1 nt such that it pairs

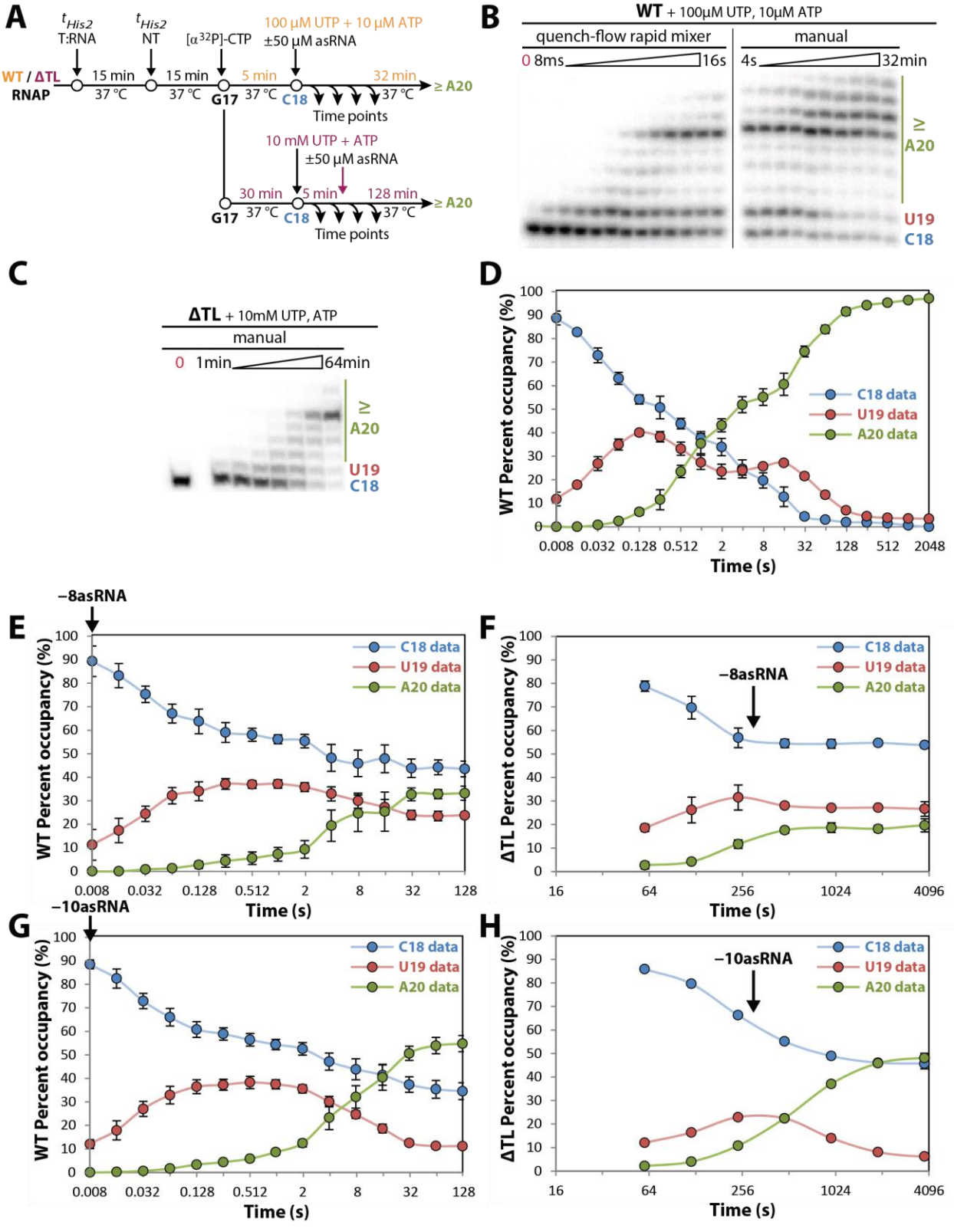
to the −9 RNA position rather than −8 also resulted in nearly 100% termination for ΔTL ECs **(Figure 2.S1C-D)**. However, shortening the asRNA by 2 nt such that it pairs to the −10 position, but no further, resulted in both modest termination and some read-through transcription for both WT and ΔTL ECs **(Figure 2.3G-H** and **Figure 2.S1*E-F*)**, enabling termination rate measurement for slow-elongating mutants.

The WT and ΔTL EC elongation kinetics were determined by fitting the proportion of C18, U19, and A20 complexes present at each timepoint in the absence of asRNA to kinetic models **(Figure 2.3B-D** and **Figure 2.4A-B)** that included C18-to-U19, U19-to-A20, and elemental pausing steps at C18 or U19 as necessary. We define elemental pausing here as a step that causes a fraction of ECs to enter a slow-elongating state (as in **Figure 2.1C**) due to an increased energy barrier to elongation. Such pausing events would increase the likelihood of termination in the presence of the T_{hp} by extending the kinetic window during which termination can occur, before elongation occurs from the paused state (*i.e.* pause escape). A fraction of WT ECs paused at C18 and U19 on t_{his2} in response to the terminator sequence, whereas ΔTL ECs elongated with simple, pseudo-first-order kinetics, consistent with the previous finding that ΔTL RNAP does not exhibit pausing behavior (36). For these and all other reaction conditions, the simplest kinetic model that exhibited a good fit to the data were used to determine rate constants **(Figure S2)**.

The rate of termination was estimated by fitting the relative concentrations of C18, U19, and ≥A20 RNA species observed in the presence of the −10asRNA to a kinetic model that included a termination step of a variable rate but set elongation and U-tract pausing rates to those observed in the absence of the −10asRNA **(Figure 2.5A-B)**. In addition, the rate of asRNA pairing, measured independently using a previously described assay (37) **(Figure 2.S3)**, and variable rates of hairpin-stabilized pausing in the RNA-duplex containing EC states were added to the model to enable fitting the data (see section **2.5.5 Kinetic fitting of**

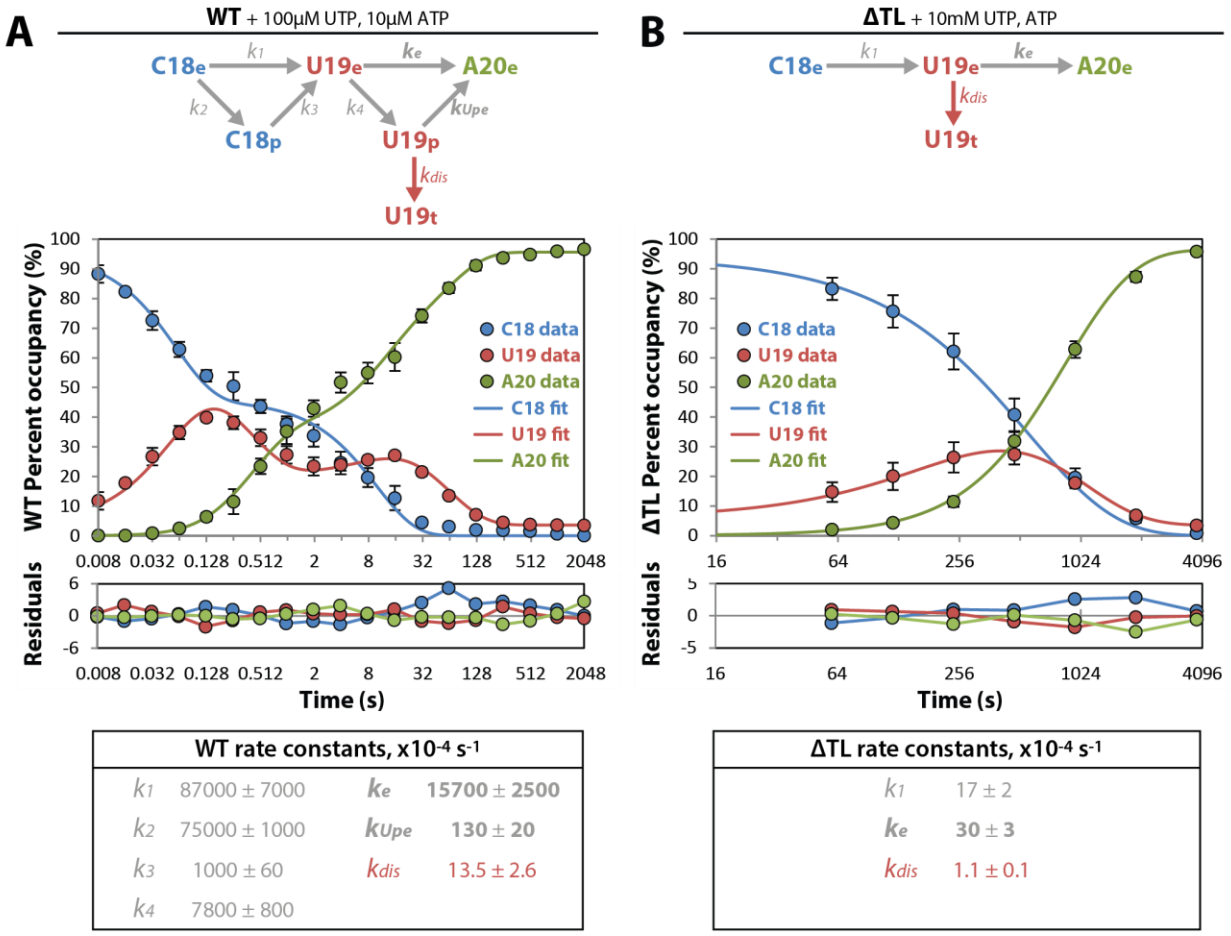
transcription assay reaction progress plots). Thus, in our estimates of termination rates, the rates of –10asRNA-bound pause escape and the termination rates at C18 and U19 were the only free parameters allowed to vary during the kinetic fitting **(Figure 2.5A-B, black arrows)**.

A comparison of the termination rates at U19 for WT and ΔTL ECs revealed that, despite the increased TE observed for ΔTL ECs **(Figure 2.2D)**, the actual rate of termination at U19 was decreased by a factor of ~75 in ECs lacking a TL ([3.5 ± 0.14] x 10^{-4} s^{-1} for ΔTL *vs.* [262 ± 20] x 10^{-4} s^{-1} for WT; **Table 2.1A**). We focus here on the U19 termination rates because the ECs at U19 are engaged in active transcription whereas the ECs at C18 were halted at the site and thus may have rearranged prior to addition of NTPs. We conclude that the intact TL significantly accelerates the rate-limiting step in the termination pathway that competes kinetically with continued elongation.


A
WT / ΔTL RNAP
t His2 T:RNA
t His2 NT
[α32P]-CTP
100 μM UTP + 10 μM ATP ±50 μM asRNA
15 min 37 °C
15 min 37 °C
5 min 37 °C
32 min 37 °C
G17
C18
≥ A20
Time points
10 mM UTP + ATP ±50 μM asRNA
30 min 37 °C
5 min
128 min 37 °C
B
WT + 100μM UTP, 10μM ATP
quench-flow rapid mixer
manual
0 8ms 16s
4s 32min
≥ A20
U19
C18
C
ΔTL + 10mM UTP, ATP
manual
0 1min 64min
≥ A20
U19
C18
D
WT Percent occupancy (%)
C18 data
U19 data
A20 data
Time (s)
E
−8asRNA
WT Percent occupancy (%)
C18 data
U19 data
A20 data
Time (s)
F
−8asRNA
ΔTL Percent occupancy (%)
C18 data
U19 data
A20 data
Time (s)
G
−10asRNA
WT Percent occupancy (%)
C18 data
U19 data
A20 data
Time (s)
H
−10asRNA
ΔTL Percent occupancy (%)
C18 data
U19 data
A20 data
Time (s)

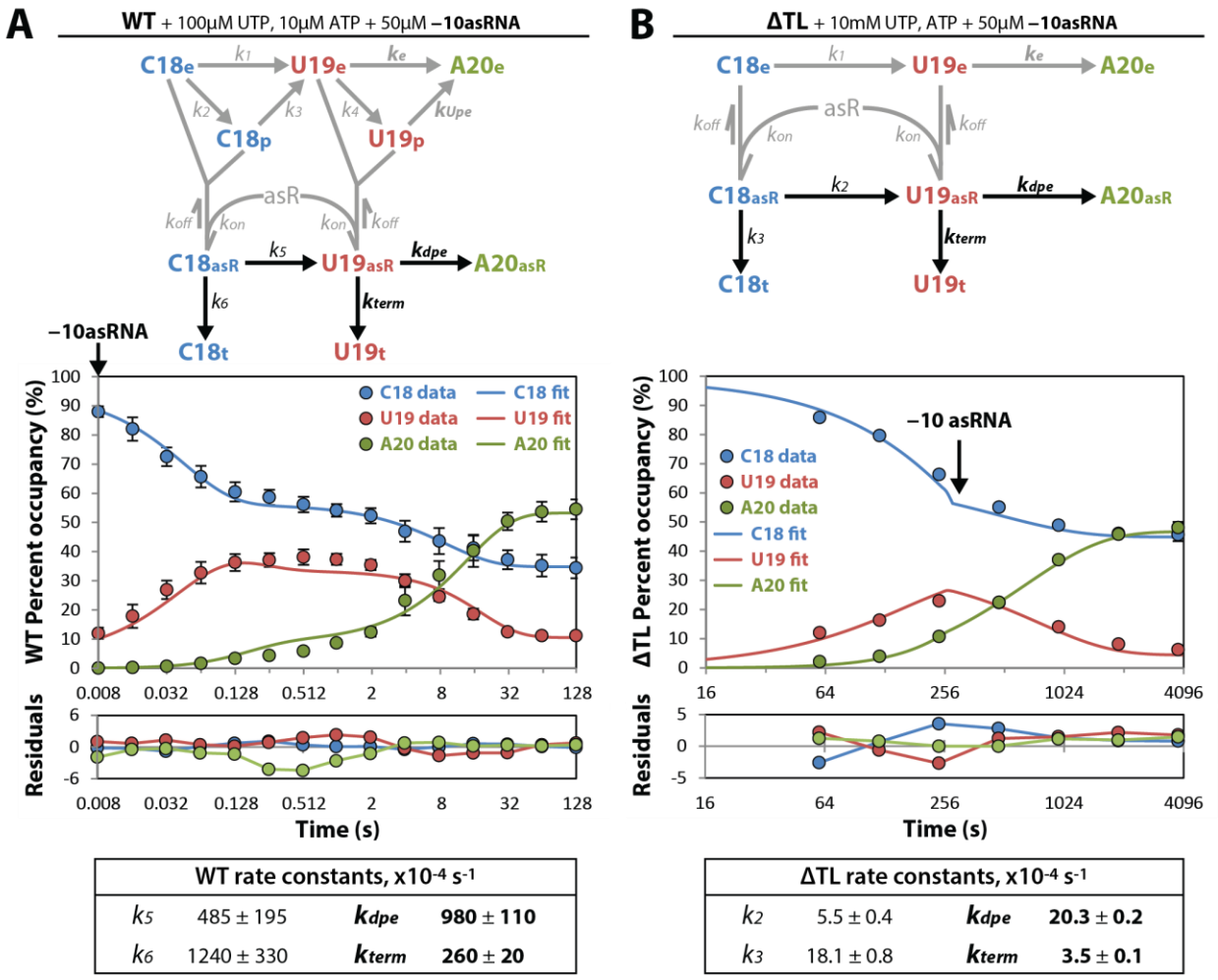
<u>Figure 2.3:</u> Limited hairpin extension enables termination rate measurements for slow RNAP mutants.

(A) Experimental setup used to measure the elongation and termination rates of WT and ΔTL ECs. **(B)** Representative gel and **(C)** reaction progress curves showing the conversion of RNAs from C18 to U19 to A20^{+} for WT ECs in the absence of asRNA. **(D-E)** Reaction progress curves for **(D)** WT ECs with −10asRNA and **(E)** ΔTL ECs with −10asRNA. Time of asRNA addition is indicated with arrows. C18, U19 and A20^{+} occupancies for each timepoint are normalized to total radioactive counts in the lane. Error bars represent SD from ≥3 independent experimental replicates; error bars are smaller than the data markers in some cases.



<u>Figure 2.4:</u> Determination of elongation rates in the absence of asRNA by kinetic fitting.

Kinetic models, the corresponding fits and the rate constants obtained from the fits are shown for **(A)** WT and **(B)** ΔTL ECs, in the absence of asRNA. Residuals between the fits and data points are shown in the lower graphs. $C18_e$, $U19_e$, and $A20_e$ denote elongating ECs with C18, U19, or A20 RNAs, respectively; $C18_p$ and $U19_p$ denote U-tract paused ECs; asR denotes free asRNA; $C18_{asR}$, $U19_{asR}$, and $A20_{asR}$ denote asRNA-bound complexes; $C18_t$ and $U19_t$ denote terminated complexes. Error bars represent SD from ≥3 independent experimental replicates.



<u>Figure 2.5:</u> The TL increases termination rate.

Kinetic models, the corresponding fits and the rate constants obtained from the fits are shown for **(A)** WT with –10asRNA and **(B)** ΔTL ECs with –10asRNA. Residuals between the fits and data points are shown in the lower graphs. $C18_e$, $U19_e$, and $A20_e$ denote elongating ECs with C18, U19, or A20 RNAs, respectively; $C18_p$ and $U19_p$ denote U-tract paused ECs; asR denotes free asRNA; $C18_{asR}$, $U19_{asR}$, and $A20_{asR}$ denote asRNA-bound complexes; $C18_t$ and $U19_t$ denote terminated complexes. Error bars represent SD from ≥3 independent experimental replicates.

Table 2.1: Effects of TL mutations on the termination pathway.

(A) Effects of TL mutations on termination efficiency (TE) and rate. **(B)** Effects of TL mutations on the termination pathway via effects on U-tract and hairpin-stabilized pausing.

A

RNAP TL mutant	asRNA	Termination		
		TE *	termination rate (s^{-1}, $x10^{-4}$), k_{term}	k_{term}(WT + –NasRNA) / k_{term}(mutant + –NasRNA) **
WT	–8asRNA	42 ± 2.4	995 ± 300	
	–10asRNA	17 ± 0.9	260 ± 20	4 ± 1 ***
ΔTL	–8asRNA	87 ± 7	> 175	
	–10asRNA	26 ± 1.6	3.5 ± 0.14	75 ± 6
ΔSI3	–8asRNA	34 ± 6.9	850 ± 150	1.2 ± 0.4
folded; Ala_6ΔSI3	–10asRNA	44 ± 8.6	14 ± 3.0	19 ± 4
unfolded; LTPPΔSI3	–10asRNA	53 ± 1.9	14 ± 1.7	19 ± 3
unfolded; LTPP	–10asRNA	47 ± 7.1	6.6 ± 0.47	40 ± 4

B

RNAP TL mutant	asRNA	[ATP]	elongation rate (s^{-1}, $x10^{-4}$), k_e	U-tract pausing			hairpin pausing	
				PE °	escape rate (s^{-1}, $x10^{-4}$), k_{Upe}	k_e / k_{Upe}	escape rate (s^{-1}, $x10^{-4}$), k_{dpe}	k_e / k_{dpe}
WT	–8asRNA	10 μM	15700 ± 2500	33 ± 2	130 ± 20	119 ± 25	980 ± 380	16 ± 7
	–10asRNA	10 μM					980 ± 110	16 ± 3
ΔTL	–10asRNA	10 mM	29.5 ± 2.7	(none)	(none)	(none)	20.3 ± 0.2	1.5 ± 0.1
ΔSI3	–8asRNA	10 μM	14200 ± 970	17 ± 2	1770 ± 570	8.0 ± 2.6	1080 ± 290	13 ± 4
folded; Ala_6ΔSI3	–10asRNA	10 μM	83 ± 4.0	(none)	(none)	(none)	18.8 ± 3.4	4.4 ± 0.8
unfolded; LTPPΔSI3	–10asRNA	10 mM	61.9 ± 2.1	(none)	(none)	(none)	21.2 ± 2.7	2.9 ± 0.4
unfolded; LTPP	–10asRNA	10 mM	41.8 ± 0.9	(none)	(none)	(none)	13.4 ± 0.3	3.1 ± 0.1

* TE = U19 / (U19 + A20) x 100 for final timepoints; ** 'N' (in '–NasRNA') = 8 or 10; *** k_{term}(WT + –8asRNA) / k_{term}(WT + –10asRNA);
° PE (pause efficiency) = percent of complexes that enter the U-tract paused state, $U19_p$. Errors represent SD from ≥3 independent experimental replicates.

2.3.4 The TL also aids termination by favoring pause entry at the termination sites and disfavoring pause escape

Our kinetic analysis also yielded estimates for the rates and efficiencies of elemental and hairpin-stabilized pausing for WT and ΔTL ECs. For WT RNAP, 33% of ECs enter an elemental paused state, and are slowed by a factor of ~120 and ~16 by elemental pausing and hairpin-stabilized pausing, respectively **(Table 2.1B)**. However, we found that the absence of the TL completely abrogated elemental pausing in response to the U-tract **(Figure 2.4B** and **Table 2.1B)**, indicating that the presence of the TL favors entry into the elemental pause in WT ECs, which would extend the kinetic window during which the T_{hp} can nucleate. We found that hairpin-stabilized pausing is also minimized in the absence of the TL, with the T_{hp} slowing extension of ΔTL ECs from U19 to A20 by only a factor of ~1.5. This is consistent with previous reports demonstrating that the TL is essential for hairpin-stabilized pausing (34, 36, 38). These results indicate that the partially formed T_{hp} does generate a hairpin-stabilized pause species in WT ECs that is significantly stabilized by the presence of the TL. We conclude that the role of the TL in establishing the elemental paused state and decreasing pause escape from the elemental and hairpin-stabilized pauses acts to aid entry of ECs into and retention within the termination pathway.

2.3.5 SI3 does not impact termination rate but aids termination efficiency

The *Escherichia coli* TL also contains a large 188-aa sequence insertion called sequence insertion 3 (SI3; **Figure 2.6A**) that is present in many γ-proteobacterial lineages and is structurally related to similar insertions in cyanobacteria (39, 40). The role of SI3 in termination was also of interest, since this insertion has been shown to affect other activities of the TL, including elongation, pausing, and intrinsic cleavage (34, 41, 42).

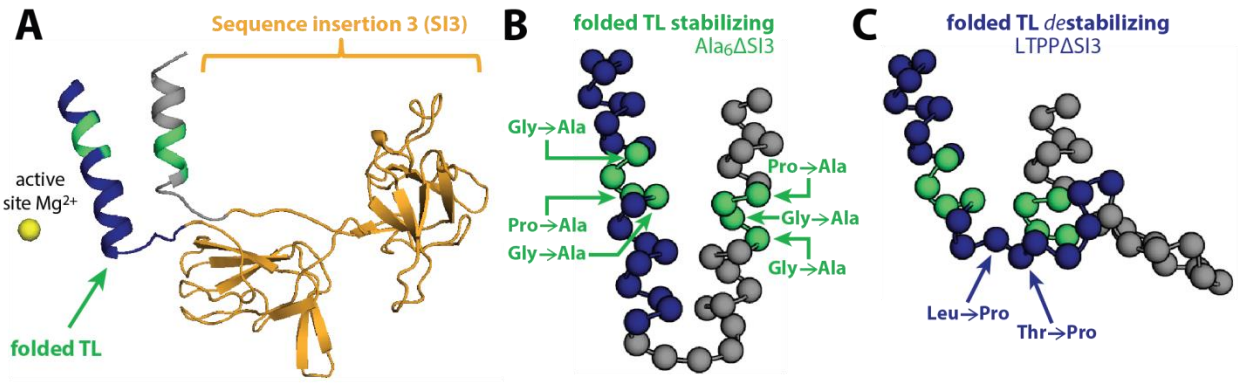
To test whether the TL sequence insertion SI3 modulates TL effects on termination, we determined the elongation, pause, and termination rates of an SI3 deletion mutant (ΔSI3) relative to WT RNAP using our termination assay. No termination was detected for ΔSI3 ECs with −10asRNA, likely due to decreased pausing of ΔSI3 ECs (see below and (38)). Thus, to compare WT and ΔSI3 ECs, we measured termination rates in response to the stronger termination signal created by annealing the −8asRNA **(Figure 2.2A)**. ΔSI3 ECs were assembled at G17 on the t_{His2} scaffold, the RNAs radiolabeled at C18 by the addition of [α^{32}P]-CTP, and elongated through U19 with 10 μM UTP & ATP in the absence or presence of −8asRNA **(Figure 2.S4A)**. Termination kinetics were measured for WT ECs with −8asRNA **(Figure 2.4A)** as described for the −10asRNA. C18, U19 and ≥A20 RNA populations were then used to determine the elongation, pause and termination rates for both WT and ΔSI3 ECs, as described above.

Deletion of SI3 had little effect on the termination rate **(Table 2.1A)**. We conclude that the effect of the TL deletion on termination rate is not a consequence of deleting SI3 in the ΔTL RNAP, but rather that the large effect of the TL on termination rate is caused by the segments of the TL known to undergo helix-to-coil transitions during nucleotide addition.

Although SI3 did not affect termination rate, it did alter TE through effects on elemental pausing. Our data show that deletion of SI3 resulted in a significant increase in the rate of pause escape from the U-tract elemental pause **(Table 2.1B)**, corresponding to a reduction in elemental pause stimulation by a factor of ~15. Deletion of SI3 additionally decreased the fraction of ECs that entered the paused states (~17% for ΔSI3 *vs.* ~33% for WT). However, in contrast to the ~2–4-fold stimulatory effect of SI3 on hairpin-stabilized pausing observed at non-terminator sites (34, 42), deletion of SI3 only reduced hairpin-stabilized pausing in response to the partial T_{hp} by a factor of ~1.2 **(Table 2.1B)**. This may be due to differences between canonical hairpin-stabilized pause hairpins, which only form RNA–RNA bp up to −12,

and the longer hairpin likely formed by the partially folded T_{hp}, since RNA–RNA base-pairing closer to the active site has been shown to reduce hairpin-stabilized pausing (43). However, as a consequence of less pausing overall, TE for ΔSI3 RNAP was 34% *vs*. 42% for WT, indicating that stimulation of pausing caused by the presence of SI3 yielded an overall increase in TE by increasing the flux of ECs into and decreasing escape from the termination pathway, even though the rate of termination was minimally affected.

Termination of WT ECs in response to the –8asRNA is more efficient than for the –10asRNA ([42 ± 2.4] % *vs*. [17 ± 0.3] %, respectively; **Table 2.1A**), as expected with the stronger termination signal. The rate of termination was also ~4-fold faster for the –8asRNA ([995 ± 300] x 10^{-4} s^{-1} for –8asRNA *vs*. [262 ± 20] x 10^{-4} s^{-1} for –10asRNA; **Table 2.1A**), indicating that a step enhanced by extension of the T_{hp} is at least partially rate-determining in EC inactivation at a terminator.



<u>Figure 2.6:</u> TL sequence insertion 3 (SI3) & TL conformational mutants.

(A) A model of SI3 and its position in the TL are shown. Models of the **(B)** folded TL stabilizing mutant Ala_6ΔSI3 and **(C)** folded TL destabilizing mutant LTPPΔSI3 are shown with spheres marking each residue. Labeled arrows describe the amino acid substitutions made in each mutant. Green residues denote the TL hinge regions; blue residues denote the active site-proximal arm of the TL.

2.3.6 Reducing TL conformational flexibility inhibits termination rate

As described above, the presence of the TL increases termination rate. However, it was unclear if a particular TL state (*i.e.*, folded into the TH as proposed by the TH–T_{hp} contact model, partially folded, or unfolded) favors termination or if the overall conformational flexibility of the TL was important rather than a particular TL state. To address this question, we measured the termination rates for previously characterized TL variants that either stabilize (Ala_6ΔSI3; **Figure 2.6B**) or destabilize (LTPPΔSI3; **Figure 2.6C**) the folded TL conformation by placing either α-helix-stabilizing alanines in the TL hinge regions, or α-helix-destabilizing prolines below the active site-proximal arm of the TL, respectively (33, 34). We studied these TL variants in the absence of SI3 so that we could distinguish the effects of the TL on termination rate from effects of SI3. As for ΔTL ECs, termination rates could not be determined for Ala6ΔSI3 and LTPPΔSI3 ECs with –8asRNA due to nearly 100% TE upon asRNA addition **(Figure 2.S5)**. Ala6ΔSI3 and LTPPΔSI3 ECs were thus assayed for termination on t_{his2} in the absence or presence of –10asRNA as described above **(Figure 2.S6)**.

If the TL must adopt a folded conformation during termination as proposed in the TH–T_{hp} contact model (1), one would predict that the termination rate for Ala_6ΔSI3 ECs would be similar to or faster than WT ECs. Instead, we found that Ala_6ΔSI3 ECs terminated ~19 times slower than WT ([14 ± 3] x 10^{-4} s^{-1} for Ala_6ΔSI3 *vs.* [262 ± 20] x 10^{-4} s^{-1} for WT; **Table 2.1A**), indicating that the TH-stabilizing mutations are not favorable for termination. Rather, this finding suggests that stabilizing the folded TL conformation (TH) decreases the rate of termination, in contradiction to the TH–T_{hp} contact model (1).

To our surprise, substitutions that disfavored TH formation also decreased termination rate. Most notably, ECs containing the TH-destabilizing LTPP substitutions, which place tandem proline residues below the hinge of the active site-proximal arm of the TL and thus

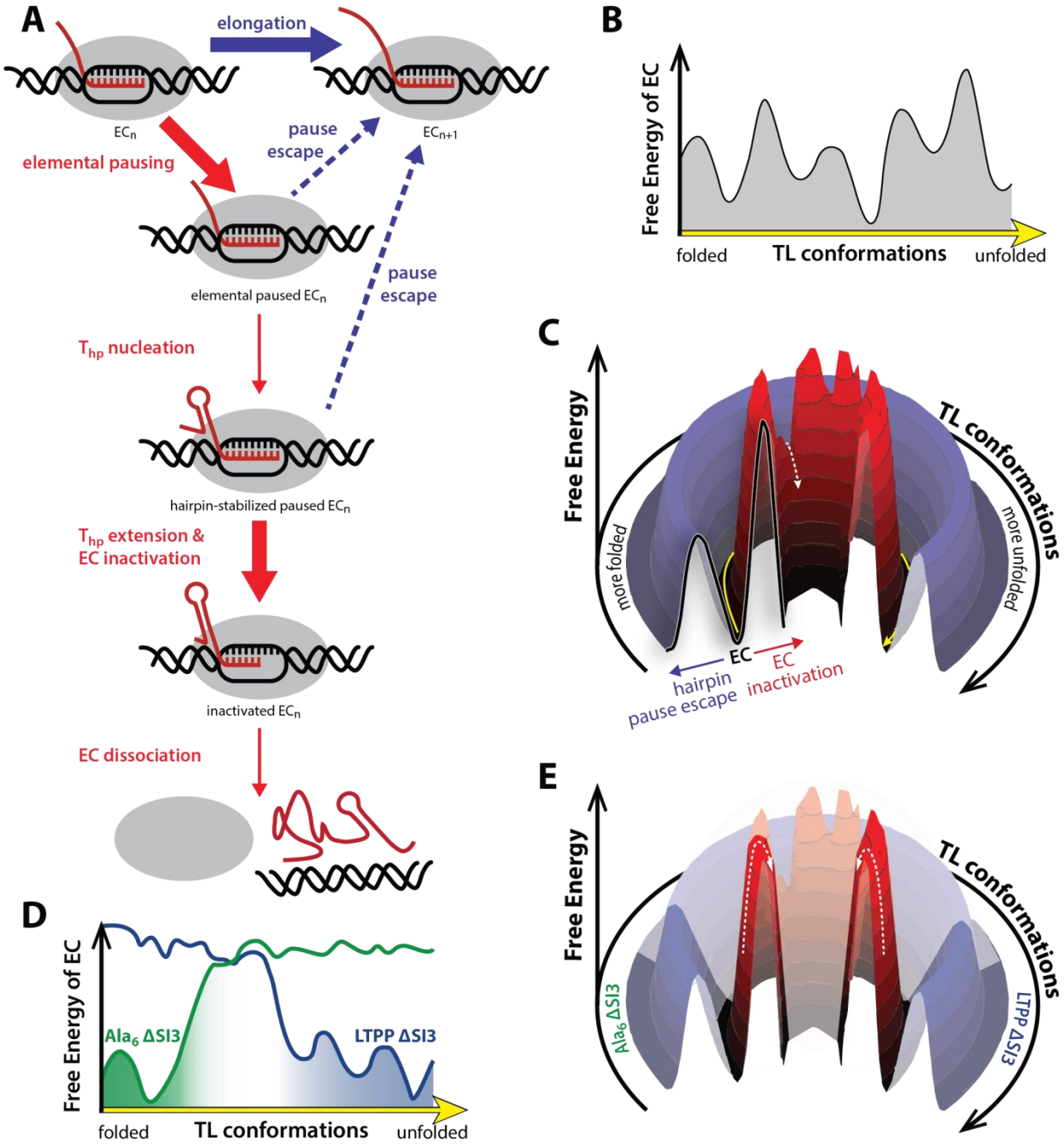
block TH formation **(Figure 2.6C)**, also decreased termination rate by a factor of ~19 ([14 ± 1.7] x 10^{-4} s^{-1} for LTPPΔSI3; **Table 2.1A**). This indicates that neither the folded nor unfolded TL *per se* is a conformation that stimulates termination rate, but rather is consistent with an interpretation that restricting TL conformational space reduces termination rate due to increased energy barriers to termination. Although these mutants increase the occupancy of the folded or unfolded TL states, less than 100% of these complexes occupy their respective favored TL state (34). That is, these variants are able to adopt more than one TL conformation, but undergo increased sampling of the folded or unfolded states due to restricted motion. The effect of these mutations on termination rate is thus likely due to the restricted dynamics of the TL, which would consequently be expected to limit overall EC dynamics through the allosteric connections of the TL with the BH, clamp, and exit channel (37) **(Figure 2.1D)**.

Although SI3 did not have a significant impact on termination rate, we wanted to determine whether the presence of SI3 affects a TL that is already conformationally restricted. To do this, we tested an LTPP TL variant in which SI3 was still present **(Figure 2.S7)**, and compared it to LTPPΔSI3. The corresponding Ala6 RNAP variant containing SI3 could not be tested due to the formation of cleavage products that were then extended into the region of analysis, confounding elongation and termination rate determination **(Figure 2.S5)**. We found that the termination rate of the LTPP TL variant was decreased by a factor of ~2 relative to LTPPΔSI3 **(Table 2.1A)**. The increased impairment of the unfolded TL in the presence of the bulky insertion further supports the idea that TL conformational dynamics allow ECs to access low energy transition states on the pathway to termination, resulting in faster overall termination rates. We also note that although the termination rates are severely decreased for the TL conformational variants (Ala_6ΔSI3, LTPPΔSI3 & LTPP) relative to WT, all three mutants terminate ~2–4 times more rapidly than ΔTL ECs. These data would suggest that in the absence of the TL, EC conformational dynamics are limited still further. TL dynamics therefore likely

play a central role in increasing the conformational space accessed by terminating ECs, thereby allowing the complexes to sample low energy pathways to termination.

2.4 DISCUSSION

Our study revealed effects of TL conformational states on the rate of intrinsic termination by exploiting an assay that derives termination rates from measurements of elongation rates, pause rates, and termination efficiency. This approach was necessary to study the effects of the TL on termination because TL alterations can profoundly impact elongation rates and pause propensities of ECs, and can thus greatly affect overall termination efficiency irrespective of effects on termination rate. Using this assay, we were able to isolate individual steps in the kinetically complex termination pathway and gain the following key insights into the role of the TL at each decision point along the termination pathway **(Figure 2.7A)**. First, the presence of the TL enhanced the rate of termination, independent of its effects on pausing and elongation. Second, the conformationally polymorphous arms of the TL, not the 188-aa SI3 module inserted in the TL, were responsible for the observed effect on termination rate. Third, both the TL and SI3 increase the flux of ECs into the termination pathway by promoting pausing. Finally, in contradiction of the TH–T_{hp} interaction model (1), we find that conformational fluctuations of the TL – that is the ability to occupy different conformations, and not the folded TH conformation (required for contact to the T_{hp}) *per se* – favor termination by accelerating the rate-limiting step of termination. Based on these findings we propose a model in which a dynamic TL facilitates increased conformational fluctuations of the EC, thereby allowing access to low energy paths to termination.



<u>Figure 2.7:</u> TL dynamics increase conformational fluctuation of ECs (the TL-facilitated dynamic EC model) and favor multiple steps on the termination pathway.

(A) The role of the TL in the defined steps of the termination pathway is depicted. The TL acts in concert with T_{hp} extension to accelerate EC inactivation. A combined effect of the TL and SI3 also increase overall termination efficiency by increasing the flux of ECs into the termination pathway, and preventing escape from the terminator by elongation or pause escape. Steps that are aided by the TL are indicated by bold arrows; steps inhibited by the TL are indicated by dashed arrows. **(B)** Theoretical energy diagram depicting the effect of TL conformation in altering EC stability. **(C)** Theoretical energy landscape

depicting the role of TL conformational dynamics in allowing the EC to sample various pathways to termination with differing energy barriers. The *white dashed arrow* indicates the lowest available energy barrier that is likely most frequently traversed by ECs; the *yellow arrow* denotes the region of the 3D projection analogous to the x-axis in **(B)**. **(D)** Theoretical energy diagram depicting the decreased EC stability with the unfolded (for Ala_6ΔSI3, *green*) or folded (for LTPPΔSI3, *blue*) TL conformations due to restricted TL dynamics. **(E)** *Highlighted areas* depict the limited set of conformations accessed by ECs when TL dynamics are restricted to the folded and unfolded conformations, as in **(D)**. These limited sampling spaces may restrict the ECs to regions that have fewer low-energy paths to termination, or that require them to overcome high local energy barriers to termination. The *white dashed arrows* indicate the lowest available energy barriers in the restricted array of TL conformations. In the absence of the TL, the ECs would similarly be unable to sample a wide range of conformations, and low energy pathways to termination that are dependent on TL dynamics would be inaccessible.

2.4.1 A model for termination aided by TL conformational dynamics

The thermodynamic model for intrinsic termination posits that the relative energy barriers to elongation versus termination determine the fraction of complexes that terminate via kinetic competition between the two pathways (3, 4) **(Figure 2.1B)**. The simplest form of this model is that ECs terminate via a single pathway involving one set of changes in the RNAP–NA contacts that constitute a major component of the energy barrier to termination. This view was sufficient prior to elucidation of the structure of ECs and the realization that modules like the clamp, BH, and TL occupy interconnected but distinct conformational states, resulting in clamp opening/closing, BH bending, and TL folding into the TH (24-27, 37, 44, 45).

Our results revealed a striking ~75-fold stimulatory effect of the TL on termination rate **(Table 2.1A)**, indicating that the TL is a crucial module in facilitating the penultimate step of termination: T_{hp} extension and EC inactivation **(Figure 2.1C)**. We propose that acceleration of termination by the TL results from its contribution to conformational fluctuations in the EC, consistent with our finding that the TL itself – rather than the folded or unfolded TL conformations *per se* – accelerates termination rate. We propose an extension of the thermodynamic model to a multistate-multipath (MS-MP) model, in which different interconverting EC conformations explore multiple, alternative termination pathways. In this view, the terminating EC exists in a family of conformational states with fluctuating RNAP–NA contacts, where in some EC states, the RNAP–NA interactions may be weakened more than in others. These conformational fluctuations would then result in variable EC stabilities **(Figure 2.7B)** and a rugged landscape of energy barriers to termination **(Figure 2.7C)**. Specifically, we propose that conformational fluctuations of the TL increase the conformational fluctuations in the terminating EC, creating multiple pathways of EC termination with differing energy barriers to EC inactivation. Variability in energy barriers caused by differing complex stabilities has precedent in studies of ligand binding, where differences between protein conformational

substates and their relative stabilities is thought to cause variations in the activation energy of ligand binding (46-49); the idea of multiple parallel routes for a single reaction is analogous to the well-documented concept of multiple protein folding pathways (48-50). The hybrid-shearing and hypertranslocation termination models (6, 17, 18) present two such examples of termination pathways, in which the interactions between the RNA and DNA are altered along with the RNAP–NA contacts. One can envision other routes to termination in which either or both of these sets of interactions are affected as a consequence of EC conformational state.

In the MS-MP model presented here, when the TL is deleted, or its conformations restricted by alterations, the number of accessible conformational states of the terminating EC is correspondingly reduced. Thus a dynamic TL that fluctuates among various conformations will enable the EC to explore multiple routes of termination **(Figure 2.7C)**, thereby increasing the overall termination rate. However, TL deletion, or restriction of TL conformational fluctuations **(Figure 2.7D)** reduces termination rate by reducing the number of available low-energy paths, or by restricting the EC to paths of relatively higher energy **(Figure 2.7E, highlighted areas)**. The termination rates obtained from our ensemble experiment are the average rates of all molecules in solution as a function of their respective barrier heights to termination. Therefore, they represent the accessibility of low-energy pathways to termination. This view of the termination pathway is similar to the original thermodynamic terminal model developed by von Hippel and co-workers (3), but adds a dimension of protein conformational states that were not fully understood when the model was developed.

2.4.2 The folded TL is unfavorable for termination, at odds with the TH–T_{hp} contact model

It was previously suggested in the TH–T_{hp} contact model that a folded TL is required for termination because a monoclonal antibody (mAb) that constrained the TL in an unfolded conformation slowed dissociation of stalled ECs (1). Based on the observation of crosslinks formed between the loop and upper stem of the T_{hp} and the TL, it was further proposed that the T_{hp} invades the main cleft of RNAP in a terminating complex to contact the folded TL (1). However, others report that the low salt conditions used in this study allowed dissociated ECs to rebind the released RNA and form binary RNA–RNAP complexes (32, 51), which are likely the source of the observed crosslink.

Our results also contradict the TH–T_{hp} contact model. We found a TL that resides predominantly in the folded state decreased the termination rate relative to WT (Ala_6ΔSI3 *vs.* WT RNAP; **Table 2.1A**), arguing against a role for a folded TL, as proposed by the TH–T_{hp} contact model. Consistently, whereas ECs lacking SI3 sample the folded TL state ~10 times more frequently than the WT TL (34), we found that deletion of SI3 has minimal effect on the termination rate, further arguing against a stimulatory role of the folded TL in the termination mechanism. Rather, we found that RNAP mutants that bias the TL conformation towards either the folded or unfolded TL states slow the termination rate by a factor of ~20 **(Table 2.1A)** – only ~4-fold faster than a complete deletion of the TL, in which TL dynamics are completely absent and ECs likely occupy a very narrow conformational space. Moreover, we found that *(i)* dissociation of folded TL-favoring Ala_6 or ΔSI3 ECs stalled at a termination site did not exhibit increased dissociation rates relative to WT (see **Appendix A**, **Figure A.2**), and *(ii)* preliminary tests suggest that folded TL-stabilizing microcin J25 (MccJ25) also inhibits EC dissociation (see **Appendix B**, **Figure B.1**) similar to the unfolded TL-stabilizing mAb used in the previous study. These findings suggest that the key contribution of the TL to termination lies

in its ability to expand EC conformational diversity rather than any particular conformation. Thus, we favor the MS-MP model, in which the TL must be dynamic to sample one of many states that have decreased energy barriers to termination.

It is also of note that the MS-MP model reconciles conflicting reports on the role of the TL in intrinsic termination versus Rho-dependent termination (1, 2); the EC would be expected to be analogously destabilized by either the Rho protein or the T_{hp}. Whereas the TL was proposed to be folded in the TH–T_{hp} contact model, it was suggested that the TL must be unfolded for Rho-dependent termination due to decreased dissociation of stalled ECs caused by the folded TL-stabilizing small molecule inhibitor tagetotoxin (Tgt). Consistent with the MS-MP model, both the mAb used in the intrinsic termination experiments and the Tgt used in Rho-dependent termination experiments would constrain TL conformation and decrease the number of available paths to termination. Therefore, it is likely that similar to intrinsic termination, Rho-dependent termination may require a dynamic TL and EC to access low energy paths to termination.

2.4.3 TL dynamics and T_{hp} extension determine the rate of EC inactivation

Our results indicate that T_{hp} extension and TL dynamics both affect the rate-limiting step in the termination pathway, since reducing either significantly slows termination rate. Moreover, these two contributions to the rate appear to be coupled, likely mediated by the energetic linkage that exists between the active site and RNA exit channel (37).

Our results confirm the previously suggested hypothesis that the second step of T_{hp} folding – T_{hp} extension – accelerates termination. Two groups have shown by independent methods that T_{hp} folding occurs in two steps: pairing of all but the terminal 2–3 bps of the T_{hp}, followed by extension of the T_{hp} by pairing of the last few bps (16, 17). Work from these as well as

the von Hippel group also suggests that extension of the T_{hp} provides the energy necessary to destabilize the EC by causing the upstream 3–4 bps of the U-tract to melt, and minimizing the RNA–DNA hybrid component of the energy barrier to termination (4, 6, 17) **(Figure 2.1B)**. Consistent with this idea, we find that increasing T_{hp} extension from ~1 bp to the canonical ~3 bps accelerates the rate of EC inactivation ~4-fold ([995 ± 300] x 10^{-4} s^{-1} for –8asRNA *vs.* [262 ± 20] x 10^{-4} s^{-1} for –10asRNA; **Table 2.1A**), indicating that when T_{hp} extension is diminished, ECs experience an increased barrier to termination. Therefore, consistent with the proposed models, T_{hp} extension likely supplies a significant portion of the activation energy required to overcome the barrier to EC inactivation **(Figure 2.7C)**, which is diminished when T_{hp} extension is limited. We thus infer that extension of the T_{hp} contributes to rate determination of EC inactivation in a manner consistent with the proposed models.

The TH–T_{hp} contact model predicts that inhibiting the contact by compromising either T_{hp} extension or TL folding would reduce termination, but that inhibiting both might have a less than an additive effect. Stated in the opposite direction, the model predicts that adding either (*i*) T_{hp} extension upto –8 or (*ii*) the TL to ΔTL ECs with –10asRNA would have a lesser effect alone than combining T_{hp} extension and the TL to facilitate the TH–T_{hp} contact. We observed the opposite relationship; adding either the TL or T_{hp} extension to –8 greatly stimulated termination, but their combined effect was less than additive. Increasing the extent of RNA-duplex extension from –10 to –8 (relative to the RNA 3′ end) accelerated termination ~4-fold in WT ECs but >50-fold in ΔTL ECs (2.6×10^{-2} s^{-1} to ~10^{-1} s^{-1} for WT *vs.* 3.5×10^{-4} s^{-1} to >1.7×10^{-2} s^{-1} for ΔTL; **Table 2.1A**), indicating that when the TL is present, the benefit of T_{hp} extension is reduced. Conversely, adding a flexible TL to ΔTL RNAP increased termination ~75-fold for –10asRNA, but only <6-fold for –8asRNA (3.5×10^{-4} s^{-1} to 2.6×10^{-2} s^{-1} for –10asRNA *vs.* >1.7×10^{-2} s^{-1} to ~10^{-1} s^{-1} for –8asRNA **Table 2.1A**), again indicating that the effects of TL addition and T_{hp} extension are less than additive. Taken together, these results suggest that

either T_{hp} extension or a flexible TL can promote termination *via* pathways with lower free energy barriers, but that combining them does not additively lower the barrier to termination. This result is readily explained by the MS-MP model, but appears at odds with the TH–T_{hp} contact model of termination. If T_{hp} extension is limited, the energy supplied by the pairing would also be decreased, but TL dynamics would enable the EC to sample alternate conformations with decreased local energy barriers to termination **(Figure 2.7C)**, thereby allowing appreciable levels of termination. If, however, TL dynamics are also restricted by mutation or TL deletion **(Figure 2.7E)**, the ECs would be unable to explore alternate routes to termination, and the decreased energy produced by the limited T_{hp} extension could result in significantly diminished termination rates depending on the heights of local energy barriers. Restriction of TL dynamics may therefore be a potential means of termination regulation by extrinsic factors; as such, SI3 presents an attractive access point as the solution-exposed domain of the TL (see section **2.4.5 SI3 as a regulatory target for intrinsic termination**).

The coupling of TL dynamics and T_{hp} extension in termination is conceivable even without invoking a TH–T_{hp} contact (1), which still lacks a description of (*i*) a path by which the T_{hp} could move from the RNA exit channel to the main cleft without dissociation, and (*ii*) how the RNA–DNA hybrid and T_{hp} can both be accommodated in the narrow main cleft of RNAP. Rather, a simple explanation for how both the T_{hp} and TL could affect a common rate-limiting step in termination is that both affect the stability of the closed *vs.* open clamp and that clamp movements could be rate-limiting for termination. Extension of exit-channel duplexes even to –12 appears to alter clamp conformation (37) and further extension may have an even greater effect (37). Extension of the T_{hp} would require the –10 RNA base that is normally flipped into a pocket formed by switch 3 (Sw3) in the post-translocated state (8) (–11 in the pre-translocated state) to be removed from the Sw3 pocket for –10 RNA–RNA base pairing in the extended T_{hp}. And as the 5 switch elements (Sw1-5) are thought to mediate opening and closing of the clamp module, switch remodeling upon T_{hp} extension may alter clamp state. Consistent with this idea,

mutations have been identified in switch elements that alter TE (28, 52). The TL also may affect clamp conformation via its contact to the BH and the interactions of the BH-TL module with the switch regions 1 and 2 (37, 38). Thus, both T_{hp} extension and TL movements could affect conformational fluctuations of the clamp that lower the energy barrier to termination.

Thus, we propose that while the movements of the TL alter the distribution of ECs on the energy surface, thereby facilitating access to low energy pathways to termination, extension of the T_{hp} provides the energy necessary to overcome the energy barrier to termination. Moreover, our data suggest that the contributions of these two processes may be coupled, providing complementary mechanisms to restructure and inactivate ECs on the termination pathway.

2.4.4 The TL promotes both elemental and hairpin-stabilized pausing at an intrinsic terminator

Two pausing steps have been proposed to occur on the termination pathway: elemental pausing in response to the sequences of the U-tract (14, 15), and hairpin-stabilized pausing mediated by the partially formed T_{hp} (16, 43, 53) **(Figure 2.1C)**, both of which would act to increase the kinetic window during which termination can occur. Elemental pausing is defined as a sequence-dependent change in EC conformational state that causes a fraction of ECs to enter a slow-elongating state. Elemental pauses can be further stabilized into longer-lived pauses by backtracking, or by the formation of a RNA duplex in the exit channel (*i.e.* hairpin-stabilized pausing). However, backtracking is unlikely to be part of the intrinsic termination pathway as it would preclude formation of the T_{hp}. Although the role of elemental pausing in termination is well-supported, the existence of a hairpin-stabilized pause at terminator sequences remains uncertain.

Our results strongly support the idea that both of these pause steps occur on the termination pathway, and further, that the TL significantly stabilizes both paused states. We find that ~30% of WT ECs enter an elemental pause for which nucleotide addition is slowed by a factor of ~120 **(Table 2.1B)**, whereas in the absence of the TL, all ECs elongate at the same rate, suggesting that ΔTL ECs do not enter an elemental paused state in response to the U-tract **(Table 2.1B)**. Our results further indicate that the partially formed T_{hp} stimulates hairpin-stabilized pausing in a TL-dependent manner. We found that nucleotide addition of WT ECs was slowed by a factor of ~16 by the presence of the T_{hp} **(Table 2.1B)**, whereas ΔTL ECs were only slowed by a factor of ~1.5. Moreover, a comparison of the Ala_6ΔSI3 and LTPPΔSI3 TL variants to WT ECs indicates that these mutations render the TL ineffective at stabilizing the hairpin-stabilized paused state **(Table 2.1B)**. These findings are consistent with previous reports that the TL significantly stabilizes hairpin-stabilized pauses (36), and that destabilizing the TL conformation favorable for hairpin-stabilized pausing minimizes both pause efficiency and duration (34, 38). Interestingly, we find that the elemental pausing at this termination sequence slows nucleotide addition more than the partial T_{hp}, which is distinct from the effect of hairpin formation at regulatory hairpin-stabilized pause signals (36). However, this finding is consistent with previous reports that hairpins extending closer to the active site than the –12 RNA base are not as effective in stimulating hairpin-stabilized pausing as hairpins positioned 1- to 2-bp further from the active site (43). Our results thus corroborate previous reports of U-tract-mediated elemental pausing, and strongly support the idea that a hairpin-stabilized paused state is an intermediate along the termination pathway. Moreover, we demonstrate that TL plays an important role in the stabilization of both paused states. The TL therefore not only stimulates the termination mechanism, but also affects TE by determining the proportion of complexes that enter the termination pathway by pausing on the U-tract, and the proportion of *those* ECs that are retained within the termination pathway and ultimately terminated **(Figure 2.7A)**.

2.4.5 SI3 as a regulatory target for intrinsic termination

Our results with the SI3 deletion mutant yielded two key findings: (*i*) that SI3 is not responsible for the role of the TL in EC inactivation at the termination site, and (*ii*) that SI3 *is* important, however, for increasing overall TE by stabilizing the elemental paused state on the terminator U-tract.

The finding that complete deletion of SI3 had minimal effect on termination rate makes the important distinction that the role of the TL in diversifying EC conformational states lies in the ~44-aa polymorphous arms of the TL, rather than in the large 188-aa insertion formed by SI3. This idea is supported by the fact that mutations that prevent proper folding/unfolding of the TL hinge regions are sufficient to decrease termination rate by a factor of ~19 (**Table 2.1A**, compare effect of ΔSI3 *vs*. WT and Ala_6ΔSI3/LTPPΔSI3 *vs*. WT). Thus SI3 is not important for enabling EC conformational fluctuations during EC inactivation.

The role of SI3 in stabilizing the U-tract elemental pause is significant. The presence of SI3 not only increased pause stabilization ~15-fold, it also increased the proportion of paused complexes from ~17% to ~33% **(Table 2.1B)**. The effect of SI3 on hairpin-stabilized pausing in response to the partial T_{hp} was not pronounced on our termination sequence **(Table 2.1B)**. However, this may be due to the stabilization of an EC conformation that is distinct from canonical hairpin-stabilized pausing as a consequence of RNA–RNA bp formed closer to the RNAP active site by the partial T_{hp}. Overall, the effect of SI3 deletion on pausing indicates that SI3 modulates TL conformation during pausing at the termination site, and acts to increase TE at the termination site. It remains to be determined whether U-tract pause-stabilization by SI3 is stabilized by interactions between SI3 and the jaw domain, as has been shown for regulatory hairpin-stabilized pausing (34).

We also note that as the surface-exposed domain of the TL, SI3 may present a binding site for extrinsic regulatory factors that can modulate the proportion of complexes that enter or escape the termination pathway. A mAb that binds SI3 has been shown to inhibit nucleotide addition and intrinsic cleavage (41), indicating that activities facilitated by the TL can be extrinsically regulated through SI3. It was also shown that another mAb that restricts TL motion – ostensibly through an SI3 epitope – decreases dissociation of ECs stalled at a termination site ~3-fold (1), suggesting that termination could similarly be regulated by SI3-binding factors. Any such role of SI3 as a regulatory access point to termination would be expected to be unique to those organisms in which the insertion is present. However, termination may be analogously regulated by alternative mechanisms in other bacterial lineages as well as in eukaryotes and archaea through interactions with the clamp or flap domains, which also present solvent-accessible surfaces as possible targets for regulation.

2.4.6 Conclusion

We have demonstrated that along with NA rearrangements, RNAP dynamics play a significant role in driving termination, and we have isolated the contribution of the TL to each decision point of the termination pathway. Some eukaryotic RNA secondary structures have also been proposed to cause termination of RNA polymerase II (54, 55), and bacterial intrinsic terminators have been shown to stimulate termination of yeast RNA polymerase II (6). Moreover the HIV-1 TAR hairpin causes termination of human RNA polymerase II transcription (56), suggesting a greater conservation of the intrinsic termination mechanism than has been appreciated previously. It has also been demonstrated that mutations in the TL of RNA polymerase III affect TE (57). It will be interesting to determine whether the role of TL dynamics in RNA-structure-mediated termination is conserved in eukaryotes and archaea.

The effect of the TL in diversifying EC states highlighted in this work is likely to impact other steps of the transcription cycle as well as termination. The TL is an essential mobile element in all multi-subunit RNA polymerases, and the conserved requirement for TL flexibility has long been understood within the context of the nucleotide addition cycle (34, 35, 38). However, it is not widely appreciated how TL dynamics may participate in the broader dynamics of RNAP throughout transcription as a whole. Although RNAP is known to undergo constant thermal fluctuations and numerous allusions to conformational changes of RNAP and other RNA polymerases have been made, no connection to the TL has yet been discussed to our knowledge. The MS-MP model thus proposes the first direct link between TL fluctuations, EC state and their impact on RNAP activity. And analogous to its effects on termination, the effects of TL dynamics on overall RNAP state will almost certainly affect other steps of the transcription cycle.

Other highly dynamic or mobile domains of RNAP such as the clamp/switch, BH, and flap domains are also likely to play important roles in the termination pathway and mechanism. This assay provides an easily adaptable method to determine what their contributions are for each step of the termination pathway, and help characterize the structural determinants that destabilize the extraordinarily stable EC at an intrinsic terminator.

2.5 Materials & Methods

2.5.1 Materials

Non-fluorescent DNA and RNA oligonucleotides **(Table 2.S2)** were obtained from Integrated DNA Technologies (IDT; Coralville, IA); the fluorescent pyrrolo-C labeled oligonucleotide (8817) was obtained from TriLink BioTechnologies (San Diego, CA). All oligonucleotides were purified by 8 M urea denaturing polyacrylamide gel electrophoresis (PAGE) before use. [α-^{32}P]-CTP was obtained from PerkinElmer Life Sciences (Warwick, RI); NTPs were obtained from Promega (Madison, WI).

2.5.2 Proteins

His_{10}-tagged core wild-type and mutant *E. coli* RNAPs **(Table 2.S1)** were purified from overexpression plasmids transformed in *E. coli* strain RL2657 as described previously (34).

2.5.3 *In vitro* EC reconstitution

The nucleic acid scaffold for EC reconstitution was formed by mixing 5 μM G17 RNA and 10 μM T-DNA (10002 and 8451, respectively; **Table 2.S2**) in reconstitution buffer (RB; 10 mM Tris-HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$) and heating to 95 °C for 2 min, cooling rapidly to 45 °C, then cooling to 25 °C in 2 °C increments for 2 min each, as described previously (58). The 17 nt RNA was designed to have 8 nt complementarity to the T-DNA at the site of reconstitution to prevent base pairing of upstream T-DNA and RNA, further preventing backtracking off of the termination site.

ECs were reconstituted by incubating 2.5 μM core *E. coli* RNAPs with 0.5 μM nucleic acid scaffold in transcription buffer (TB; 20 mM Tris-OAc, pH 8.0, 75 mM NaOAc, 1 mM $Mg(OAc)_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 2.5% glycerol, and 25 μg of acetylated bovine

serum albumin/mL) for 15 min at 37 °C. NT-DNA (8450; **Table 2.S2**) was then added at 1.5 μM and incubated for another 15 min at 37 °C. Fully complementary NT-DNA was used to allow the energy of DNA reannealing upon termination and bubble collapse to contribute to the termination energetics.

2.5.4 *In vitro* transcription

ECs were diluted 5-fold in TB to 100 nM ECs, and radiolabeled by incorporation labeling with 0.1 μM [α-^{32}P]-CTP and 0.9 μM CTP for 5 min (for WT, ΔSI3, and Ala_6ΔSI3 ECs) or 30 min (for ΔTL, LTPPΔSI3, and LTPP ECs) to extend ECs to C18. Transcription was restarted by addition of diluted ECs to an equal volume of TB containing UTP and ATP at various concentrations (as indicated in figures), in the presence or absence of 50 μM asRNA (100 μM asRNA for Ala_6ΔSI3 ECs). All NTP mixes were supplemented with $Mg(OAc)_2$ at a concentration equivalent to the NTP concentration to prevent Mg^{2+} depletion effects. Timepoints longer than 8 s were stopped manually with an equal volume of 2x stop buffer (10 M urea, 50 mM EDTA, 90 mM Tris-borate buffer, pH 8.0, 0.02% bromophenol blue and 0.02% xylene cyanol). Timepoints shorter than 8 s were taken with rapid mixing quench-flow apparatus (RQF-3; KinTek Corporation, Austin, TX) by injecting diluted ECs in one sample loop, and UTP and ATP in TB in the presence or absence of 100 μM asRNAs in the other sample loop. Quench-flow reactions were quenched with 2 M HCl, and neutralized immediately to pH 7.8 by the addition of 3 M Tris base. RNA products were purified by phenol:chloroform extraction followed by ethanol precipitation, and resuspended in 1x stop buffer. RNA products from all timepoints were resolved by 8 M urea denaturing PAGE. Gels were exposed to phosphorimager screens, scanned using the Typhoon PhosphorImager and quantified using ImageQuant software (GE Healthcare).

For dissociation assays **(Figure 2.2D)**, reconstituted ECs were diluted 5-fold in TB and tethered to paramagnetic Co^{2+}-beads (Dynabeads from Thermo Fisher Scientific Biosciences

Inc.) through the His_{10}-tag on RNAP by incubating for 10 min at room temperature followed by 10 min at 37 °C. Bead-bound ECs were washed with TB to remove unbound RNAP and nucleic acids, and reactions restarted by UTP and ATP addition in the presence or absence of 50 μM asRNA. Supernatant timepoints were taken by magnetic partitioning, and supernatant and whole reaction timepoints were stopped with an equal volume of 2x stop buffer. RNA products were separated by 8 M urea denaturing PAGE, and visualized by phosphorimaging, as described above.

2.5.5 Kinetic fitting of transcription assay reaction progress plots

Reaction progress plots were constructed by determining the percent of each RNA species (*e.g.* C18, U19, etc.) for each timepoint, which represents the proportion of ECs positioned at each RNA position (*i.e.* percent occupancy) as a function of time. All RNAs longer than A20 were summed into the A20 fraction, since the kinetics at these later positions do not impact kinetics at U19.

The reaction progress plots were then fit to kinetic models (such as in **Figure 2.4A**) using KinTek Explorer (KinTek Corporation, Austin, TX), which uses numerical integration to fit the rate equations. All replicates were input in KinTek Explorer to obtain global fits and minimize technical errors. All replicates were also fit with the same models individually to obtain error estimates for the rate constants obtained by global fitting.

Reactions in the absence of asRNA were fit first to determine the elongation kinetics (and pausing, where applicable) at the termination site in the absence of termination. Data were first fit to the simplest model (below):

$$C18_e \rightarrow U19_e \rightarrow A20_e$$

where C18 is extended to U19 with a single elongation rate, and U19 is extended to A20 with a single rate. The simulations were seeded with initial rate constant guesses for the two reactions, obtained from single exponential fits for C18 disappearance and A20 appearance, respectively. All elongation kinetic models also required a slow dissociation step to adequately fit the model (see **Figure 2.4B**), likely due to the instability of the U-rich termination hybrid **(Figure 2.S2)**. If the simplest model provided a poor fit to the data (evidenced by high residuals with systematic error **(Figure 2.S2)**), more steps were added to the model to help describe the data, such as pausing steps for the C18 and/or U19 species (as in **Figure 2.4A** or **Figure 2.S4C**). In these cases, double exponential fits of the relevant reaction were used to seed the simulations, where the slower rates were used to seed the nucleotide addition rates from the paused complexes (*e.g.* $C18_p \rightarrow U19_e$ in **Figure 2.4A**).

In conditions with asRNA, elongation and pause kinetics were assumed to be unchanged, since these occur before asRNA binding (compare **Figure. 2.4B** and **Figure 2.5B**), and the relevant rate constants were thus constrained in these fits. Binding- (k_{on}) and off- (k_{off}) rates for asRNA, asRNA concentration, and time of asRNA addition were input into the models and also constrained. Since we observed some amount of terminator escape from C18 and U19, as well as some termination (evidenced by plateaus formed by C18 and U19 with a percent occupancy of >0%; see **Figure 2.5**), 4 reactions were added to these conditions: termination at C18, termination at U19, elongation of asRNA-bound C18, and asRNA-bound U19 (as in **Figure 2.5B, black arrows**). The only free parameters in these fits were thus the rate constants for these 4 reactions, obtained by fitting the relevant kinetic models to their datasets by global and individual fitting.

2.S1 Supplementary Methods

2.S1.1 Stopped-flow asRNA binding rate measurements

The fluorescent nucleic acid scaffold for asRNA binding assay was formed by mixing 25 μM pyrrolo-C (pC) labeled RNA and 50 μM T-DNA (8817 and 8451, respectively; **Table 2.S2**) in RB, annealing by heating to 95 °C and rapid cooling to 25 °C, as described above.

ECs were reconstituted by incubating 5 μM core *E. coli* RNAPs with 2.5 μM fluorescent nucleic acid scaffold in chloride transcription buffer (TB1.75; 20 mM Tris-Cl, pH 8.0, 75 mM NaCl, 1 mM $MgCl_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 2.5% glycerol, and 25 μg of acetylated bovine serum albumin/mL) for 15 min at 37 °C. NT-DNA (8450; **Table 2.S2**) was then added at 7.5 μM and incubated for another 15 min at 37 °C. ECs were diluted 5-fold in TB1.75 and extended to U19 by incubation with 100 μM UTP at 37 °C for 5 min for WT/ΔSI3 ECs, or with 1 mM UTP at 37 °C for 1 hr for ΔTL ECs.

ECs were then injected into one loading syringe of the stopped-flow apparatus (SF-300X; KinTek Corporation, Austin, TX), and asRNA (−8asRNA; 8713, **Table 2.S2**) in TB1.75 was loaded in the other syringe. Upon initiating rapid mixing at 37 °C, pC fluorescence was excited at 337 nm (2.4 nm bandwidth), and emission was monitored in real time through a 400 nm long-pass filter (Edmund Optics Inc., Barrington, NJ). The kinetics of pC fluorescence quenching were determined from n ≥3 replicates by fitting to a single exponential equation [1] and averaging rates (k_{obs}) obtained from all replicates:

$$F_t = F_0 + Ae^{-k_{obs}t} \quad [1]$$

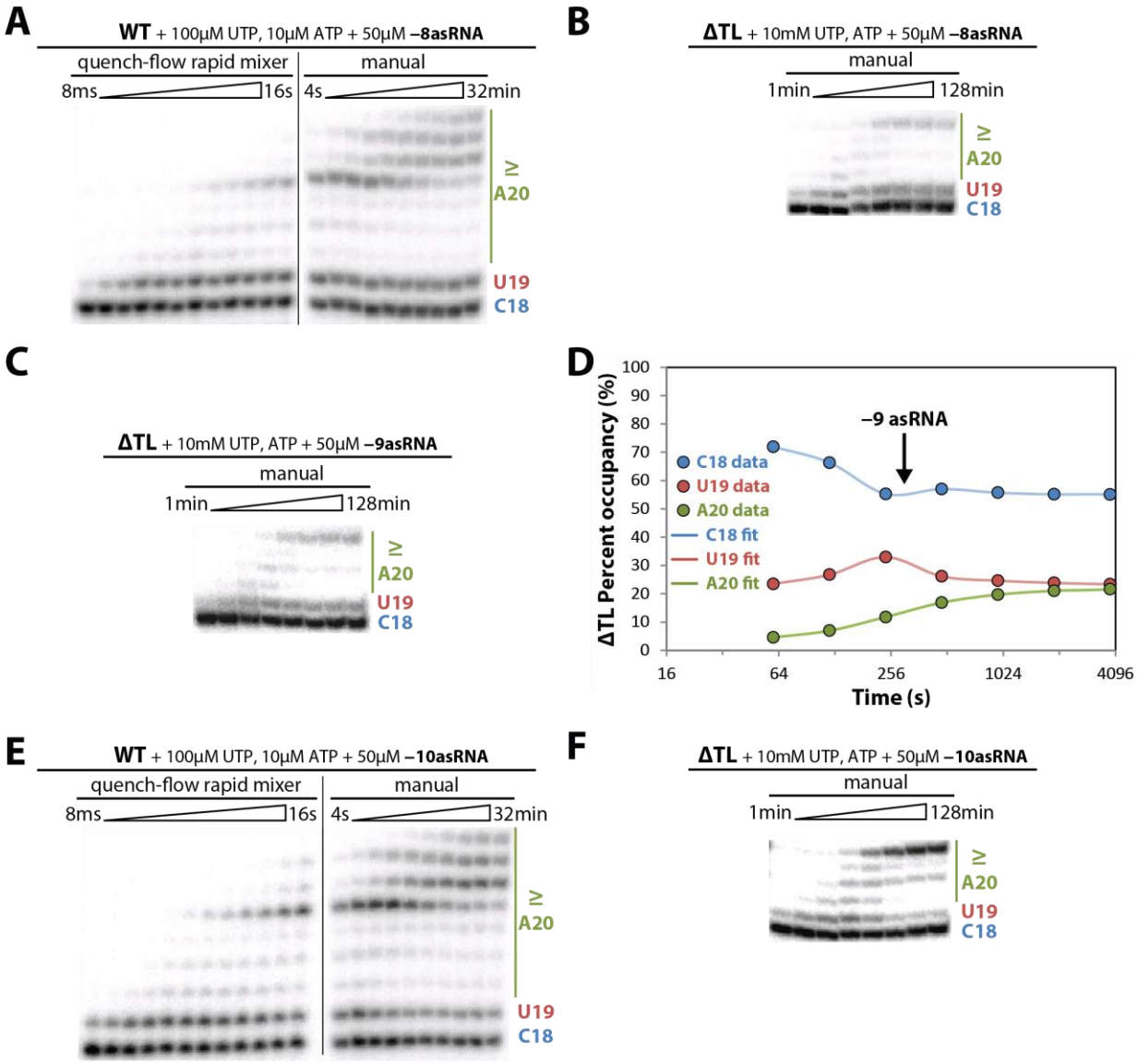
where t = time (s); F_o = initial pC fluorescence intensity; F_t = pC fluorescence intensity at time t; A = amplitude of the fluorescence change k_{obs} is the pseudo-first-order observed rate of the decrease in pC fluorescence.

To obtain a binding rate constant, k_{obs} was determined at various asRNA concentrations ranging from 1–16 μM. The resulting concentration dependence was fit to equation [2]:

$$k_{obs} = k_{on}[asRNA] + C \quad [2]$$

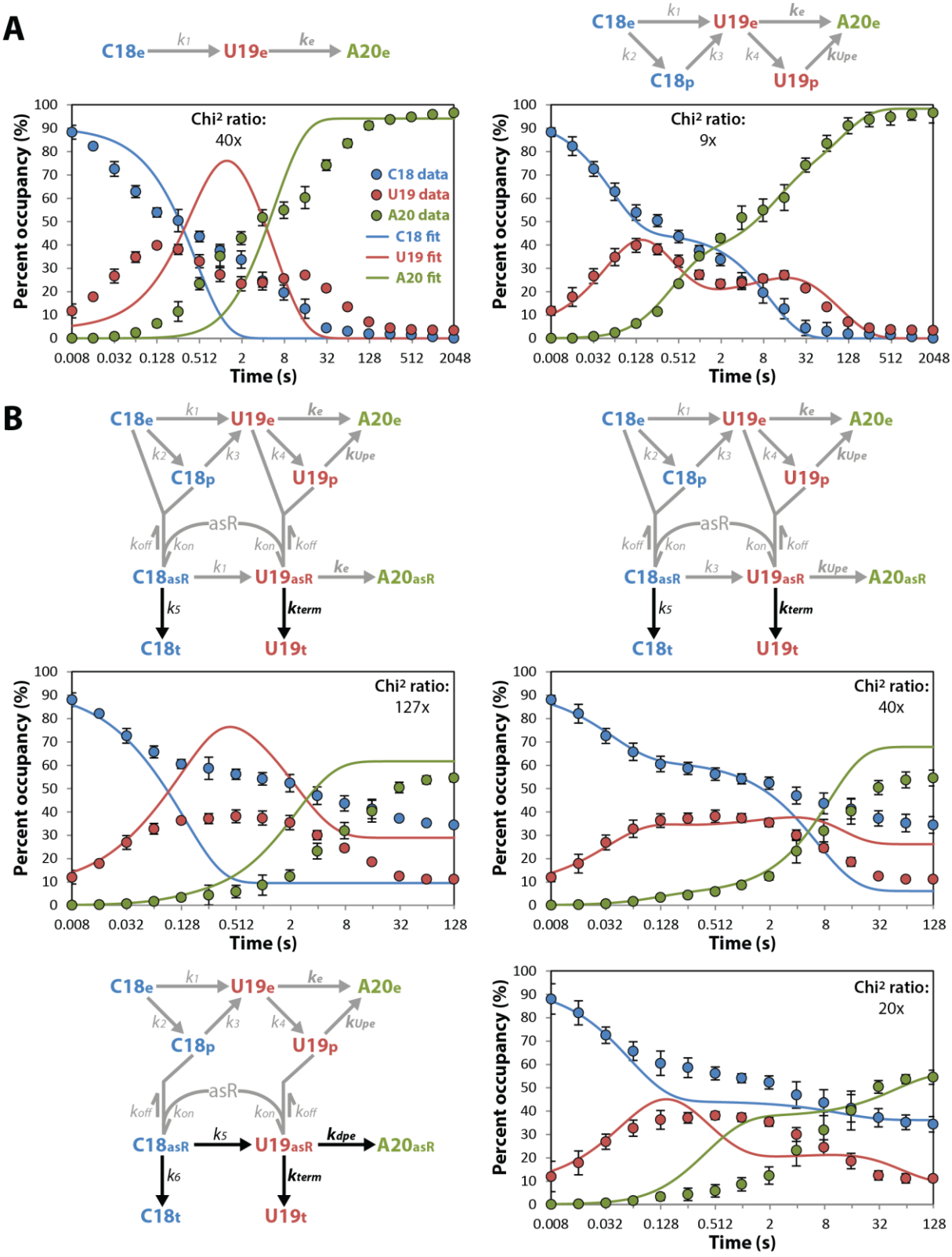
where k_{on} is the bimolecular binding rate constant, calculated from the slope of the line; C is a constant. Off-rate measurements (k_{off}) were obtained from previous work (37). Rates obtained for ΔTL ECs were also used for Ala_6ΔSI3, LTPPΔSI3, and LTPP ECs, since the elongation and termination kinetics of these RNAP variants are sufficiently slow that they are unaffected by large changes in the asRNA binding and dissociation rates.

2.S2 Supplementary Figures & Tables



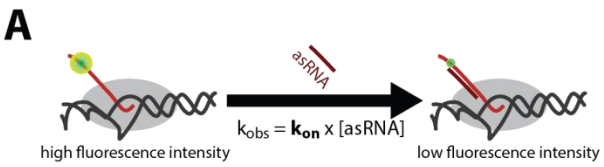
<u>Figure 2.S1:</u> –8 and –9asRNAs direct nearly 100% TE from ΔTL ECs, but –10asRNAs allow some terminator read-through.

(A) & **(B)** Representative gel panels showing the termination kinetics of WT and ΔTL ECs, respectively, with –8asRNA. **(C)** Representative gel panel, and **(D)** reaction progress plot for ΔTL ECs in the presence of –9asRNA.Time of asRNA addition is indicated with an arrow. Nearly 100% of ΔTL ECs at C18 & U19 at time of –9asRNA addition are terminated, as with –8asRNA **(Figure 2.3F)**. **(E)** & **(F)** Representative gel panels showing the termination kinetics of WT and ΔTL ECs, respectively, with –10asRNA.


A
C18e
k1
U19e
ke
A20e
C18p
k2
k3
k4
kUpe
U19p
Chi² ratio:
40x
Chi² ratio:
9x
Percent occupancy (%)
Time (s)
C18 data
U19 data
A20 data
C18 fit
U19 fit
A20 fit
B
asR
koff
kon
C18asR
U19asR
A20asR
k5
kterm
C18t
U19t
Chi² ratio:
127x
Chi² ratio:
40x
kdpe
k6
Chi² ratio:
20x

<u>Figure 2.S2:</u> Simpler kinetic models do not produce good fits to the data.

Simpler kinetic models with fewer free parameters and their corresponding best-fits to the data are shown for **(A)** WT ECs in the absence of asRNA, **(B)** WT ECs with –10asRNA, **(C)** ΔTL ECs in the absence of asRNA, and **(D)** ΔTL ECs with –10asRNA. Steps that are physiologically justifiable (based on known pause propensities of WT and mutant TL RNAPs) were added to the kinetic models as necessary until the fits exhibited a good fit to the data. Kinetic models for all other RNAP variants and reaction conditions were analogously determined by sequential adaptation of the model.
Chi^2 ratio = Chi^2(simple model) / Chi^2(final model). We note also that although the fits in panel **(A**, right**)**, and panel **(C)** appear to produce relatively good fits to the data, the addition of a slow dissociation step decreased the Chi^2 ratio significantly, but did not result in large changes in the other rates obtained from the fit.



B

	k_{on} ($\times 10^{-1}$ s^{-1})
WT	2.0 ± 0.05
ΔTL	1.0 ± 0.01
ΔSI3	1.3 ± 0.04

Figure 2.S3: asRNA binding rates measured directly.

(A) Schematic of fluorescence quenching assay used to determine apparent asRNA on-rates (k_{on}); assay designed and described previously (37) (see section **2.S1 Supplementary Methods**). Briefly, WT, ΔTL, and ΔSI3 ECs were reconstituted with pyrrolo-C-labeled RNA (8817; **Table 2.S2**) at position C18, then elongated with UTP to U19. U19 ECs were then mixed with varying concentrations of asRNA, and quench curves were obtained using KinTek Stopped-Flow apparatus. These data were then fit to single exponential equations, and subsequently used to determine k_{on}, summarized in **(B)**. These on-rates were used for asRNA kinetic fits. Errors represent SD from ≥3 independent experimental replicates.

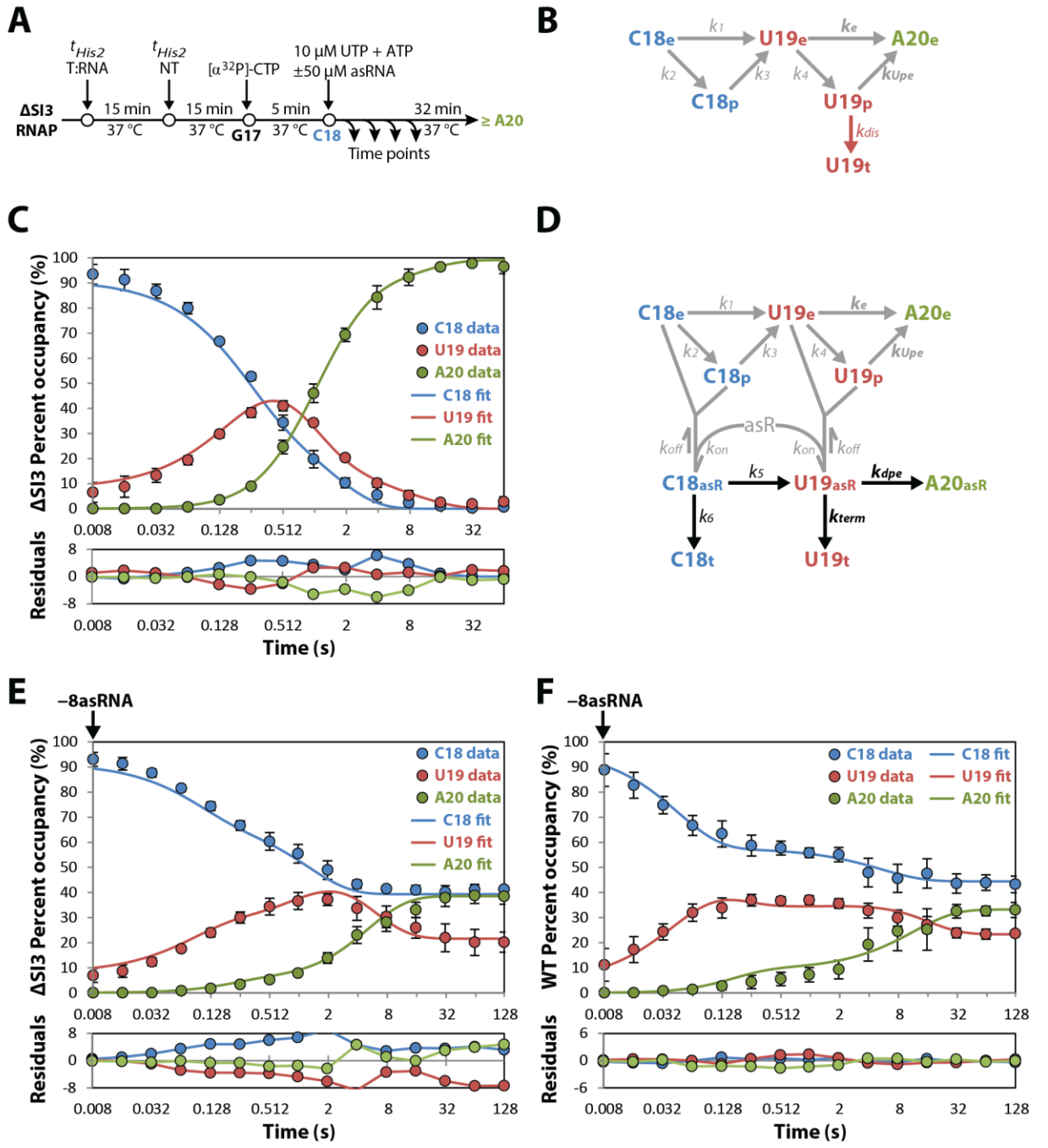
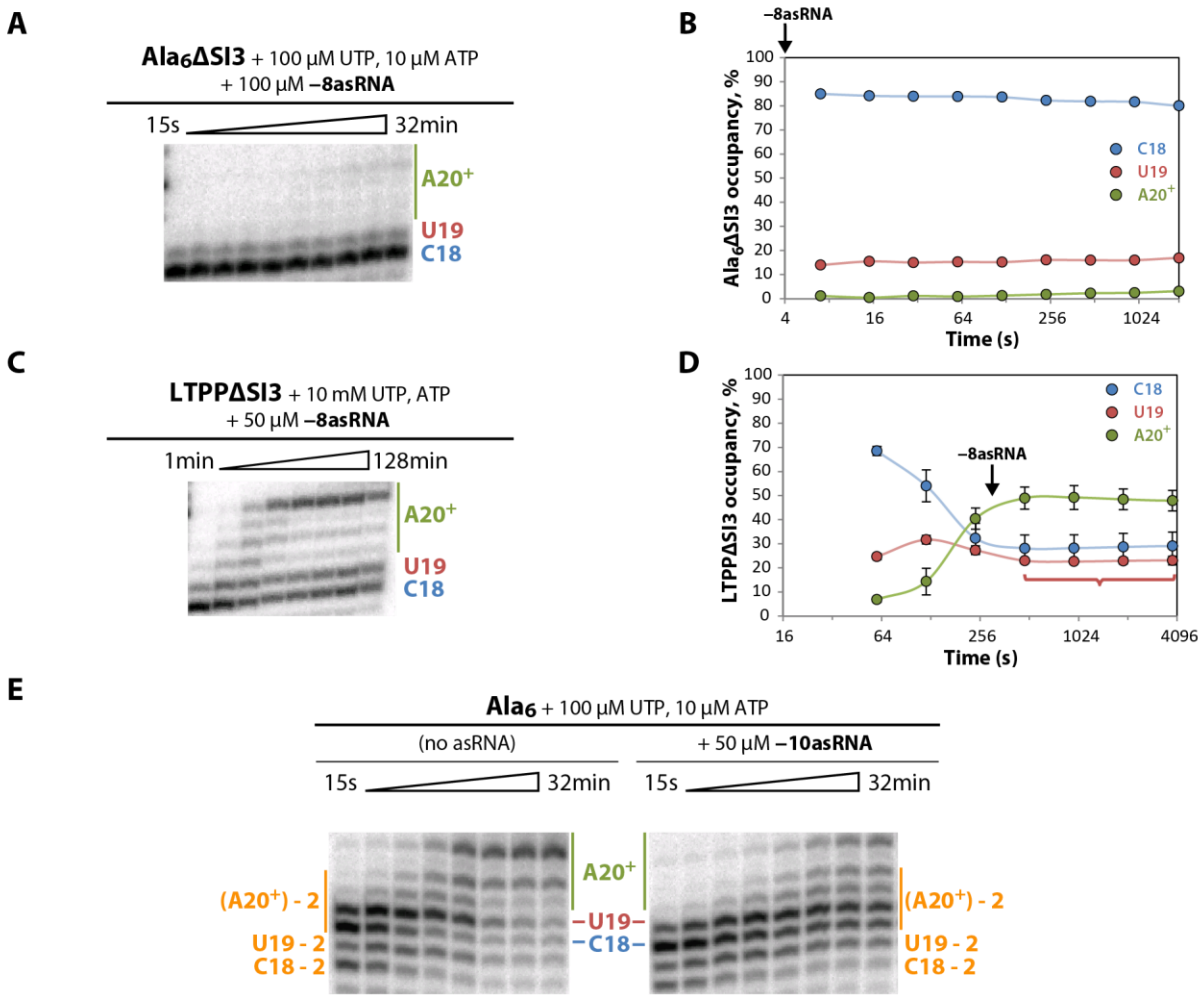


Figure 2.S4: TL sequence insertion 3 (SI3) has minimal contribution to termination rate.

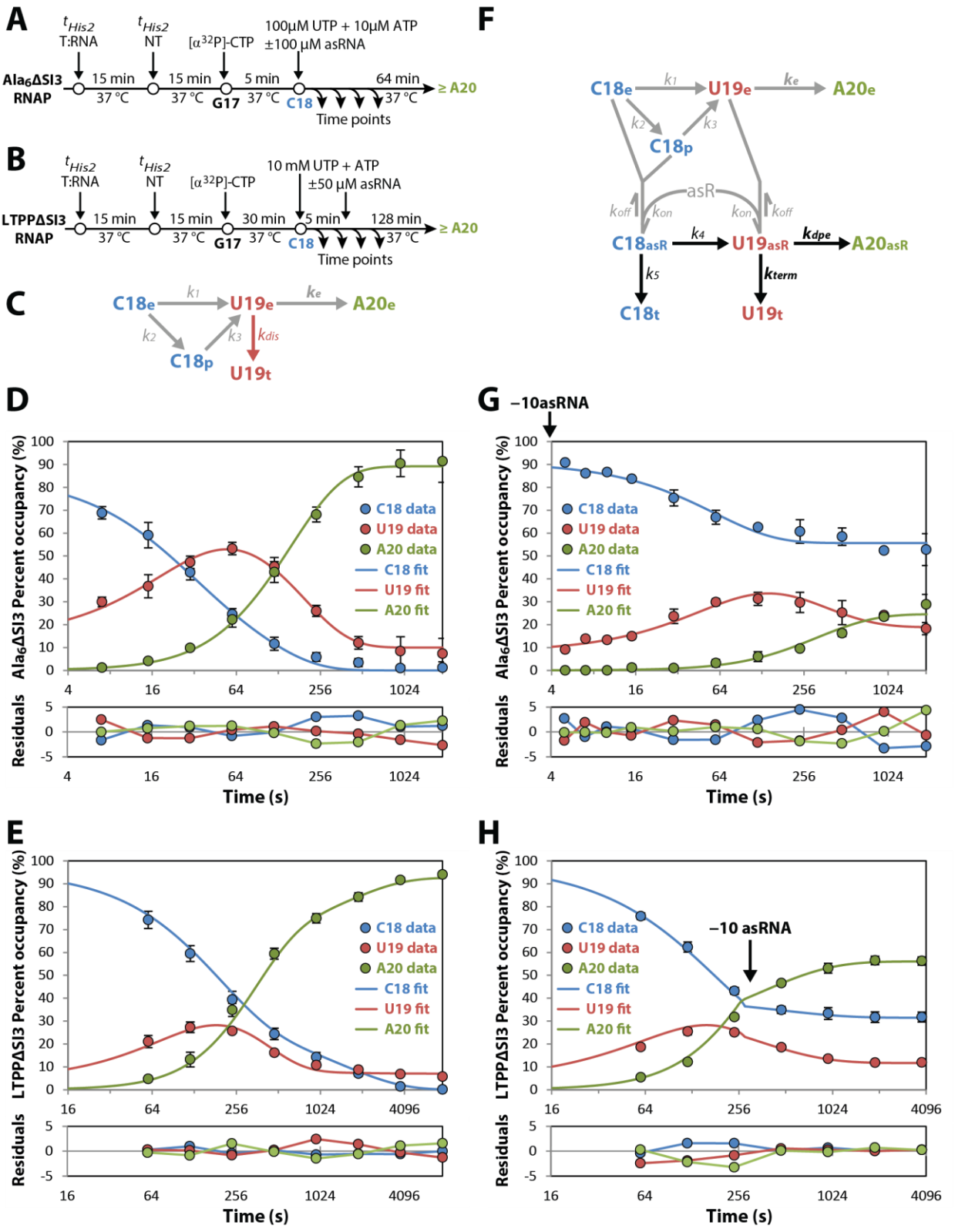
(A) Schematic representation of experimental set up used to measure elongation and termination rates for ΔSI3 ECs. **(B)** Kinetic model and **(C)** the corresponding fit for ΔSI3 ECs in the absence of asRNA. **(D)** Kinetic model and **(E)** the corresponding fit for ΔSI3 ECs with −8asRNA. **(F)** Kinetic fit for WT ECs with −8asRNA. WT reactions were carried out as shown in **Figure 2.3A**. Time of asRNA addition is indicated with black arrows. Residuals between the fits and data points are shown in the lower graphs. $C18_e$, $U19_e$, and $A20_e$ denote elongating ECs with C18, U19, or A20 RNAs, respectively; $C18_p$ and $U19_p$ denote U-tract paused ECs; asR denotes free asRNA; $C18_{asR}$, $U19_{asR}$, and $A20_{asR}$ denote asRNA-bound

complexes; $C18_t$ and $U19_t$ denote terminated complexes. Errors represent SD from ≥3 independent experimental replicates.



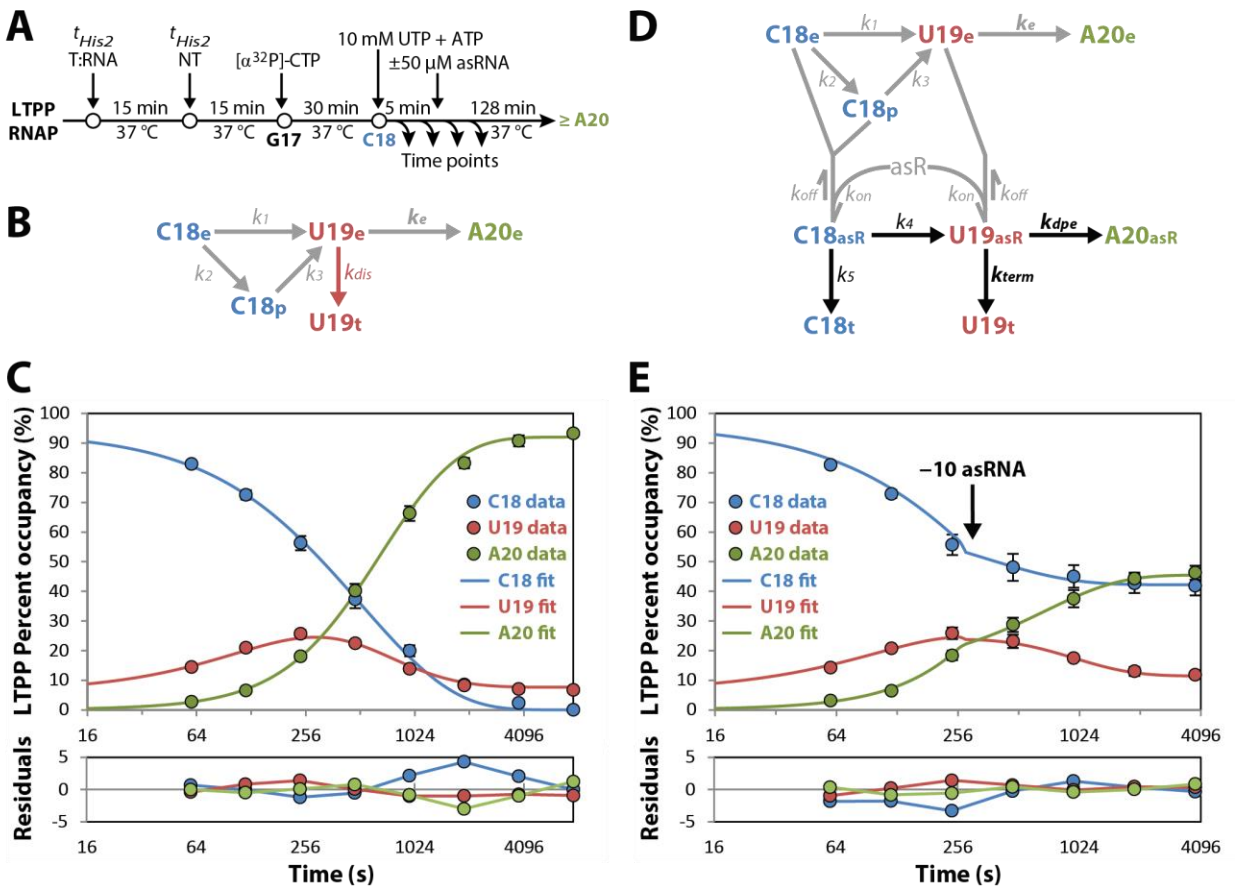
<u>Figure 2.S5:</u> Ala_6ΔSI3 ECs with –8asRNAs, LTPPΔSI3 ECs with –8asRNAs, and Ala_6 ECs are not suited for termination rate determination.

Representative gels and reaction progress curves for **(A-B)** Ala_6ΔSI3 ECs with –8asRNA and **(C-D)** LTPPΔSI3 ECs with –8asRNA. Times of asRNA addition are indicated with arrows. Nearly all ECs at U19 at the time of –8asRNA addition terminate, as evident from the immediate plateau in fraction U19 ECs (*indicated by red bracket for LTPPΔSI3*). C18, U19 and $A20^+$ occupancies for each timepoint are normalized to total radioactive counts in the lane. **(E)** Representative gel panels for Ala_6 ECs in the presence and absence of –10asRNA. Prominent RNA cleavage products are formed in both conditions that are then extended into the region of analysis, confounding elongation and termination rate determination. The lengths of the cleaved and extended bands are indicated with *orange* labels. Error bars represent SD from ≥3 independent experimental replicates.


A
t_{His2} T:RNA
t_{His2} NT
[α32P]-CTP
100µM UTP + 10µM ATP ±100 µM asRNA
Ala6ΔSI3 RNAP
15 min 37 °C
15 min 37 °C
5 min 37 °C
64 min 37 °C
G17
C18
≥ A20
Time points
B
t_{His2} T:RNA
t_{His2} NT
[α32P]-CTP
10 mM UTP + ATP ±50 µM asRNA
LTPPΔSI3 RNAP
15 min 37 °C
15 min 37 °C
30 min 37 °C
5 min
128 min 37 °C
G17
C18
≥ A20
Time points
C
C18e
U19e
A20e
C18p
U19t
k1
k2
k3
ke
kdis
F
C18e
U19e
A20e
C18p
asR
C18asR
U19asR
A20asR
C18t
U19t
k1
k2
k3
ke
koff
kon
kon
koff
k4
kdpe
k5
kterm
D
Ala6ΔSI3 Percent occupancy (%)
C18 data
U19 data
A20 data
C18 fit
U19 fit
A20 fit
Residuals
Time (s)
G
−10asRNA
Ala6ΔSI3 Percent occupancy (%)
C18 data
U19 data
A20 data
C18 fit
U19 fit
A20 fit
Residuals
Time (s)
E
LTPPΔSI3 Percent occupancy (%)
C18 data
U19 data
A20 data
C18 fit
U19 fit
A20 fit
Residuals
Time (s)
H
LTPPΔSI3 Percent occupancy (%)
−10 asRNA
C18 data
U19 data
A20 data
C18 fit
U19 fit
A20 fit
Residuals
Time (s)

<u>Figure 2.S6:</u> Conformationally biased TL mutants impair termination rate.

(A) & **(B)** Schematic representation of experimental set up used to measure elongation and termination rates for Ala_6ΔSI3 and LTPPΔSI3 ECs, respectively. **(C)** Kinetic model and the corresponding fits for **(D)** Ala_6ΔSI3 and **(E)** LTPPΔSI3 ECs in the absence of asRNA. **(F)** Kinetic model and the corresponding fits for **(G)** Ala_6ΔSI3 and **(H)** LTPPΔSI3 ECs with –10asRNA. Time of asRNA addition is indicated with black arrows, and is accounted for in the kinetic fits. Residuals between the fits and data points are shown in the lower graphs. $C18_e$, $U19_e$, and $A20_e$ denote elongating ECs with C18, U19, or A20 RNAs, respectively; $C18_p$ and $U19_p$ denote U-tract paused ECs; asR denotes free asRNA; $C18_{asR}$, $U19_{asR}$, and $A20_{asR}$ denote asRNA-bound complexes; $C18_t$ and $U19_t$ denote terminated complexes. Errors represent SD from ≥3 independent experimental replicates.



<u>Figure 2.S7:</u> The presence of SI3 further inhibits termination of unfolded TL mutants.

(A) Schematic representation of experimental set up used to measure LTPP elongation and termination rates. **(B)** Kinetic model and **(C)** the corresponding fit for LTPP ECs in the absence of asRNA. **(D)** Kinetic model and **(E)** the corresponding fit for LTPP ECs with −10asRNA. Time of asRNA addition is indicated with a black arrow, and is accounted for in the kinetic fits. Residuals between the fits and data points are shown in the lower graphs. $C18_e$, $U19_e$, and $A20_e$ denote elongating ECs with C18, U19, or A20 RNAs, respectively; $C18_p$ and $U19_p$ denote U-tract paused ECs; asR denotes free asRNA; $C18_{asR}$, $U19_{asR}$, and $A20_{asR}$ denote asRNA-bound complexes; $C18_t$ and $U19_t$ denote terminated complexes. Errors represent SD from ≥3 independent experimental replicates.

Table 2.S1: RNAP variants and overexpression plasmids

Plasmid #	Plasmid Name	RNAP TL mutant	Description	Source
2956	pRM756	WT	Expresses wild-type *E. coli* RNAP (α2ββ′ω) with His_{10} tag on β′ C-terminus	(34)
2995	pRM795	WT	Expresses wild-type *E. coli* RNAP (α2ββ′ω) with His_{10} and HMK tags on β′ C-terminus	(38)
2978	pRM778	ΔTL	pRM756 derivative expressing mutant *E. coli* RNAP (α2ββ′ω) containing β' Δ931-1137 with SNAP-tag on β′ C-terminus	this work & (34)
2959	pRM759	ΔSI3	pRM756 derivative expressing mutant *E. coli* RNAP (α2ββ′ω) containing β' Δ943-1130	(34)
2996	pRM796	Ala_6ΔSI3	pRM795 derivative expressing mutant *E. coli* RNAP (α2ββ′ω) containing β' Ala_6 (924A, 926A, 927A, 11336A, 1137A, 1139A), and ΔSI3 (Δ945-1132)	(34)
2964	pRM780	LTPPΔSI3	pRM756 derivative expressing mutant E. coli RNAP (α2ββ′ω) containing β' L930P, T931P, and ΔSI3 (Δ945-1130) with SNAP-tag on β′ C-terminus	this work
2961	pRM761	LTPP	pRM756 derivative expressing mutant E. coli RNAP (α2ββ′ω) containing β' L930P, T931P	(34)

Table 2.S2: Oligonucleotides used in this study.

Oligo #	RNA / DNA	Sequence (5' to 3')	Use
8450	DNA	GGTCAGTACGTCCTGCTTTGTGCTAAAAGAGATTCAGAGTCTTCCAGTGGTGCATGAACG	NT-strand for t_{his2} reconstitution
8451	DNA	CGTTCATGCACCACTGGAAGACTCTGAATCTCTTTTAGCACAAAGCAGGACGTACTGACC	T-strand for t_{his2} reconstitution
10002	RNA	UCAUCCGGCGCUUUGUG	G17 RNA strand for t_{his2} reconstitution
8713	RNA	AGCGCCGGAUGA	−8asRNA
10428	RNA	CGCCGGAUGA	−10asRNA
8703	DNA	AGCGCCGGATGA	asDNA **(Figure 2.2D)**
8812	RNA	UACCAUACUGUU	nasRNA **(Figure 2.2D)**
8811	DNA	TACCATACTGTT	nasDNA **(Figure 2.2D)**
8714	RNA	GCGCCGGAUGA	−9asRNA **(Figure 2.S1)**
8817	RNA	UCAU**X**CGGCGCUUUGUGC	pyrrolo-C (**X**) labeled RNA for reconstituting scaffolds for asRNA binding rate experiments in **Figure 2.S3**

2.7 REFERENCES

1. Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S, & Nudler E (2007) An allosteric path to transcription termination. *Molecular cell* 28(6):991-1001.

2. Epshtein V, Dutta D, Wade J, & Nudler E (2010) An allosteric mechanism of Rho-dependent transcription termination. *Nature* 463(7278):245-249.

3. Wilson KS & von Hippel PH (1994) Stability of Escherichia coli transcription complexes near an intrinsic terminator. *J Mol Biol* 244(1):36-51.

4. Yager TD & von Hippel PH (1991) A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. *Biochemistry* 30(4):1097-1118.

5. Wilson KS & von Hippel PH (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A* 92(19):8793-8797.

6. Komissarova N, Becker J, Solter S, Kireeva M, & Kashlev M (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Molecular cell* 10(5):1151-1162.

7. Korzheva N*, et al.* (2000) A structural model of transcription elongation. *Science* 289(5479):619-625.

8. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, & Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448(7150):157-162.

9. Nudler E, Avetissova E, Markovtsov V, & Goldfarb A (1996) Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science* 273(5272):211-217.

10. Wilson KS, Conant CR, & von Hippel PH (1999) Determinants of the stability of transcription elongation complexes: interactions of the nascent RNA with the DNA template and the RNA polymerase. *J Mol Biol* 289(5):1179-1194.

11. Korzheva N, Mustaev A, Nudler E, Nikiforov V, & Goldfarb A (1998) Mechanistic model of the elongation complex of Escherichia coli RNA polymerase. *Cold Spring Harb Symp Quant Biol* 63:337-345.

12. Sidorenkov I, Komissarova N, & Kashlev M (1998) Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Molecular cell* 2(1):55-64.

13. Martin FH & Tinoco I, Jr. (1980) DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res* 8(10):2295-2299.

14. Gusarov I & Nudler E (1999) The mechanism of intrinsic transcription termination. *Molecular cell* 3(4):495-504.

15. Yarnell WS & Roberts JW (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* 284(5414):611-615.

16. Lubkowska L, Maharjan AS, & Komissarova N (2011) RNA folding in transcription elongation complex: implication for transcription termination. *J Biol Chem* 286(36):31576-31585.

17. Larson MH, Greenleaf WJ, Landick R, & Block SM (2008) Applied force reveals mechanistic and energetic details of transcription termination. *Cell* 132(6):971-982.

18. Santangelo TJ & Roberts JW (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Molecular cell* 14(1):117-126.

19. Peters JM, Vangeloff AD, & Landick R (2011) Bacterial transcription terminators: the RNA 3'-end chronicles. *J Mol Biol* 412(5):793-813.

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

21. Ray-Soni A, Bellecourt MJ, & Landick R (2016) Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual review of biochemistry* 85:319-347.

22. Toulokhonov I & 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(5):1125-1136.

23. Reynolds R, Bermudez-Cruz RM, & Chamberlin MJ (1992) Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators. *J Mol Biol* 224(1):31-51.

24. Chakraborty A*, et al.* (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science* 337(6094):591-595.

25. Gnatt AL, Cramer P, Fu J, Bushnell DA, & Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* 292(5523):1876-1882.

26. Weixlbaumer A, Leon K, Landick R, & Darst SA (2013) Structural basis of transcriptional pausing in bacteria. *Cell* 152(3):431-441.

27. Sekine S, Murayama Y, Svetlov V, Nudler E, & Yokoyama S (2015) Ratcheting of RNA polymerase toward structural principles of RNA polymerase operations. *Transcription* 6(3):56-60.

28. Rutherford ST, Villers CL, Lee JH, Ross W, & Gourse RL (2009) Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev* 23(2):236-248.

29. Vassylyev DG*, et al.* (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448(7150):163-168.

30. Vassylyev DG*, et al.* (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* 417(6890):712-719.

31. Wang D, Bushnell DA, Westover KD, Kaplan CD, & Kornberg RD (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127(5):941-954.

32. Kashlev M & Komissarova N (2002) Transcription termination: primary intermediates and secondary adducts. *J Biol Chem* 277(17):14501-14508.

33. Zhang J, Palangat M, & Landick R (2010) Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat Struct Mol Biol* 17(1):99-104.

34. Windgassen TA, *et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

35. Mejia YX, Nudler E, & Bustamante C (2015) Trigger loop folding determines transcription rate of Escherichia coli's RNA polymerase. *Proc Natl Acad Sci U S A* 112(3):743-748.

36. Toulokhonov I, Zhang J, Palangat M, & Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Molecular cell* 27(3):406-419.

37. Hein PP, *et al.* (2014) RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nat Struct Mol Biol* 21(9):794-802.

38. Nayak D, Voss M, Windgassen T, Mooney RA, & Landick R (2013) Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Molecular cell* 50(6):882-893.

39. Lane WJ & Darst SA (2010) Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J Mol Biol* 395(4):671-685.

40. Iyer LM, Koonin EV, & Aravind L (2004) Evolution of bacterial RNA polymerase: implications for large-scale bacterial phylogeny, domain accretion, and horizontal gene transfer. *Gene* 335:73-88.

41. Zakharova N, Bass I, Arsenieva E, Nikiforov V, & Severinov K (1998) Mutations in and monoclonal antibody binding to evolutionary hypervariable region of Escherichia coli RNA polymerase beta' subunit inhibit transcript cleavage and transcript elongation. *J Biol Chem* 273(38):24912-24920.

42. Artsimovitch I, Svetlov V, Murakami KS, & 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(14):12344-12355.

43. Chan CL, Wang D, & Landick R (1997) Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *J Mol Biol* 268(1):54-68.

44. Bar-Nahum G, *et al.* (2005) A ratchet mechanism of transcription elongation and its control. *Cell* 120(2):183-193.

45. Darst SA, *et al.* (2002) Conformational flexibility of bacterial RNA polymerase. *Proc Natl Acad Sci U S A* 99(7):4296-4301.

46. Ansari A, *et al.* (1985) Protein states and proteinquakes. *Proc Natl Acad Sci U S A* 82(15):5000-5004.

47. Frauenfelder H, Sligar SG, & Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254(5038):1598-1603.

48. Liu SQX, Y.H.; Ji, X.L.; Tao, Y.; Tan, D.Y.; Zhang, K.Q.; Fu, Y.X. (2012) Protein Folding, Binding and Energy Landscape: A Synthesis. *Protein Engineering*, ed Kaumaya PTPRijeka, Croatia), pp 207–252.

49. Henzler-Wildman K & Kern D (2007) Dynamic personalities of proteins. *Nature* 450(7172):964-972.

50. Noe F, Schutte C, Vanden-Eijnden E, Reich L, & Weikl TR (2009) Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc Natl Acad Sci U S A* 106(45):19011-19016.

51. Berlin V & Yanofsky C (1983) Release of transcript and template during transcription termination at the trp operon attenuator. *J Biol Chem* 258(3):1714-1719.

52. Weilbaecher R, Hebron C, Feng G, & Landick R (1994) Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev* 8(23):2913-2927.

53. Datta K & von Hippel PH (2008) Direct spectroscopic study of reconstituted transcription complexes reveals that intrinsic termination is driven primarily by thermodynamic destabilization of the nucleic acid framework. *J Biol Chem* 283(6):3537-3549.

54. Zamudio JR, Kelly TJ, & Sharp PA (2014) Argonaute-bound small RNAs from promoter-proximal RNA polymerase II. *Cell* 156(5):920-934.

55. Xie M, *et al.* (2013) Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* 155(7):1568-1580.

56. Harwig A, Jongejan A, van Kampen AH, Berkhout B, & Das AT (2016) Tat-dependent production of an HIV-1 TAR-encoded miRNA-like small RNA. *Nucleic Acids Res* 44(9):4340-4353.

57. Rijal K & Maraia RJ (2016) Active Center Control of Termination by RNA Polymerase III and tRNA Gene Transcription Levels In Vivo. *PLoS Genet* 12(8):e1006253.

58. Kyzer S, Ha KS, Landick R, & Palangat M (2007) Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J Biol Chem* 282(26):19020-19028.

Chapter 3:

Investigating the mechanism of intrinsic termination enhancement by NusA

Chapter contributions:

A.R.S. performed all experiments (except for Figure 3.4), experimental design and analysis, and prepared all text/figures; Rachel A. Mooney performed experiment in Figure 3.4 and copyedited the text.

3.1 INTRODUCTION

Intrinsic transcription termination is one of the two major pathways for termination in bacteria that is encoded in the genome. It is caused by the formation of an RNA stem-loop (or, terminator hairpin, T_{hp}) structure in the enzyme, which through a series of steps results in dissociation of the elongating transcription complex (EC; see **Chapter 1** for a detailed description of the mechanism). The majority of evidence regarding the mechanism of intrinsic termination was obtained using purified *in vitro* systems, that lack cellular transcription factors. This strategy allowed careful investigation of the termination mechanism of the basal transcription machinery, without the potentially complicating effects of dissociable transcription factors. However, as our knowledge of the basal termination mechanism advances, it is prudent to determine the roles of some the transcription factors that may regulate termination *in vivo*.

Of particular interest is the universal bacterial and archaeal transcription factor NusA, essential in *Escherichia coli* (1). NusA is thought to be associated with all active ECs that have released the initiation factor sigma (2), and thus is an important EC modifier to consider in the termination mechanism. Importantly, NusA modulates multiple transcriptional activities including elongation, pausing, and both intrinsic and Rho-dependent termination (3-18). Additionally, when complexed with other transcription factors, NusA can also act as an anti-termination factor that inhibits intrinsic termination (19).

E. coli NusA consists of six domains **(Figure 3.1)**: the N-terminal domain (NTD), three consecutive RNA-binding domains (S1, KH1, and KH2), and two acidic repeat domains (AR1 and AR2) (20-22). Whereas the NTD, S1, KH1, and KH2 domains are highly conserved, the AR domains are present in most γ-proteobacteria, but absent in archaea and most other bacterial lineages; these AR domains are dispensable for *E. coli* growth (23). NusA binds RNAP near the RNA exit channel through contacts between *(i)* NusA-NTD and the tip of the RNAP flap module

(RNAP-FT) (16), and *(ii)* NusA-AR1 and NusA-AR2 and the C-terminal domain of one of the two RNAP alpha subunits (RNAP-αCTD), where NusA-AR2 forms the tighter interaction with RNAP-αCTD (24-27). The AR domains of *E. coli* NusA are auto-inhibitory, and binding to the RNAP-αCTD unmasks NusA activity (21, 24).

NusA activity is able to aid hairpin-stabilized pausing by promoting hairpin formation. Docking of NusA at the opening of the RNA exit channel perfectly positions it for stimulating RNA hairpin formation in the exit channel, and stabilizes the interaction between the RNAP-FT and RNA hairpin that stimulates hairpin-stabilized pausing (28). Consistently, NusA has been shown to interact with an RNA hairpin in the exit channel (16, 18, 29-31), and aid intermolecular RNA duplex formation in the exit channel *in vitro* in an RNAP-FT-dependent manner (32). NusA can also promote intrinsic termination at terminators with weak T_{hp}s or interruptions in the U-tract (4-6, 16, 18, 28, 33-35). In fact, some *Bacillus subtilis* intrinsic terminators with remarkably weak T_{hp}s were identified as NusA-dependent *in vivo* (35).

An alternate model suggested that NusA aids hairpin-stabilized pausing and intrinsic termination by displacing upstream ssRNA from a binding site on RNAP, making it available for RNA hairpin formation in the exit channel (6). However, NusA effects on pausing and termination can be mediated through intermolecular RNA duplexes lacking upstream ssRNA (16, 36), indicating that *(i)* NusA activity does not rely on the displacement of ssRNA from a binding site on RNAP, and that *(ii)* NusA is able to act as effectively with A-form RNA duplexes as a binding substrate. These findings thus support the model in which the primary activity of NusA in hairpin-stabilized pausing and termination is to stimulate formation of the RNA hairpin in the RNA exit channel.

Interestingly, single-molecule optical trapping assays revealed that NusA also decreases the rate of pause-free elongation, ostensibly by increasing the energy barrier to forward translocation of ECs, and increases the probability of entry into an elemental paused state

without affecting pause duration (37). Consistently, stimulation of elemental pause half-life at a termination site was estimated to contribute less than 20% of the NusA effect on termination efficiency (TE) (6). However, it remains to be determined whether NusA affects entry into the elemental paused state, or stabilization of a hairpin-stabilized pause intermediate on the termination pathway.

Notably, NusA-NTD is necessary and sufficient to recapitulate the effects of full length NusA (FL-NusA) on both hairpin-stabilized pausing and RNA duplex formation *in vitro* (16, 32). However, NusA-NTD was not able to recapitulate the full effect of FL-NusA on termination enhancement *in vitro* (16), suggesting a possible role for NusA in termination in addition to stimulation of T_{hp} formation.

A-tracts upstream of the T_{hp} **(Figure 1.1)** complementary to the terminator U-tracts have been shown to increase TE *in vivo* (35, 38). However, A-tracts did not have an effect on TE with purified core RNAP (*i.e.* lacking initiation factor sigma) *in vitro* (33), indicating a requirement of cellular transcription factors to mediate the A-tract effect. Because pairing of the A- and U-tracts must occur after the complete formation and extension of the T_{hp} (see discussion in section **1.4.4c Reconciling observed effects of A-tracts on termination efficiency**, and **Figure 1.3**), NusA poses an interesting candidate for the mediation of A-tract effects due to its putative role in stabilizing the interactions between RNAP-FT and the T_{hp}. Stabilization of the RNAP-FT–T_{hp} interaction may slow EC dissociation and allow reversal of T_{hp} extension, thought to be irreversible *in vitro*, thus enabling the pairing of A- and U-tracts to favor EC dissociation in competition with T_{hp} melting (*i.e.* the reverse of T_{hp} extension).

Given the detailed mechanistic nature of key remaining questions, we reasoned that the ability to isolate the rates of the steps along the intrinsic termination pathway would be instrumental in addressing some of these ideas. The termination assay I developed (see **Chapter 2**) could allow us to determine the effects of NusA at each of these steps, and the

ability to titrate the concentration of antisense RNAs (asRNAs; complementary to the RNA in the exit channel of RNAP) that form the upstream arm of the T_{hp} would enable us to identify effects of NusA that are independent of effects on hairpin formation. This assay thus presented an attractive method to adapt for the study of NusA effects on intrinsic termination.

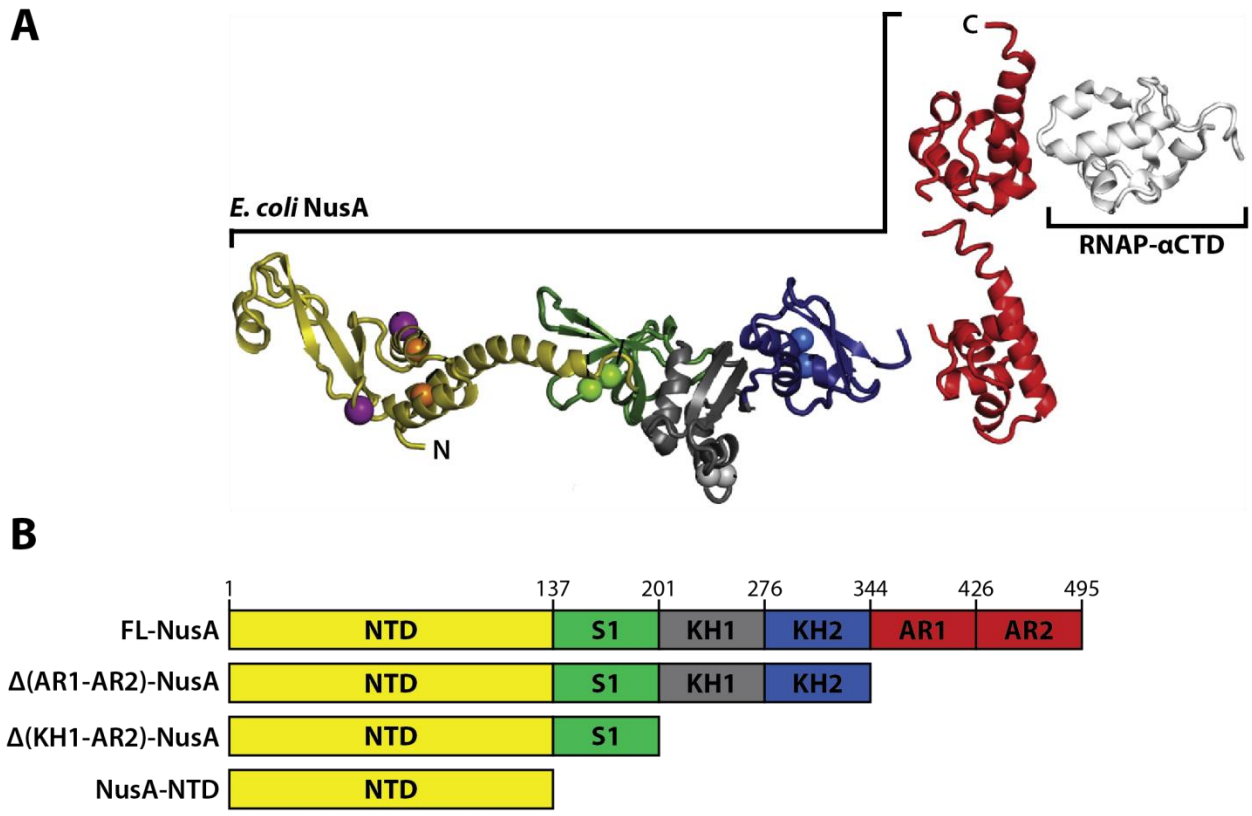


Figure 3.1: NusA domain structure and truncation mutants used in this study

(adapted from Ha *et al.* (2010) JMB. 401:708-725.)

(A) Structural model of NusA developed by Ha *et al.*, with domains colored as in **(B)**. The NTD, S1, KH1, and KH2 domains were modeled using the crystal structure of *Thermotoga maritima* NusA (PDB 1l2f). NMR structures of *E. coli* NusA were used for the AR1and AR2 domains (PDB 1wcl), and their interactions with the RNAP α subunit C-terminal domain (RNAP- αCTD; PDB 2jzb). **(B)** NusA truncation mutants used in this work. FL-NusA, full-length NusA; NusA-NTD, NusA N-terminal domain.

3.2 Results & Discussion

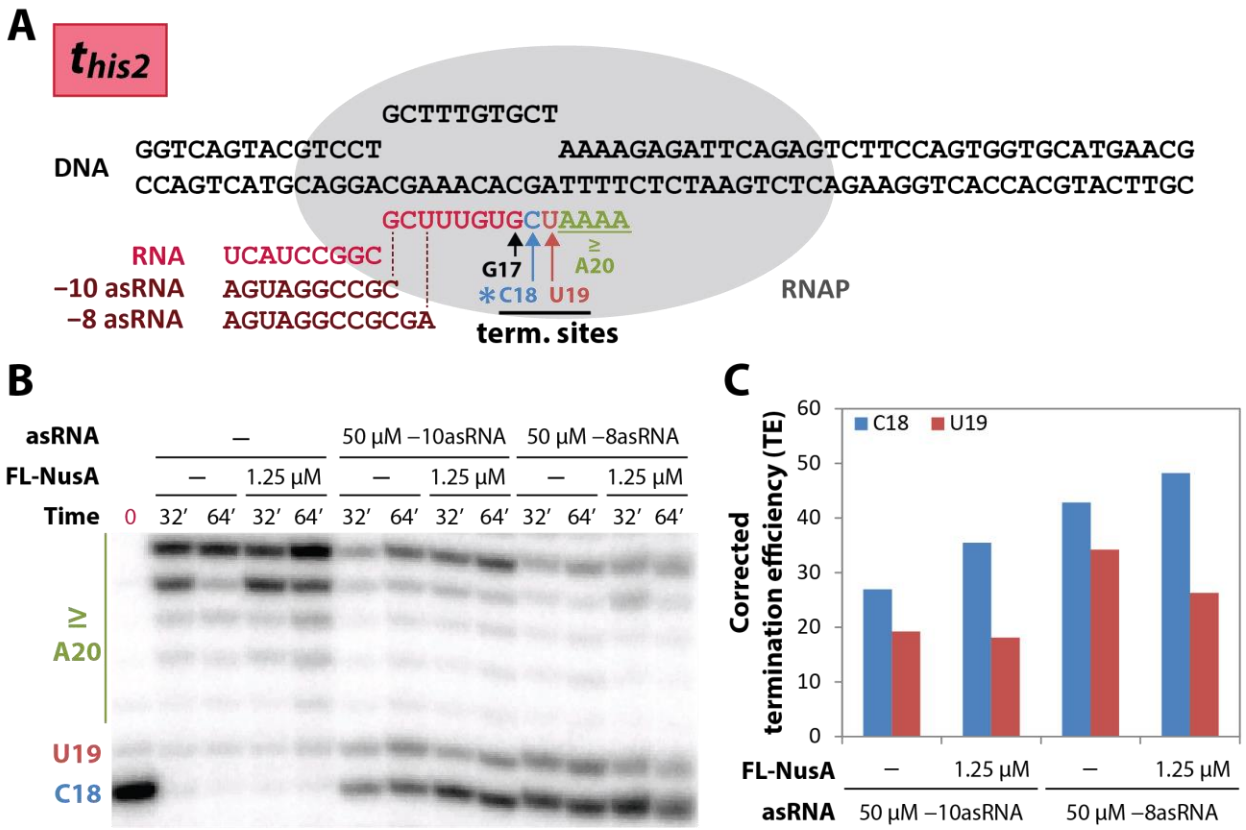
3.2.1 FL-NusA does not increase TE from termination scaffold t_{his2} with saturating asRNA

The ability of NusA to aid intrinsic termination in response to intermolecular RNA duplexes at low asRNA concentrations was shown previously (36), consistent with its role in stimulating hairpin formation in the RNA exit channel (32, 34). Because NusA-NTD is sufficient for stimulating hairpin formation, but is unable to recapitulate the full effect of FL-NusA on intrinsic termination, we wanted to ask whether NusA can aid termination when the asRNA is saturating.

To test whether FL-NusA could enhance TE at saturating asRNA concentrations on the termination scaffold used in our assay (t_{his2}; **Figure 3.2A**, see also **Chapter 2**), ECs were assembled with WT RNAP at position G17, and the RNAs were then radiolabeled by addition of [α-^{32}P]-CTP. The ECs were then incubated with NusA or an equivalent volume of NusA buffer, and elongated through the termination position U19 by incubation with UTP and ATP, in the absence or presence of –8asRNA (which allows T_{hp} extension; see sections **1.4.3 Terminator Hairpin Extension**, and **2.3.1 Measuring termination from elongation-compromised RNAP mutants**) or –10asRNA (which allows only limited T_{hp} extension). Both asRNA alternatives were used to determine whether any effect of NusA is dependent on T_{h}p extension. The effect of NusA on TE was minimal under these conditions, further complicated by apparently opposing effects at the two termination sites, C18 and U19 **(Figure 3.2B-C)**.

We reasoned that there were two possibilities for why NusA did not have a strong effect under these conditions: *(i)* the main role of NusA in termination is the stimulation of T_{hp} formation, and therefore NusA does not exhibit a large effect on TE when asRNA is saturating and NusA cannot increase the rate of RNA duplex formation in the exit channel, or *(ii)* the RNA

emerging from the RNA exit channel is too short in this scaffold to provide a good binding substrate for NusA. The t_{his2} RNA duplex only extends to –19 when the EC is positioned at U19. However, it was shown previously that the effect of NusA on pausing is enhanced when the RNA duplex is extended from –19 (with respect to the RNA 3′ end in the active site as –1) to –20 or –21 (34), consistent with a need for a longer RNA substrate. To eliminate the possibility that the short RNA substrate in the t_{his2} ECs was the reason for the lack of NusA activity on termination, we next tested the activity of NusA with an extended RNA duplex.



<u>Figure 3.2:</u> NusA does not enhance termination efficiency (TE) from the t_{his2} terminator scaffold

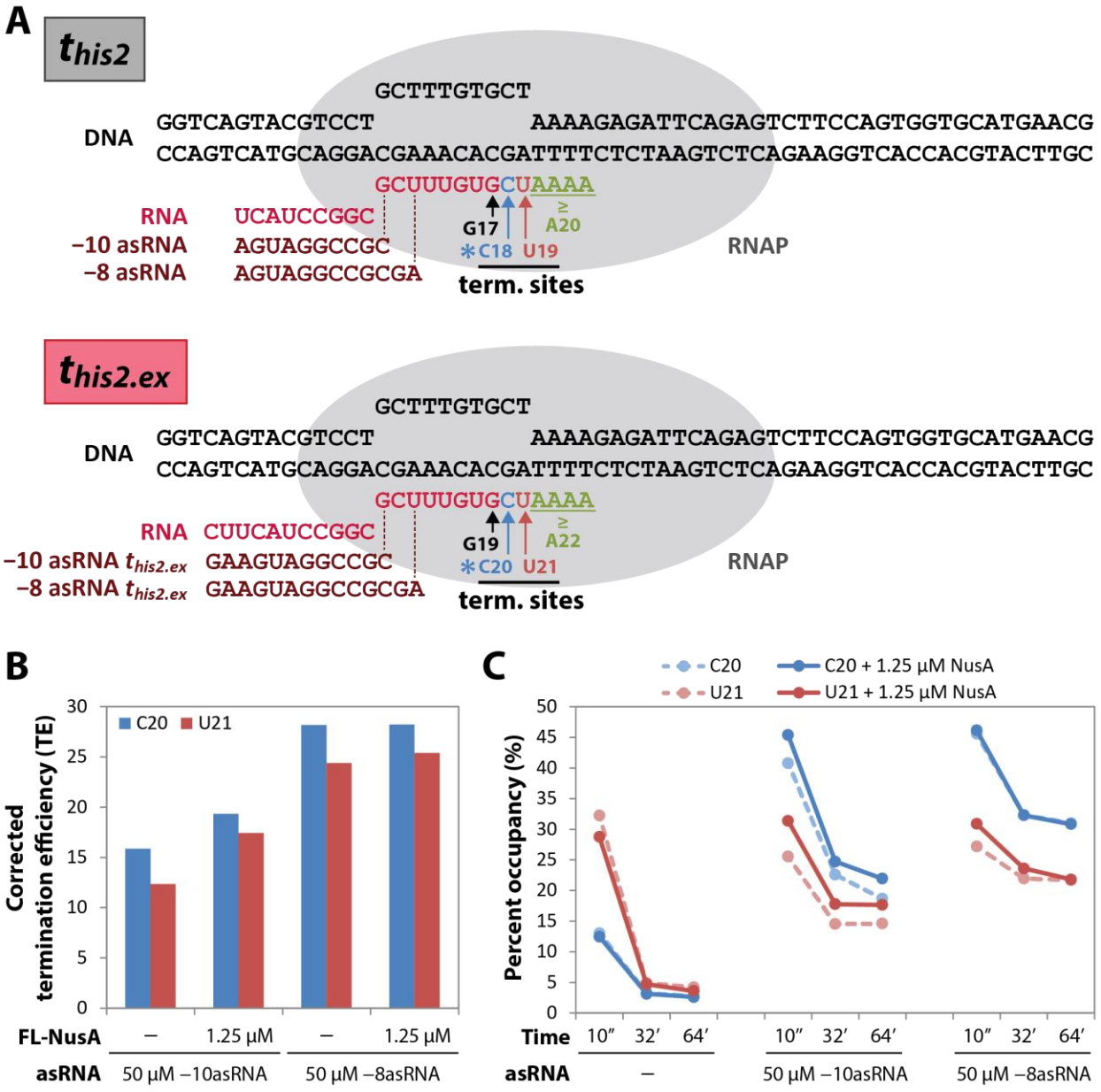
(A) The t_{his2} terminator scaffold designed for our scaffold-based termination assay (see also **Chapter 2**). **(B)** Saturating FL-NusA does not increase TE on t_{his2}. ECs were assembled with WT RNAP at position G17, and RNAs radiolabeled by addition of [α-^{32}P]-CTP. ECs were incubated with 1.25 μM NusA or an equivalent volume of NusA buffer, and transcription restarted by the addition of 100 μM UTP and 10 μM ATP in transcription buffer TB (see section **3.3 Materials & Methods**), in the absence or presence of 50 μM −10asRNA or −8asRNA. **(C)** Quantification of the gel panel shown in **(B)**.
TE(C18) = C18 / (C18+U19+A20) * 100; TE(U19) = U19 / (U19+A20) * 100.
Corrected TE = TE(+NusA) − TE(−NusA).

3.2.2 FL-NusA does not increase TE from t_{his2} with saturating asRNA, but may stimulate hairpin-stabilized pausing

To test if the short RNA substrate in the t_{his2} ECs prevented NusA activity on termination, we lengthened the RNA duplex of the T_{hp} by 2 bps to −21 (when the EC is positioned at U21; $t_{his2.ex}$ **Figure 3.3A**), thereby creating a better RNA substrate in the EC for NusA binding.

To test whether FL-NusA could enhance TE at saturating asRNA concentrations on this extended termination scaffold, ECs were assembled with WT RNAP at position G19, and the RNAs radiolabeled by addition of [α-^{32}P] CTP. The ECs were then incubated with NusA and elongated through U19 in the absence or presence of −1oasRNA($t_{his2.ex}$) or −8asRNA($t_{his2.ex}$), as before. As with t_{his2}, the effect of NusA on TE was minimal under these conditions, albeit with a slightly increased enhancement of TE with −1oasRNA($t_{his2.ex}$) **(Figure 3.3B)**.

We noted that a comparison of the 5 s timepoints indicated a small but discernable increase in EC occupancy at the C20 and U21 positions in the presence of NusA with −1oasRNA($t_{his2.ex}$) or −8asRNA($t_{his2.ex}$), but not in the absence of asRNA **(Figure 3.3C)**. From this NusA effect, we inferred that NusA was able to bind to and act on the $t_{his2.ex}$ ECs. Moreover, the increased occupancy of ECs at C20 and U21 in the presence of asRNA suggests that NusA has a small effect in favoring pausing on the termination pathway in response to RNA duplexes in the exit channel. However, this effect on pausing did not appear to be sufficient to substantially increase TE.



<u>Figure 3.3:</u> Extension of the RNA duplex on the t_{his2} terminator scaffold does not allow NusA-mediated termination enhancement

(A) The t_{his2} (top) and $t_{his2.ex}$ (bottom) termination scaffolds and corresponding asRNAs are shown. **(B)** Saturating FL-NusA does not increase termination efficiency (TE) on $t_{his2.ex}$. ECs were assembled with WT RNAP at position G19, and RNAs radiolabeled by addition of [α-^{32}P]-CTP. ECs were incubated with 1.25 µM NusA or an equivalent volume of NusA buffer, and transcription restarted by the addition of 100 µM UTP and 10 µM ATP in transcription buffer TB (see section **3.3 Materials & Methods**), in the absence or presence of 50 µM −10asRNA or −8asRNA. Timepoints were taken at 10 s, 32 min, and 64 min. TE(C20) = C20 / (C20+U21+A22) * 100; TE(U21) = U21 / (U21+A22) * 100. Corrected TE = TE(+NusA) − TE(−NusA). **(C)** Percent occupancy (%Occupancy) of ECs at 10 s timepoint indicate decreased elongation rates in the presence of asRNA and NusA, suggesting that NusA was able to

bind ECs, and that NusA may stabilize a paused intermediate on the termination pathway.
%Occupancy(C20) = C20 / (C20+U21+A22) * 100; %Occupancy(U21) = U21 / (C20+U21+A22) * 100.

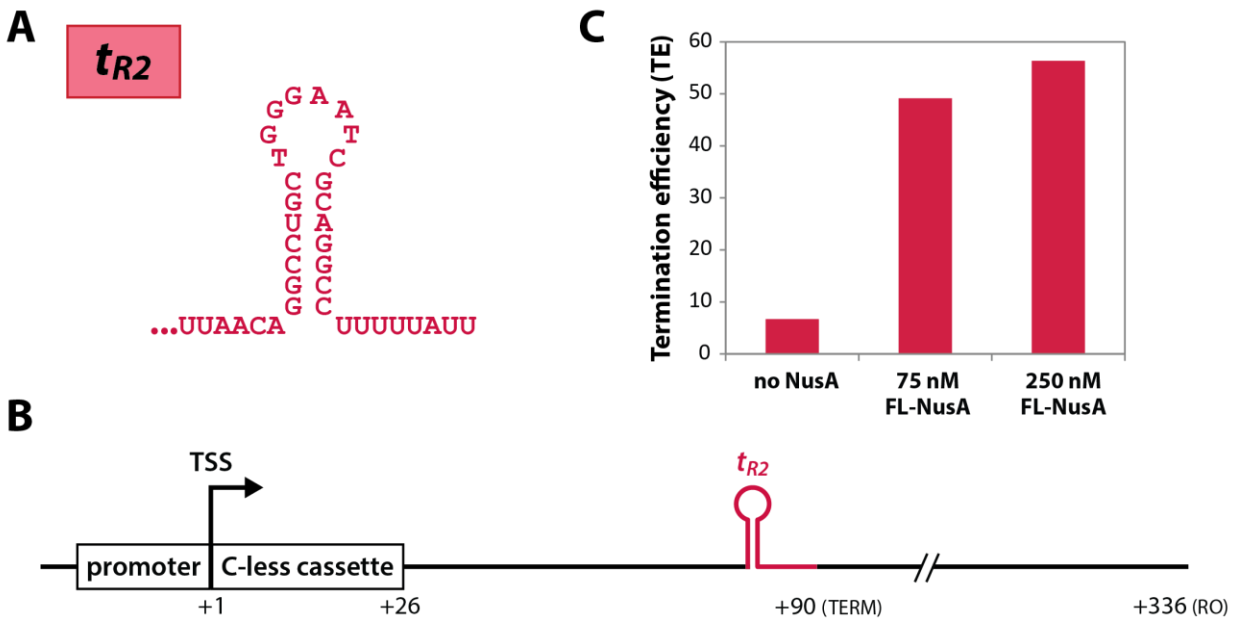
3.2.3 A scaffold-based system is unable to recapitulate NusA-effects on TE observed during promoter-initiated transcription

In a promoter-initiated assay, FL-NusA was shown to increase TE at the t_{R2} terminator from ~3% to ~35%, translating to a ~10-fold increase in TE due to NusA (**Figure 3.4**: unpublished results from Rachel Mooney; see also methods section **3.5.6**). Given the significantly smaller magnitude of NusA effects on TE observed on $t_{his2.ex}$, we designed a termination scaffold analogous to $t_{his2.ex}$ with the sequences from t_{R2} ($t_{R2.ex}$; **Figure 3.5A-B**).

To test whether FL-NusA enhances TE at saturating asRNA concentrations on $t_{R2.ex}$, ECs were assembled with WT RNAP at position A19 with [γ-^{32}P]-ATP-labeled RNAs. The ECs were then incubated with 500 nM NusA or an equivalent volume of NusA buffer, and then elongated through the termination positions U20 and U21 by incubation with UTP and GTP, in the absence or presence of −10asRNA($t_{R2.ex}$) or −8asRNA($t_{R2.ex}$). No effect of NusA was detectable for either asRNA **(Figure 3.5C)**, in stark contrast to the large effect of NusA on termination at t_{R2} during promoter-initiated transcription.

We noted that the buffer conditions were also different between the scaffold-based (TB; see methods section **3.5.3**) and promoter-initiated assays (TBtR2; see methods section **3.5.3**), with a relatively high concentration of destabilizing chloride ions in TBtR2 whereas acetate was used as the anion in the scaffold-based assays. Therefore we repeated the scaffold-based assay in the TBtR2 buffer to eliminate the possibility of diminished NusA effects in the scaffold-based assay arising as a consequence of the buffer conditions. To test whether NusA can stimulate RNA duplex formation on $t_{R2.ex}$, as seen for other termination scaffolds previously (36), both high (50 μM) and low (500 nM) concentrations of the asRNAs were tested. No effect of NusA was detectable at either high or low asRNA concentrations **(Figure 3.5D)** in the TBtR2 transcription buffer. Notably, there was very high background termination in TBtR2 even in the

absence of asRNA, as expected from a combination of high chloride ion concentration and the extremely unstable termination hybrid of $t_{R2.ex}$. We concluded that this set of terminator sequences was unable to recapitulate the effect of NusA in a scaffold-based system.



<u>Figure 3.4:</u> NusA significantly enhances termination at the t_{R2} terminator during promoter-initiated transcription

(unpublished data from Rachel Mooney)

(A) Sequence of the λt_{R2} terminator. **(B)** Promoter-initiated termination assay template structure. The construct contains a C-less cassette to allow ECs to be halted at position +26 (A26) by withholding CTP. Termination at t_{R2} produces a 90 nt product (TERM), and terminator bypass produces a 336 nt template run-off product (RO). **(C)** FL-NusA increases termination efficiency (TE) at t_{R2} in the promoter-initiated termination assay. Halted A26 ECs were formed by incubating 25 nM t_{R2} template and 62.5 nM RNAP holoenzyme with 150 μM ApU, 5 μM ATP, UTP, and GTP, and 0.6 μM (20 μCi) [α-^{32}P]-GTP in transcription buffer TBtR2 (see section **3.3 Materials & Methods**). A26 ECs were then incubated with 500 nM NusA or an equal volume of NusA storage buffer, and transcription restarted by addition of 400 μM NTPs, 100 μg rifampicin/mL and 0.5 units rRNasin/μL in TBtR2. Termination efficiency (TE) = TERM / (TERM+RO) * 100.

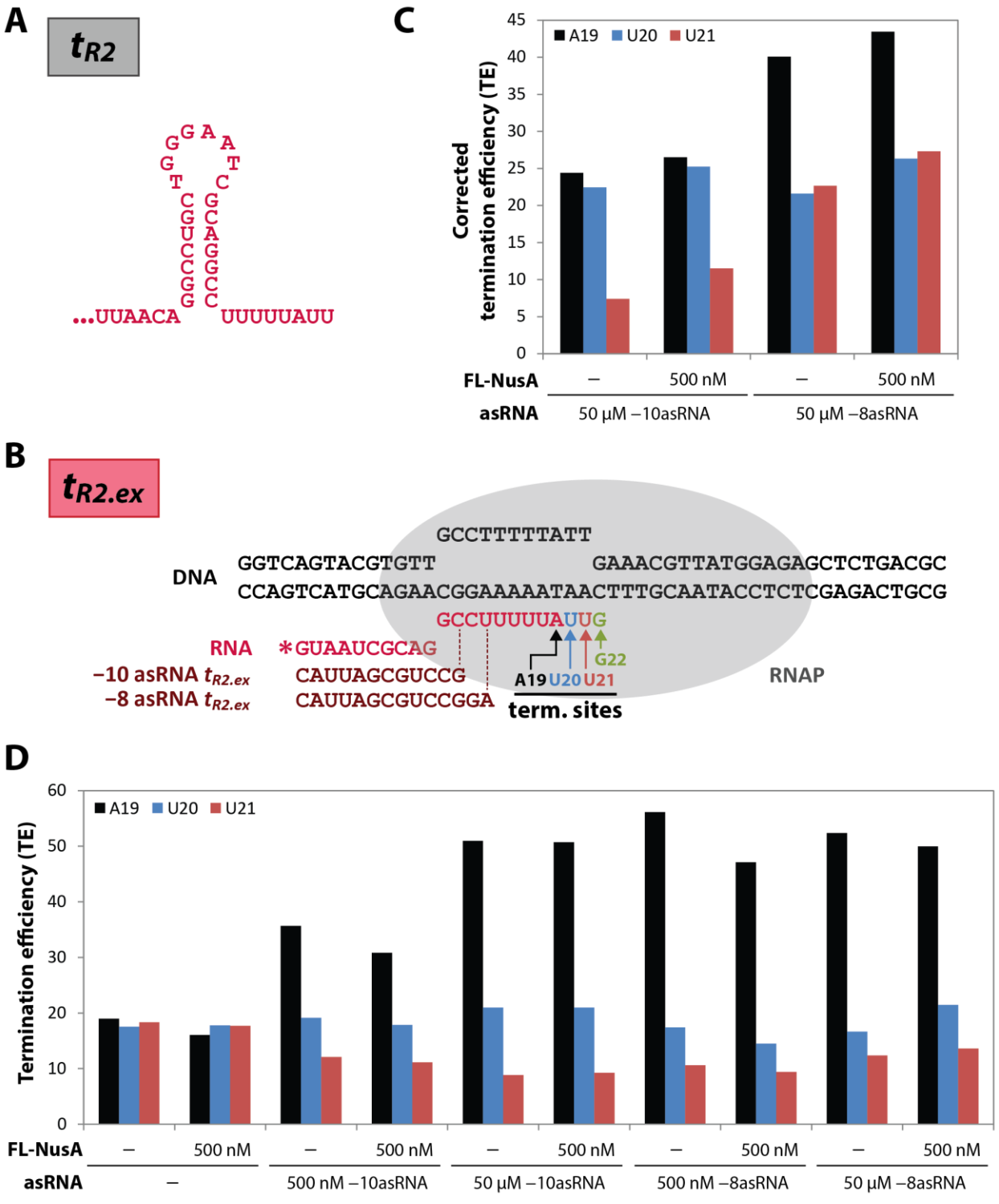


Figure 3.5: The $t_{R2.ex}$ terminator scaffold does not recapitulate NusA-mediated termination enhancement

(A) The sequence of the λt_{R2} terminator and **(B)** the $t_{R2.ex}$ termination scaffold and corresponding asRNAs designed based on t_{R2} are shown. **(C)** FL-NusA showed minimal effect on termination efficiency (TE)

from $t_{R2.ex}$. ECs were assembled with WT RNAP at position A19 with [γ-^{32}P]-ATP-end-labeled RNA. ECs were incubated with 500 nM NusA or an equivalent volume of NusA buffer, and transcription restarted by the addition of 100 μM UTP and 10 μM GTP in transcription buffer TB (see section **3.3 Materials & Methods**), in the absence or presence of 50 μM −1oasRNA or −8asRNA. **(D)** NusA-mediated TE enhancement on $t_{R2.ex}$ was not detectable with low asRNA concentrations, or in transcription buffer TBtR2 (see section **3.3 Materials & Methods**). Transcription assays were performed as in **(C)** in TBtR2, and in the absence or presence of 500 nM or 50 μM asRNA. TE(A19) = A19 / (A19+U20+U21+G22) * 100
TE(U20) = U20 / (U20+U21+G22) * 100; TE(U21) = U21 / (U21+G22) * 100.
Corrected TE = TE(+NusA) − TE(−NusA). Background TE was too high in transcription buffer TBtR2 (panel **(D)**) to calculate corrected TE.

3.2.4 Competition with secondary structures in the upstream RNA does not explain NusA activity at t_{R2}

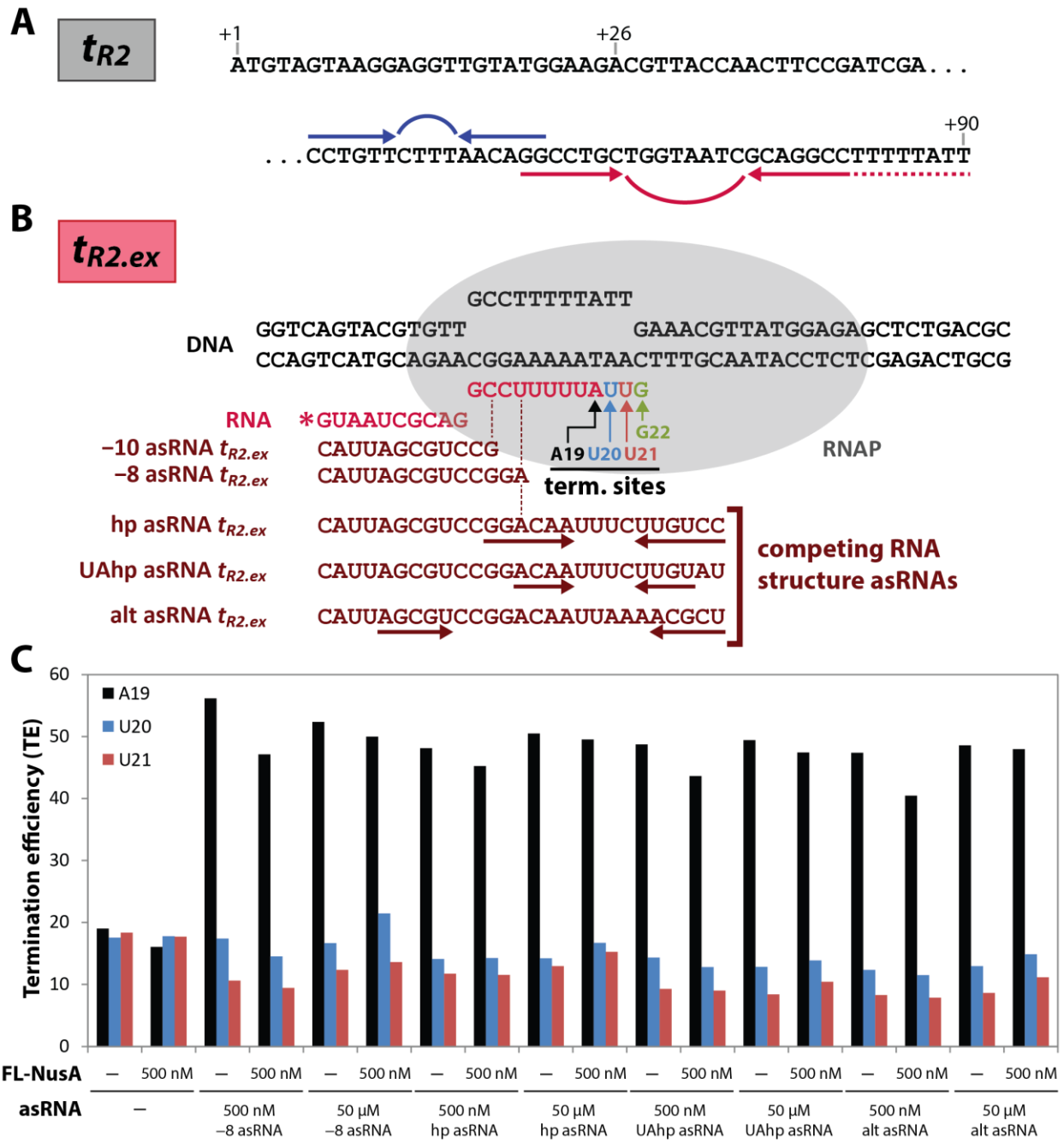
We considered that one possible reason for the inability of $t_{R2.ex}$ to recapitulate the effect of NusA on termination at t_{R2} during promoter-initiated transcription could be the absence of RNA upstream of the RNA duplex in the $t_{R2.ex}$ scaffold-based system. Consistent with this idea, the presence of upstream ssRNA has been shown to favor hairpin-stabilized pausing (34), likely through interactions with the S1 and KH domains that form a positive RNA binding surface continuous with the opening of the RNA exit channel (21). Moreover, as the formation of upstream RNA structures that compete with T_{hp} formation can also reduce TE (17, 39), NusA could increase TE by preventing upstream RNA structure formation through its interactions with ssRNA, or by favoring formation of the T_{hp}.

We used the co-transcriptional RNA folding prediction algorithm KineFold (40) to identify putative RNA structures that could form in the promoter-initiated t_{R2} termination assay. Interestingly, this analysis revealed one key RNA structure that would be expected to directly compete with extension of the t_{R2} T_{hp} **(Figure 3.6A)**. The relative stabilities of the two hairpins determined by MFold (41) ($\Delta G = -13.4$ kcal/mol for the T_{hp} *vs.* $\Delta G = -5.0$ kcal/mol) would suggest that the competing RNA structure may have a small but detectable effect in disfavoring termination.

To test whether NusA could enhance termination in the presence of competing structures, we designed asRNAs that *(i)* mimicked the RNA structure formed in the promoter-initiated transcript (hp asRNA($t_{R2.ex}$); **Figure 3.6B**), *(ii)* eliminated the base-pairing potential of the 2 bps in the competing RNA structure that would directly compete with T_{hp} extension (UAhp asRNA($t_{R2.ex}$); **Figure 3.6B**), or *(iii)* formed an RNA structure of similar stability to hp asRNA, but that competes for base pairing with the middle of the T_{hp} stem rather than with the

base of the T_{hp} (alt asRNA($t_{R2.ex}$); **Figure 3.6B**). The effect of NusA with these asRNAs was then determined as for −10asRNA($t_{R2.ex}$) and −8asRNA($t_{R2.ex}$). However, no effect of NusA was detectable with these asRNAs, at both high and low asRNA concentrations **(Figure 3.6C)**.

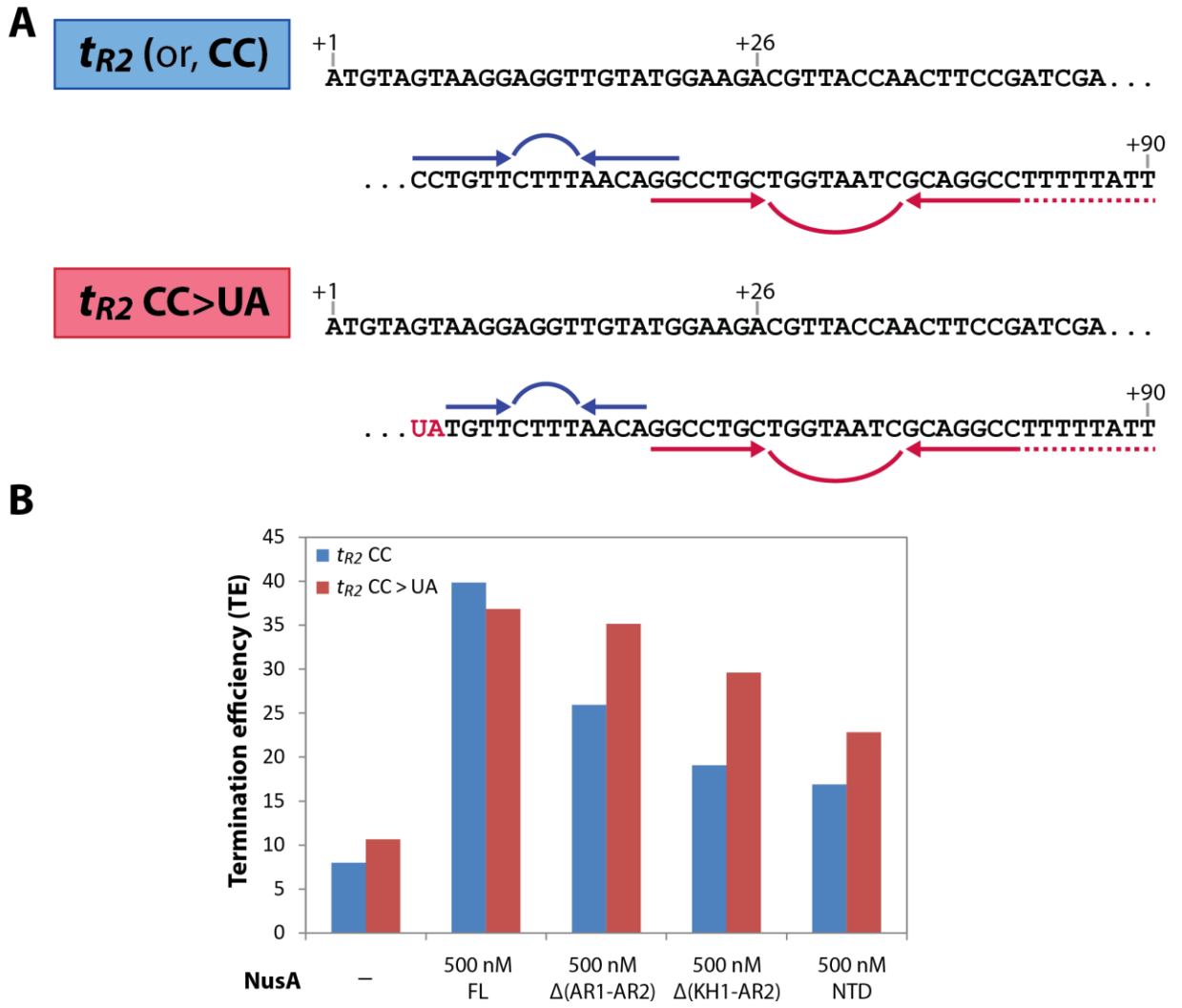
To test whether weakening the putative competing RNA structure in the promoter-initiated system would *(i)* increase TE in the absence of NusA, and *(ii)* eliminate or decrease the NusA-mediated TE enhancement, we modified the transcription template analogously to the UAhp asRNA (CC>UA mutation; **Figure 3.7A**). Halted A26 ECs were formed by incubating the t_{R2} (CC) or t_{R2} CC>UA template with RNAP holoenzyme and ApU, ATP, UTP, and GTP, and [α-^{32}P]-GTP. The halted A26 ECs were then incubated with NusA or an equal volume of NusA storage buffer, and transcription was restarted by the addition of NTPs. To ensure single-round transcription reactions, rifampicin was added with the NTPs. To our surprise, eliminating direct competition between the T_{hp} and competing RNA structure did not improve TE in the absence of NusA **(Figure 3.7B)**. As with the scaffold-system, there was no change in the NusA effect in the presence or absence of the CC>UA mutation. From these two experiments, we concluded that stabilization of the t_{R2} T_{hp} against competing RNA structures is the not the main role of NusA in termination enhancement.



<u>Figure 3.6:</u> The presence of competing RNA structures does not facilitate NusA-mediated termination enhancement on the t_{his2} terminator

(A) RNA-folding analysis of the transcript produced by the promoter-initiated termination assay reveals a putative RNA structure that may compete with t_{R2} T_{hp} formation. *Red arrows and arc* denote the t_{R2} T_{hp} stem and loop, respectively; *blue arrows and arc* denote the stem and loop of the putative competing RNA structure. **(B)** The $t_{R2.ex}$ termination scaffold is shown with asRNAs designed to mimic the competing RNA structure depicted in **(A)** (hp asRNA), weaken the competing RNA structure (UAhp asRNA), or form an alternate competing structure of similar stability (alt asRNA). **(C)** Competing RNA

structure formation in the asRNA does not recapitulate NusA-mediated enhancement of termination efficiency (TE) observed in the promoter-initiated termination assay. ECs were assembled with WT RNAP at position A19 with [γ-^{32}P]-ATP-end-labeled RNA. ECs were incubated with 500 nM NusA or an equivalent volume of NusA buffer, and transcription restarted by the addition of 100 μM UTP and 10 μM GTP in transcription buffer TBtR2 (see section **3.3 Materials & Methods**), in the absence or presence of 500 nM or 50 μM asRNA. TE(A19) = A19 / (A19+U20+U21+G22) * 100
TE(U20) = U20 / (U20+U21+G22) * 100; TE(U21) = U21 / (U21+G22) * 100.



<u>Figure 3.7:</u> Weakening competing RNA structure upstream of the terminator hairpin (T_{hp}) does not decrease NusA-mediated termination enhancement

(A) The transcripts produced by the promoter-initiated termination assay with the original tR2 sequence context (t_{R2} CC; top) and the mutation (t_{R2} CC>UA; bottom) are shown. *Red arrows and arc* denote the t_{R2} T_{hp} stem and loop, respectively; *blue arrows and arc* denote the stem and loop of the predicted competing RNA structure. The CC>UA mutation is expected to weaken the competing RNA structure and eliminate direct competition for base-pairing between the t_{R2} T_{hp} and the competing structure. **(B)** The CC>UA mutation does not increase termination efficiency (TE) in the absence of NusA, but protects against a slight NusA-TE-enhancement defect in the Δ(AR1-AR2)-NusA deletion mutant. Halted A26 ECs were formed by incubating 25 nM t_{R2} CC or t_{R2} CC>UA template with 62.5 nM RNAP holoenzyme, and 150 μM ApU, 5 μM ATP, UTP, and GTP, and 0.6 μM (20 μCi) [α-^{32}P]-GTP in transcription buffer TBtR2 (see section **3.3 Materials & Methods**). A26 ECs were then incubated with 500 nM WT or mutant NusA or an equal volume of NusA storage buffer, and transcription restarted by addition of 400 μM NTPs, 100 μg rifampicin/mL and 0.5 units rRNasin/μL in TBtR2. Termination efficiency (TE) = TERM / (TERM+RO) * 100, where TERM is the 90 nt termination product, and RO is the 336 nt termination bypass product.

3.2.5 Weakening competing RNA structures protects against the defect in TE-enhancement caused by deletion of the NusA AR domains

Given the modular structure of NusA, we were interesting in determining whether weakening the competing RNA structure had differential effects on FL-NusA versus incremental truncation mutants of NusA **(Figure 3.1)**. To this end, we determined TE at t_{R2} during promoter-initiated transcription on the CC and CC>UA templates with Δ(AR1-AR2)-NusA, Δ(KH1-AR2)-NusA, and NusA-NTD (*i.e.* Δ(S1-AR2)-NusA). Consistent with previously reported effects of NusA truncation on TE-enhancement (16, 21), we observed an incremental decrease in NusA-mediated TE-enhancement on the CC template upon the deletion of each module **(Figure 3.7B)**. In the presence of the CC>UA mutation, however, deletion of the NusA AR domains did not decrease NusA-mediated TE-enhancement, as observed on the CC template. The simplest interpretation of this result is that deletion of the AR domains decreased the affinity of NusA for the EC, which becomes more relevant when upstream RNA structures can compete with T_{hp} formation and increase terminator bypass slightly. However determining the exact mechanism, and the steps affected by AR deletion will require further investigation.

3.2.6 NusA does not facilitate A-tract-mediated TE-enhancement

Finally, we wanted to determine whether the effect of A- and U-tract pairing on TE-enhancement observed *in vivo* is facilitated by NusA. To test this idea, we designed asRNAs with A-tract extensions on the 5′ end that either *(i)* mimic the A-tract present in the endogenous t_{R2} sequence context and in the t_{R2} promoter-initiated assay (tR2A asRNA($t_{R2.ex}$); **Figure 3.8A**), or *(ii)* are completely complementary to the U-tract (A asRNA($t_{R2.ex}$); **Figure 3.8A**). The effect of NusA with these asRNAs was then determined as for −10asRNA($t_{R2.ex}$) and −8asRNA($t_{R2.ex}$). There appeared to be a modest NusA-mediated enhancement of TE for both A-tract asRNAs

(Figure 3.8B-C). However, it is unclear whether the effect was due to pairing of the A- and U-tracts, or simply due to the presence of upstream ssRNA. It will be instructive to repeat the experiment with both A-tract and non-A-tract 5′ sequence extensions on the asRNA.

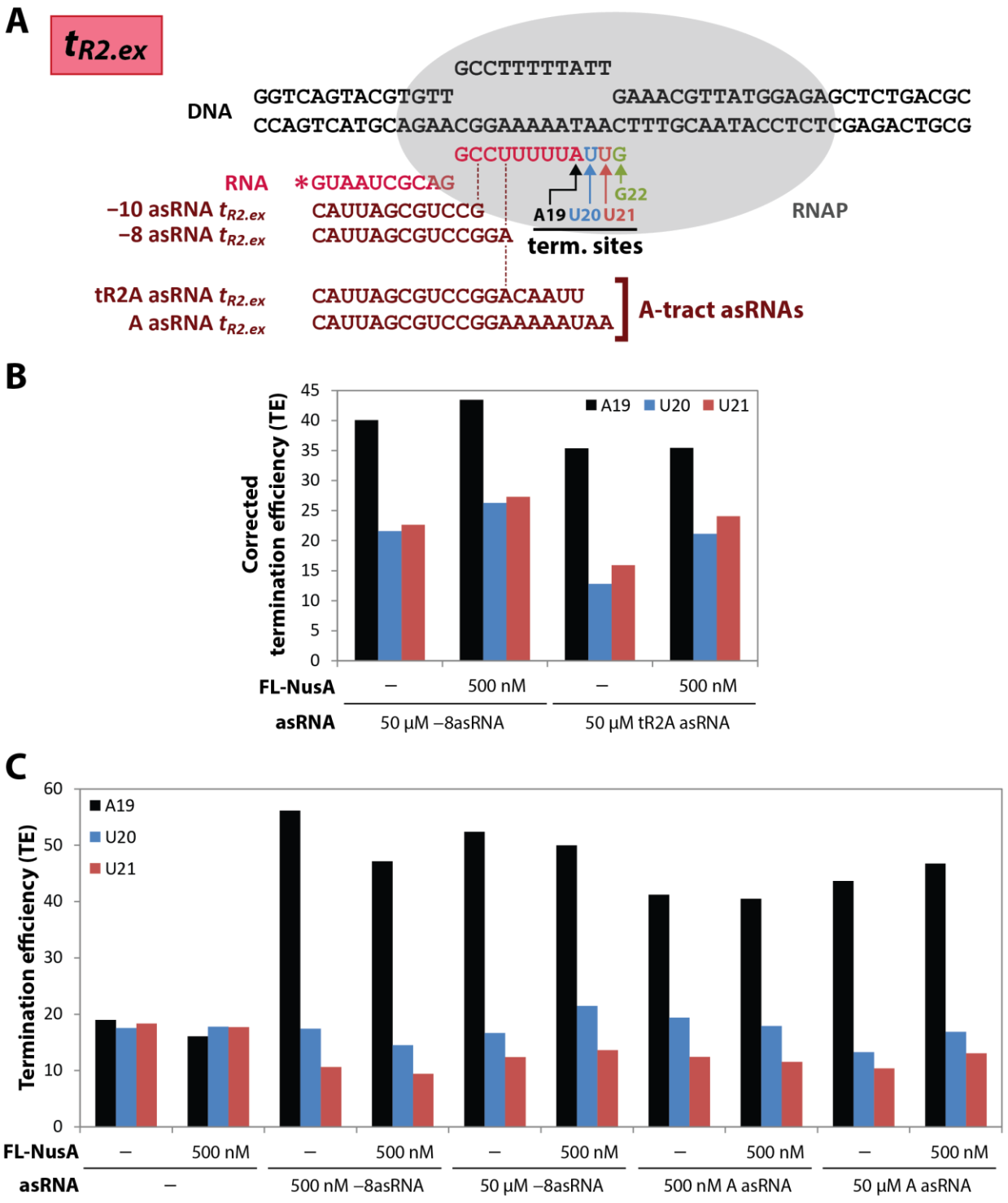


Figure 3.8: NusA does not facilitate A-tract mediated termination enhancement *in vitro*

(A) The $t_{R2.ex}$ termination scaffold is shown with asRNAs designed either to mimic the endogenous A-tract-like sequence upstream of the t_{R2} terminator (tR2A asRNA; compare with **Figure 3.4A**), or be completely complementary to the U-tract sequence (A asRNA). **(B-C)** A-tract-mediated enhancement of

TE is not facilitated by NusA in transcription buffer TB **(B)** or TBtR2 **(C)** (see section **3.3 Materials & Methods**). ECs were assembled with WT RNAP at position A19 with [γ-^{32}P]-ATP-end-labeled RNA. ECs were incubated with 500 nM NusA or an equivalent volume of NusA buffer, and transcription restarted by the addition of 100 μM UTP and 10 μM GTP in TB or TBtR2, in the absence or presence of 500 nM or 50 μM asRNA. TE(A19) = A19 / (A19+U20+U21+G22) * 100 TE(U20) = U20 / (U20+U21+G22) * 100; TE(U21) = U21 / (U21+G22) * 100. Corrected TE = TE(+NusA) – TE(–NusA). Background TE was too high in TBtR2 (panel **(C)**) to calculate corrected TE.

3.2.7 Future experiments of interest

Despite testing numerous sequences, we did not detect a strong NusA effect using this scaffold-based transcription system. In contrast, Shankar *et al.* previously described NusA effects on TE on a scaffold system with sequences similar to ours (36). However, two key differences exist between their assay setup and ours, which may explain their higher sensitivity to NusA. First, Shankar *et al.* measured EC dissociation over 10 min from ECs stalled at the termination site, allowing ample time for any NusA activity to take effect in the absence of elongation; in our assay setup ECs can escape the termination sites by elongation much more rapidly. Second, in the Shankar *et al.* assay, the ECs were halted at the termination site using cleavage-defective EcoRI-E111Q as a roadblock that contacted the downstream end of RNAP, where it could potentially affect both EC conformation and translocation state. One possibility would therefore be to replicate the Shankar *et al.* assay setup. However, the types of questions that can be addressed by this system are limited since it does not allow assessment of the contributions of elongation, pausing, and pause escape. Rather, thoughtful experiments using the promoter-initiated termination assay may provide further insights, which can then be applied to the more manipulable scaffold-based assay system. Genome-scale RNA-seq data is also available for a *Bacillus subtilis* NusA-depletion strain (35); further analysis of these data comparing the relative rates of termination (rather than TEs, as done in the published analysis) may also be informative. As such, I have outlined a set of analyses that could be applied to this dataset in **Appendix C**.

One key difference between the promoter-initiated and scaffold-based assays is the presence of initiation factor sigma. Whereas RNAP holoenzyme (RNAP core, $\alpha_2\beta\beta'\omega$ + sigma, σ) was used in the promoter-initiated assays to enable transcription initiation, RNAP core was used in the scaffold-based assays because core RNAP is sufficient for elongation. Given that sigma has been shown to stay associated with a subset of ECs after the transition from initiation to

elongation (42), and that sigma also binds the RNAP-FT (43), NusA and sigma compete for binding to the RNAP-FT during elongation (2, 44). One possible explanation for the specificity of NusA-mediated TE enhancement for promoter-initiated transcription assays (rather than scaffold-based assays) is that the presence of sigma inhibits termination on t_{R2}, and that the major effect of NusA on TE-enhancement depends on the displacement of sigma. This may be tested by *(i)* comparing the effect of NusA on termination from the scaffold-based system with RNAP core versus holoenzyme, or *(ii)* using the promoter-initiated system with a paramagnetic bead-tethered RNAP core, so sigma can be washed off from the A26 halted complexes.

Melting of the upstream 3 rU–dA bps of the U-tract upon extension of the T_{hp} has been shown to be important for termination (45-47) (see section **1.4.3a** and **Figure 1.3**).The small difference between the NusA effects on $t_{his2.ex}$ with the shorter (–1oasRNA($t_{his2.ex}$)) versus the longer (–8asRNA($t_{his2.ex}$)) RNA duplex **(Figure 3.3B)** may also suggest an effect of NusA in aiding melting of the upstream U-tract when T_{hp} extension is limited. This is consistent with previous reports of increased TE with NusA when the upstream 3 bps of the U-tract are interrupted with non- rU–dA bps (6, 35). However, the minimal effect of NusA on the $t_{R2.ex}$ scaffold was not specific to the short RNA duplex created by a –1oasRNA **(Figure 3.5C)**, inconsistent with a significant role of NusA in aiding melting of the upstream bps of the U-tract. This idea will require more direct testing by comparing scaffolds with and without non- rU–dA interruptions in the upstream U-tract. NusA may increase TE by more than one mechanism, any of which may become relevant depending on the terminator sequence and the state of RNAP.

3.3 Materials & Methods

3.3.1 Materials

DNA and RNA oligonucleotides **(Table 3.S1)** were obtained from Integrated DNA Technologies (IDT; Coralville, IA). All oligonucleotides were purified by 8 M urea denaturing polyacrylamide gel electrophoresis (PAGE) before use. [α-^{32}P]-CTP and [γ-^{32}P]-ATP were obtained from PerkinElmer Life Sciences (Warwick, RI); NTPs were from Promega (Madison, WI).

3.5.2 Proteins

E. coli RNAP core enzyme ($\alpha_2\beta\beta'$) (48), *E coli* σ^{70} (49), and wild-type and truncation mutant *E. coli* NusA proteins were purified as described previously (16) (see also **Table D.S1**). RNAP holoenzyme was formed by incubating 4 μM σ^{70} with 2 μM RNAP core in RNAP storage buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol, and 40% glycerol) supplemented with 10 mM dithiothreitol for 30 min at 30 °C.

3.5.3 *In vitro* EC reconstitution for scaffold-based assays

The nucleic acid scaffolds for EC reconstitution was formed by mixing 5 μM RNA and 10 μM T-DNA **(Table 3.S1)** in reconstitution buffer (RB; 10 mM Tris-HCl, pH 7.9, 40 mM KCl, and 5 mM $MgCl_2$) and heating to 95 °C for 2 min, cooling rapidly to 45 °C, then cooling to 25 °C in 2 °C increments for 2 min each, as described previously (50). For experiments with $t_{R2.ex}$, RNAs were first [γ-^{32}P]-ATP-end-labeled by incubating 50 μM RNA with~1 μM [γ-^{32}P]-ATP, 50 μM ATP, and T4 PNK (New England Biolabs, Ipswich, MA) for 30 min at 30 °C in T4 PNK buffer provided by the manufacturer.

ECs were reconstituted by incubating 2.5 μM *E. coli* RNAP core with 0.5 μM nucleic acid scaffold in either transcription buffer TB (20 mM Tris-OAc pH 8.0, 75 mM NaOAc, 1 mM

$Mg(OAc)_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 2.5% glycerol, and 25 μg of acetylated bovine serum albumin/mL) or TBtR2 (20 mM Tris-HCl pH 8.0, 20 mM NaCl, 14 mM $MgCl_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 10 mM dithiothreitol, and 10 μg of acetylated bovine serum albumin/mL), as indicated in figure legends, for 15 min at 37 °C. NT-DNA **(Table 3.S1)** was then added at 1.5 μM and incubated for another 15 min at 37 °C.

3.5.4 *In vitro* transcription with reconstituted ECs

ECs were diluted 5-fold in TB or TBtR2 (as indicated in figure legends) to 100 nM ECs. ECs for t_{his2} and $t_{his2.ex}$ experiments were then radiolabeled by incorporation labeling with ~0.1 μM [α-^{32}P]CTP and 0.9 μM CTP for 5 min. ECs were then incubated with 1.25 μM or 500 nM NusA or an equal volume of NusA storage buffer (20 mM Tris-HCl, pH7.9, 0.1 mM EDTA, 100 mM NaCl, 10 mM $MgCl_2$, 40% glycerol, and 1 mM DTT; (16)) for >3 min at 37 °C. Transcription was restarted by addition of diluted ECs to an equal volume of TB or TBtR2 (as indicated in figure legends) containing 2x concentrations of NTPs in the presence or absence of asRNAs, to achieve a final concentration of 100 μM UTP and 10 μM ATP (for t_{his2} and $t_{his2.ex}$) or 100 μM UTP and 10 μM GTP (for $t_{R2.ex}$), in the presence or absence of 50 μM or 500 nM asRNA. All NTP mixes were supplemented with $Mg(OAc)_2$ at a concentration equivalent to NTP concentration to prevent Mg^{2+} depletion effects. Reactions were stopped with an equal volume of 2x stop buffer (10 M urea, 50 mM EDTA, 90 mM Tris-borate buffer, pH 8.0, 0.02% bromophenol blue and 0.02% xylene cyanol).

RNA products were resolved by 8 M urea denaturing PAGE. Gels were exposed to phosphorimager screens, scanned using the Typhoon PhosphorImager and quantified using ImageQuant software (GE Healthcare).

3.5.5 Plasmid and linear template construction for promoter-initiated transcription

Plasmid pRM1065, encoding the template for the promoter-initiated t_{R2} termination assay, was a kind gift from Rachel A. Mooney (University of Wisconsin–Madison, WI). Plasmid pARS01, encoding the mutated CC>UA template for the promoter-initiated t_{R2} termination assay, was constructed from pRM1065 by Gibson assembly (Gibson Assembly HiFi 1-step kit, New England Biolabs, Ipswich, MA) from 2 fragments with 20 bps overlapping sequence at both ends. The CC>UA mutation was introduced through the forward PCR primer for fragment 1. Fragment 1 was designed to also contain the t_{R2} terminator and the p15A origin; fragment 2 was designed to contain the kanamycin resistance marker and the λP_R promoter required for the promoter-initiated termination assay. PCR primers used for amplification of fragment 1 were 5'-CGTTACCAACTTCCGATCGATATGTTCTTTAACAGGCCTGCTG (11340), and 5'-CAGAGATTTTGAGACACAACGTG (11341); fragment 2 was amplified using primers 5'-GTTGTGTCTCAAAATCTCTGATGTT (11342), and 5'-TCGATCGGAAGTTGGTAACG (11343).

Linear transcription templates for promoter-initiated assays were generated by PCR amplification from supercoiled plasmid DNA using PCR primers 5'-GCTGCCTGCACTAATGTTCC (9001) and 5'-CAGTTCCCTACTCTCGCATG (645), and purified by spermine precipitation.

3.5.6 *In vitro* promoter-initiated transcription

Halted A26 ECs were formed by incubating 25 nM t_{R2} CC (linear pRM1065 template) or t_{R2} CC>UA (linear pARS01 template) template and 62.5 nM RNAP holoenzyme in TBtR2 and combined with 150 μM ApU, 5 μM ATP, UTP, and GTP, and 0.6 μM (20 μCi) [α-^{32}P]-GTP for 15 min at 37°C. Halted A26 ECs were then incubated with 500 nM wild-type or truncation mutant NusA or an equal volume of NusA storage buffer for >3 min at 37 °C. Transcription was restarted by addition of halted ECs to an equal volume of TBtR2 containing 2x concentrations of

NTPs rifampicin and RNasin, to achieve a final concentration of 400 μM NTPs, 100 μg rifampicin/mL and 0.5 units rRNasin/μL (Promega, Fitchburg, WI). Reactions were stopped after 10 min with an equal volume of 2x stop buffer. RNA products were resolved by 8 M urea denaturing PAGE, and visualized as described above.

Table 3.S1: Oligonucleotides used in this study.

Oligo #	RNA / DNA	Sequence (5' to 3')	Use
8450	DNA	GGTCAGTACGTCCTGCTTTGTGCTAAAAGAGATTCAGAGTCTTCCAGTGGTGCATGAACG	NT-strand for t_{his2} and $t_{his2.ex}$ reconstitution
8451	DNA	CGTTCATGCACCACTGGAAGACTCTGAATCTCTTTTAGCACAAAGCAGGACGTACTGACC	T-strand for t_{his2} and $t_{his2.ex}$ reconstitution
10002	RNA	UCAUCCGGCGCUUUGUG	RNA strand for t_{his2} reconstitution
10428	RNA	CGCCGGAUGA	**−10**asRNA for t_{his2}
8713	RNA	AGCGCCGGAUGA	**−8**asRNA for t_{his2}
11187	RNA	CUUCAUCCGGCGCUUUGUG	RNA strand for $t_{his2.ex}$ reconstitution
11188	RNA	CGCCGGAUGAAG	**−10**asRNA($t_{his2.ex}$)
11189	RNA	AGCGCCGGAUGAAG	**−8**asRNA($t_{his2.ex}$)
11185	DNA	GGTCAGTACGTGTTGCCTTTTTATTGAAACGTTATGGAGAGCTCTGACGC	NT-strand for $t_{R2.ex}$ reconstitution
11186	DNA	GCGTCAGAGCTCTCCATAACGTTTCAATAAAAAGGCAAGACGTACTGACC	T-strand for $t_{R2.ex}$ reconstitution
11732	RNA	GCGAUCGCAGGCCUUUUUA	RNA strand for $t_{R2.ex}$ reconstitution
11741	RNA	GCCUGCGAUUAC	**−10**asRNA($t_{R2.ex}$)
11742	RNA	AGGCCUGCGAUUAC	**−8**asRNA($t_{R2.ex}$)
11316	RNA	CCUGUUCUUUAACAGGCCUGCGAUUAC	**hp**asRNA($t_{R2.ex}$)
11317	RNA	UAUGUUCUUUAACAGGCCUGCGAUUAC	**UA**asRNA($t_{R2.ex}$)
11318	RNA	UCGCAAAAUUAACAGGCCUGCGAUUAC	**alt**asRNA($t_{R2.ex}$)
11314	RNA	UUAACAGGCCUGCGAUUAC	**tR2A**asRNA($t_{R2.ex}$)
11315	RNA	AAUAAAAAGGCCUGCGAUUAC	**A**asRNA($t_{R2.ex}$)

3.4 REFERENCES

1. Nakamura Y & Uchida H (1983) Isolation of conditionally lethal amber mutations affecting synthesis of the nusA protein of Escherichia coli. *Mol Gen Genet* 190(2):196-203.

2. Mooney RA*, et al.* (2009) Regulator trafficking on bacterial transcription units in vivo. *Mol Cell* 33(1):97-108.

3. Greenblatt J, McLimont M, & Hanly S (1981) Termination of transcription by nusA gene protein of Escherichia coli. *Nature* 292(5820):215-220.

4. Farnham PJ, Greenblatt J, & Platt T (1982) Effects of NusA protein on transcription termination in the tryptophan operon of Escherichia coli. *Cell* 29(3):945-951.

5. Schmidt MC & Chamberlin MJ (1987) nusA protein of Escherichia coli is an efficient transcription termination factor for certain terminator sites. *J Mol Biol* 195(4):809-818.

6. Gusarov I & Nudler E (2001) Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell* 107(4):437-449.

7. Yakhnin AV & Babitzke P (2002) NusA-stimulated RNA polymerase pausing and termination participates in the Bacillus subtilis trp operon attenuation mechanism invitro. *Proc Natl Acad Sci U S A* 99(17):11067-11072.

8. Sigmund CD & Morgan EA (1988) Nus A protein affects transcriptional pausing and termination in vitro by binding to different sites on the transcription complex. *Biochemistry* 27(15):5622-5627.

9. Lau LF, Roberts JW, & Wu R (1983) RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. *J Biol Chem* 258(15):9391-9397.

10. Schmidt MC & Chamberlin MJ (1984) Binding of rho factor to Escherichia coli RNA polymerase mediated by nusA protein. *J Biol Chem* 259(24):15000-15002.

11. Chen CY & Richardson JP (1987) Sequence elements essential for rho-dependent transcription termination at lambda tR1. *J Biol Chem* 262(23):11292-11299.

12. Kainz M & Gourse RL (1998) The C-terminal domain of the alpha subunit of Escherichia coli RNA polymerase is required for efficient rho-dependent transcription termination. *J Mol Biol* 284(5):1379-1390.

13. Carlomagno MS & Nappo A (2003) NusA modulates intragenic termination by different pathways. *Gene* 308:115-128.

14. Cardinale CJ*, et al.* (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science* 320(5878):935-938.

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

16. Ha KS, Toulokhonov I, Vassylyev DG, & 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(5):708-725.

17. Larson MH, Greenleaf WJ, Landick R, & Block SM (2008) Applied force reveals mechanistic and energetic details of transcription termination. *Cell* 132(6):971-982.

18. Toulokhonov I, Artsimovitch I, & Landick R (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* 292(5517):730-733.

19. Santangelo TJ & Artsimovitch I (2011) Termination and antitermination: RNA polymerase runs a stop sign. *Nat Rev Microbiol* 9(5):319-329.

20. Shin DH*, et al.* (2003) Crystal structure of NusA from Thermotoga maritima and functional implication of the N-terminal domain. *Biochemistry* 42(46):13429-13437.

21. Worbs M, Bourenkov GP, Bartunik HD, Huber R, & Wahl MC (2001) An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol Cell* 7(6):1177-1189.

22. Gopal B*, et al.* (2001) Crystal structure of the transcription elongation/anti-termination factor NusA from Mycobacterium tuberculosis at 1.7 A resolution. *J Mol Biol* 314(5):1087-1095.

23. Schauer AT, Carver DL, Bigelow B, Baron LS, & Friedman DI (1987) lambda N antitermination system: functional analysis of phage interactions with the host NusA protein. *J Mol Biol* 194(4):679-690.

24. Mah TF, Kuznedelov K, Mushegian A, Severinov K, & Greenblatt J (2000) The alpha subunit of E. coli RNA polymerase activates RNA binding by NusA. *Genes Dev* 14(20):2664-2675.

25. Mah TF, Li J, Davidson AR, & Greenblatt J (1999) Functional importance of regions in Escherichia coli elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA. *Mol Microbiol* 34(3):523-537.

26. Traviglia SL, Datwyler SA, Yan D, Ishihama A, & Meares CF (1999) Targeted protein footprinting: where different transcription factors bind to RNA polymerase. *Biochemistry* 38(48):15774-15778.

27. Eisenmann A, Schwarz S, Prasch S, Schweimer K, & Rosch P (2005) The E. coli NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. *Protein Sci* 14(8):2018-2029.

28. Toulokhonov I & 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(5):1125-1136.

29. Yang X*, et al.* (2009) The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Rep* 10(9):997-1002.

30. Wang D, Severinov K, & 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(16):8433-8438.

31. Chan CL & 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(1):25-42.

32. Hein PP*, et al.* (2014) RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nat Struct Mol Biol* 21(9):794-802.

33. Wilson KS & von Hippel PH (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A* 92(19):8793-8797.

34. Kolb KE, Hein PP, & Landick R (2014) Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. *J Biol Chem* 289(2):1151-1163.

35. Mondal S, Yakhnin AV, Sebastian A, Albert I, & Babitzke P (2016) NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat Microbiol* 1:15007.

36. Shankar S, Hatoum A, & Roberts JW (2007) A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. *Mol Cell* 27(6):914-927.

37. Zhou J, Ha KS, La Porta A, Landick R, & Block SM (2011) Applied force provides insight into transcriptional pausing and its modulation by transcription factor NusA. *Mol Cell* 44(4):635-646.

38. Chen YJ*, et al.* (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat Methods* 10(7):659-664.

39. Lubkowska L, Maharjan AS, & Komissarova N (2011) RNA folding in transcription elongation complex: implication for transcription termination. *J Biol Chem* 286(36):31576-31585.

40. Xayaphoummine A, Bucher T, & Isambert H (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res* 33(Web Server issue):W605-610.

41. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13):3406-3415.

42. Bar-Nahum G & Nudler E (2001) Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. *Cell* 106(4):443-451.

43. Zuo Y & Steitz TA (2015) Crystal structures of the E. coli transcription initiation complexes with a complete bubble. *Mol Cell* 58(3):534-540.

44. Greenblatt J & Li J (1981) Interaction of the sigma factor and the nusA gene protein of E. coli with RNA polymerase in the initiation-termination cycle of transcription. *Cell* 24(2):421-428.

45. Gusarov I & Nudler E (1999) The mechanism of intrinsic transcription termination. *Mol Cell* 3(4):495-504.

46. Komissarova N, Becker J, Solter S, Kireeva M, & Kashlev M (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Mol Cell* 10(5):1151-1162.

47. d'Aubenton Carafa Y, Brody E, & Thermes C (1990) Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* 216(4):835-858.

48. Windgassen TA*, et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

49. Gribskov M & Burgess RR (1983) Overexpression and purification of the sigma subunit of Escherichia coli RNA polymerase. *Gene* 26(2-3):109-118.

50. Kyzer S, Ha KS, Landick R, & Palangat M (2007) Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J Biol Chem* 282(26):19020-19028.

Chapter 4:
Conclusions and future directions

Chapter contributions:

A.R.S. wrote all text; Rachel A. Mooney copyedited the text.

4.1 Conclusions and Significance of This Work

In striking contrast to the simplicity of the intrinsic termination signal, the mechanism by which ECs are terminated in response to the terminator involves a complex, multi-step pathway. As I described in **Chapter 1**, the sequence determinants as well as the mechanistic roles of individual components of the structured RNA termination signal have been extensively characterized. Moreover, with the exception of A-tract effects on termination efficiency (TE), the importance of the components of the RNA terminator have been validated by genome-scale analyses of terminator sequences and their TEs *in vivo* (1, 2). However, the role of structural rearrangements in the ~450 kDa enzyme in response to the termination signal remains almost completely undescribed. Although efforts have been made to identify modules of RNAP that contribute to the stability of the EC, it is unknown whether these modules play any role in termination aside from the final step in the termination pathway – EC dissociation. Proposals of a step in the termination pathway that functionally inactivates the EC active site prior to EC dissociation further suggest that domains that affect EC stability may not necessarily be involved in determining TE at a terminator. With recent advances in structural information about ECs and RNAP holoenzyme (3-9) as well as carefully crafted biochemical and biophysical analyses of dynamic RNAP modules (10-12), we are now better prepared to make predictions about the roles of different RNAP modules and design targeted RNAP mutants to test these ideas biochemically.

The chief goal of this work was to develop a system by which one could study the rate at which ECs become transcriptionally inactive, and in so doing, begin to identify mobile RNAP modules that are involved in facilitating the termination mechanism. The ability to measure termination *rate* was particularly important because TE is kinetically determined, and mutations of RNAP that affect elongation could impact TE readouts even without significantly affecting the termination mechanism itself. In fact, as intrinsic termination occurs by a multi-

step pathway, there are several points at which RNAP could bypass termination by elongating to the next sequence position; RNAP mutation could affect any of these steps. To this end, I setup a system by which elongation and termination kinetics could be monitored separately, and then, using kinetic modeling, I was able to isolate individual rates of elongation, pausing and termination on the intrinsic termination pathway. The system thus not only reports whether RNAP modules impact the termination mechanism, but also has the added benefit of reporting how each step and decision point along the termination pathway are affected by these modules. Finally, this assay provides a platform easily amenable to the study of various mutations in RNAP or the terminator sequence, and thus may be used to interrogate the roles of other RNAP modules in termination.

The development of this termination assay enabled me to determine the significant roles of one such module of RNAP – the trigger loop (TL). The active-site proximal, polymorphous TL module was previously implicated in intrinsic termination due to defects in EC dissociation observed upon mutation of the TL (13). A specific model for TL involvement was also proposed (the TH–T_{hp} contact model), in which it was argued that inactivation of the EC at a termination site requires direct contact between a folded TL and the terminator hairpin (T_{hp}), which was proposed to occupy the main cleft of RNAP (13). Our results contradict this model, and instead indicate that TL dynamics are crucial for termination rate enhancement (see **Chapter 2**). We also determined that the dynamics of the ~44-aa polymorphous arms of the TL – not the 188-aa domain inserted in the TL (SI3) – are responsible for TL effects on termination rate. We suggest that TL dynamics expand the conformational diversity of the EC, thereby allowing the EC to access conformational states that face lower energy barriers to termination (multistate-multipath model; MS-MP model). Our findings thus highlight the idea that the conformational fluctuations that constantly occur in a protein in solution can have profound effects on activation barriers to biological processes (14-16) (up to 75-fold in the case of the TL during termination). As such, extrinsic regulators likely act by capturing and stabilizing particular

conformations of the protein that experience either increased or decreased barriers to their inherent biological functions.

By using my termination assay, I was also able to isolate the effects of the TL and SI3 on the elongation and pausing rates on the termination pathway (see **Chapter 2**). We found that the TL and SI3 both increase TE by favoring the first step in the termination pathway – entry of ECs into the elemental paused state – and by inhibiting escape from the termination pathway by elemental or hairpin-stabilized pause escape. As a solvent-exposed domain of the TL, SI3 thus presents a regulatory target by which extrinsic factors could bind to ECs and modulate TE. Interestingly, SI3-like domains are present in a wide spectrum of bacteria, including pathogenic species in the Proteobacteria, Chlamydiae, Spirochaetes, and Fusobacteria phyla (17, 18), but are absent in human RNA polymerases. Our findings combined with previously described roles of SI3 in elongation, hairpin-stabilized pausing, and intrinsic cleavage (10, 19, 20) indicate that SI3 could present a potential drug-target for these bacterial pathogens.

Finally, as I described in **Chapter 3**, I began to investigate the role of NusA in termination. NusA is thought to be associated with all active ECs *in vivo*, and has been shown to enhance hairpin-stabilized pausing and intrinsic termination by stimulating hairpin formation in the RNA exit channel of RNAP (2, 21-24). Whereas the N-terminal domain of NusA (NusA-NTD) has been shown to be sufficient to reproduce the effects of full-length NusA (FL-NusA) on pausing, NusA-NTD is unable to completely recapitulate the effects of FL-NusA on intrinsic termination, even at saturating concentrations (22). The insufficiency of NusA-NTD for termination enhancement suggests that NusA likely has other roles in termination other than promoting T_{hp} formation. My termination assay seemed well-suited to address this question since at saturating concentrations of the antisense RNA oligonucleotide (asRNA) that creates an artificial T_{hp} together with the RNA positioned in the exit channel, the rate of T_{hp} formation would no longer be rate-determining for termination. Therefore any effects of NusA observed

under these conditions would reflect effects of NusA on termination that are separate from its role in stimulating hairpin formation. However, I found that the scaffold-based system was unable to recapitulate the dramatic effects of NusA on termination from actively transcribing ECs that we observed in a promoter-initiated termination assay. Although, the specific role(s) of NusA in termination remain elusive, my results point to a possible role of the antagonistic relationship between the initiation factor sigma (σ) and NusA for binding to the EC and affecting TE (see **Chapter 3**). Ideas to directly test this idea and perform analyses of existing RNA-seq datasets from NusA-depletion experiments are described in **Chapter 3** and **Appendix C**, respectively.

4.2 Possible Future Directions of Interest

4.2.1 What roles do other RNAP modules play in intrinsic termination?

EC conformational fluctuations during termination are likely to be impacted by the dynamics of multiple modules of RNAP. A number of mobile elements of RNAP are attractive candidates to investigate.

The clamp module

Opening of the clamp module, which forms many stabilizing contacts with the nucleic acids (NA) in its closed conformation, has been proposed to be important for loosening of contacts to the NA and EC dissociation (8, 12, 25). Whether clamp movement affects preceding steps on the termination pathway that could affect TE remains to be determined. However, consistent with a

role of the clamp in steps preceding EC dissociation, TE is affected by mutations in the switch domains (Sw1-5) (26, 27), thought to control clamp opening and closure (25).

Use of the closed clamp-stabilizing transcription factor, RfaH, may provide one way to interrogate the effects of clamp movement on the termination pathway. RfaH is a paralog of the universal transcription factor NusG (see section **1.3 Impact of Transcription Factors on Termination**) that stabilizes the closed clamp conformation through contacts to the clamp helices and lobe domains of RNAP, preventing clamp opening (28-30). RfaH has been shown to suppress pausing (23, 24, 29) and inhibit RNA duplex formation in the exit channel (24). It will be interesting to determine whether inhibition of clamp opening by RfaH impacts elemental pausing, pause escape, and EC inactivation at termination sites, aside from its effects on elongation and hairpin formation. By studying elongation and termination kinetics using the termination assay developed in **Chapter 2** with saturating RfaH, and modeling in the rates of RNA duplex formation in the presence of RfaH measured previously (24), the effects of clamp closure on these steps may be addressed.

The switch domains (Sw1-5)

The TE-altering point mutations isolated in the switch domains (26, 27) indicate that even small changes in individual switches are sufficient to alter termination or elongation rates, and thereby impact TE. It would therefore be informative to assess the effects of individual switch mutations by our termination assay, and simultaneously determine the effects of these mutations on clamp state. Our lab has developed a RNAP variant that reports clamp state in the presence of an oxidizing agent by forming one of two possible disulfide crosslinks between three engineered cysteine residues, depending on whether the clamp is open or closed (see individual cysteine pairs in (24)). By introducing switch mutations of interest into this triple-cysteine RNAP, we can determine the effect of the mutation on clamp state, and thus distinguish termination effects mediated by clamp movement from effects specific to switch dynamics.

As I described in Chapter 1, we propose that the interaction between the hydrophobic pocket formed by Sw3 and the –10 RNA base may limit complete extension of the T_{hp} upon hairpin nucleation (see section **1.4.3 Terminator Hairpin Extension**). The importance of this interaction in regulating EC inactivation may be tested by introducing a proline or charged residue in the Sw3 pocket to prevent binding to the –10 RNA base, and then measuring the effect on the rate of EC inactivation by our assay. It is possible that introducing a phenylalanine or tyrosine at the edge of the Sw3 pocket could instead stabilize the interaction by forming pi-stacking interactions with the –10 RNA base. These experiments would similarly need to be complemented with the triple-cysteine reporter assays to determine the effects of these mutations on the clamp in the presence and absence of an RNA duplex in the exit channel. However, previous work from our lab as well as others indicate that the clamp may be "wedged" open by the presence of an RNA duplex in the exit channel (8, 24), thus any observed effects of Sw3 on EC inactivation may act independently of clamp motion since the nucleated hairpin already causes clamp opening. It would also be instructive to determine whether changes in –10 RNA–Sw3 interactions alter the rate of T_{hp} extension directly. The use of a fluorescent base analog-based fluorescence quenching assay – similar to the assay used to measure on-rates of the asRNAs (see section **2.S1.1 Stopped-flow asRNA binding rate measurements** and (24)) – with the fluorescent probe located at –10 would be ideal in theory. However, one possible challenge in using such a setup is quenching of the fluorescent –10 RNA base due to interactions with the Sw3 pocket.

The bridge helix (BH)

Concerted movements of the BH and TL accelerate nucleotide addition and translocation but also stabilize paused intermediates (9, 31-33). Similar to the TL, mutations in the BH affect clamp conformation and RNA duplex formation in the exit channel (9, 24). Moreover, a kink

formed in the BH during pausing occludes entry of the +1 template base into the active site, directly impeding elongation (9). All of these roles of the BH suggest that it too may have a large impact on steps of the termination mechanism. It is also particularly attractive to consider the possibility that kinking of the BH could inactivate the EC by directly blocking translocation. Whether concerted BH and TL dynamics or a particular stabilized state of the BH favors termination would also be of interest, and could be addressed with the use of targeted BH mutations in our termination assay. These experiments could also be complemented with the clamp triple-cysteine reporter assay as well as another reporter assay developed in the lab to interrogate TL state.

4.2.2 Identifying other RNAP domains of interest by cryo-EM

The power of cryo-EM in elucidating structural rearrangements of protein complexes is becoming increasingly apparent as technological advances rapidly improve the technique (34). Visualization of elongating ribosomes by cryo-EM revealed the various structures adopted during translation elongation as well as their relative populations (35). By ordering the different conformational states of the ribosome in the elongation cycle, the authors could then determine the thermodynamic stabilities of each intermediate and predict the energetics for each step of translation elongation. The application of cryo-EM to the study of yeast RNA polymerase I initiation and elongation (36-38) identified both large-scale conformational changes of polymerase elements as well as rearrangements of modules as small as the BH and lid loop (see analogous 'lid' element of bacterial RNAP in **Figure 1.2**), thus validating the use of this technique to investigate RNAP structures. It will therefore be highly informative to study terminating complexes by cryo-EM. Because only meta-stable intermediates are captured by cryo-EM, it may also be useful to rapidly freeze termination reactions from a range of different temperatures, with the goal of *(i)* identifying the most physiologically abundant conformations

(from reactions at physiological temperatures), and *(ii)* enriching for relatively unstable termination intermediates that may be too transient to capture at physiological temperatures (from reactions at low temperatures). The low temperature cryo-EM in particular may help define as-of-yet unidentified steps in the termination pathway that are not accompanied by changes in NA structure.

As I alluded to in **Chapter 2**, studies of termination complexes by cryo-EM are likely to complement rather than supplant the information obtained from biochemical measurements of termination rate. The importance of TL dynamics in termination demonstrated by this work indicates that TL dynamics produce a rugged landscape of energy barriers to termination. Since the relative thermodynamic stabilities of different EC conformations need not be correlated to the heights of their corresponding energy barriers to termination, the population information obtained by cryo-EM may not be sufficient to predict the energetics between termination intermediates. Therefore, a combination of cryo-EM analyses of termination complexes along with biochemical interrogation of carefully designed RNAP mutants by our termination assay may provide a powerful way to define multiple steps and pathways of the intrinsic termination mechanism.

4.2.3 Is the MS-MP model conserved in factor-dependent termination?

The TL has also been implicated in Rho-dependent termination (39). Interestingly, in contrast to the requirement for a folded TL for intrinsic termination suggested by the TH–T_{hp} contact model (13), the model for Rho-dependent termination proposed by the same group suggested that the TL must be *un*folded for efficient termination. Whereas decreased EC dissociation was observed for unfolded TL-trapping monoclonal antibodies (mAbs) in intrinsic termination, the folded TL-trapping small-molecule tagetotoxin (Tgt) was shown to decreased Rho-dependent termination. However, the reciprocal tests were not performed for either termination pathway.

The MS-MP model proposed in this work provides a possible alternative explanation for these observations: Rather than preventing the TL from adopting a particular (folded or unfolded) conformation for Rho-dependent versus intrinsic termination, termination was inhibited in these experiments due to restricted TL dynamics, irrespective of the conformation in which the TL was stabilized.

The idea that TL dynamics favor Rho-dependent termination needs testing. It is conceivable that our termination assay could be adapted to address this question. EC elongation kinetics could be determined at a low NTP concentration over a few basepairs of a well characterized Rho-terminator (*e.g.* λt_{R1}) in the presence of mAbs or small-molecule inhibitors that the trap the TL in various conformations. By reconstituting these ECs on a scaffold with an upstream *rut* site in the RNA, the TEs of the trapped-TL RNAPs could then be determined at the same NTP concentrations for each mAb/inhibitor at each termination site. Such an assay would likely require optimization to achieve the right NTP concentrations to *(i)* enable both elongation of trapped-TL RNAPs and ATP-dependent Rho translocation, but *(ii)* sufficiently slow elongation to stimulate termination at the sites for which elongation rates were determined.

Mfd is another termination factor that acts to terminate transcription by removing ECs stalled at DNA lesions from the chromosome (40). Therefore, unlike intrinsic and Rho-dependent termination, the activity of Mfd is not thought to require specific sequence signatures, although the stability of DNA downstream of the stall site may have small effects on Mfd activity since it must be unwound for Mfd-dependent termination (41). The various stabilizing contacts of the EC may impact Mfd-termination, however, and RNAP conformational dynamics may also be of importance. It would therefore be interesting to determine whether TL dynamics have any bearing on this mechanism as well. A scaffold-based termination assay is highly amenable to studying Mfd-dependent termination. ECs with mutations in the TL (such as those described in **Chapter 2**) could be reconstituted on stable scaffold sequences, and tethered

to paramagnetic Co^{2+} beads through a His-tag on RNAP. By comparing the relative rates at which Mfd dissociates these various mutant TL ECs and releases radiolabeled RNAs into solution, we could assess whether the role of TL dynamics in termination is conserved for the Mfd-dependent termination mechanism.

Targeted testing of these and other RNAP modules in the different termination pathways will be critical for gaining a more complete understanding of these termination mechanisms. The subsequent investigation of the roles of dissociable transcription factors on termination may then help illuminate how these mechanisms are regulated *in vivo*.

4.3 REFERENCES

1. Chen YJ, *et al.* (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat Methods* 10(7):659-664.

2. Mondal S, Yakhnin AV, Sebastian A, Albert I, & Babitzke P (2016) NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat Microbiol* 1:15007.

3. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, & Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448(7150):157-162.

4. Vassylyev DG, *et al.* (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448(7150):163-168.

5. Vassylyev DG, *et al.* (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* 417(6890):712-719.

6. Wang D, Bushnell DA, Westover KD, Kaplan CD, & Kornberg RD (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127(5):941-954.

7. Korzheva N, *et al.* (2000) A structural model of transcription elongation. *Science* 289(5479):619-625.

8. Sekine S, Murayama Y, Svetlov V, Nudler E, & Yokoyama S (2015) Ratcheting of RNA polymerase toward structural principles of RNA polymerase operations. *Transcription* 6(3):56-60.

9. Weixlbaumer A, Leon K, Landick R, & Darst SA (2013) Structural basis of transcriptional pausing in bacteria. *Cell* 152(3):431-441.

10. Windgassen TA, *et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

11. Nayak D, Voss M, Windgassen T, Mooney RA, & Landick R (2013) Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol Cell* 50(6):882-893.

12. Chakraborty A, *et al.* (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science* 337(6094):591-595.

13. Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S, & Nudler E (2007) An allosteric path to transcription termination. *Mol Cell* 28(6):991-1001.

14. Frauenfelder H, Sligar SG, & Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254(5038):1598-1603.

15. Noe F, Schutte C, Vanden-Eijnden E, Reich L, & Weikl TR (2009) Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc Natl Acad Sci U S A* 106(45):19011-19016.

16. Wang H, *et al.* (2007) Protein dynamics control the kinetics of initial electron transfer in photosynthesis. *Science* 316(5825):747-750.

17. Iyer LM, Koonin EV, & Aravind L (2004) Evolution of bacterial RNA polymerase: implications for large-scale bacterial phylogeny, domain accretion, and horizontal gene transfer. *Gene* 335:73-88.

18. Lane WJ & Darst SA (2010) Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J Mol Biol* 395(4):671-685.

19. Zakharova N, Bass I, Arsenieva E, Nikiforov V, & Severinov K (1998) Mutations in and monoclonal antibody binding to evolutionary hypervariable region of Escherichia coli RNA polymerase beta' subunit inhibit transcript cleavage and transcript elongation. *J Biol Chem* 273(38):24912-24920.

20. Artsimovitch I, Svetlov V, Murakami KS, & 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(14):12344-12355.

21. Toulokhonov I, Artsimovitch I, & Landick R (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* 292(5517):730-733.

22. Ha KS, Toulokhonov I, Vassylyev DG, & 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(5):708-725.

23. Kolb KE, Hein PP, & Landick R (2014) Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. *J Biol Chem* 289(2):1151-1163.

24. Hein PP*, et al.* (2014) RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nat Struct Mol Biol* 21(9):794-802.

25. Gnatt AL, Cramer P, Fu J, Bushnell DA, & Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* 292(5523):1876-1882.

26. Weilbaecher R, Hebron C, Feng G, & Landick R (1994) Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev* 8(23):2913-2927.

27. Rutherford ST, Villers CL, Lee JH, Ross W, & Gourse RL (2009) Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev* 23(2):236-248.

28. Sevostyanova A, Belogurov GA, Mooney RA, Landick R, & Artsimovitch I (2011) The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol Cell* 43(2):253-262.

29. Svetlov V, Belogurov GA, Shabrova E, Vassylyev DG, & Artsimovitch I (2007) Allosteric control of the RNA polymerase by the elongation factor RfaH. *Nucleic Acids Res* 35(17):5694-5705.

30. Belogurov GA*, et al.* (2007) Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol Cell* 26(1):117-129.

31. Bar-Nahum G*, et al.* (2005) A ratchet mechanism of transcription elongation and its control. *Cell* 120(2):183-193.

32. Toulokhonov I, Zhang J, Palangat M, & Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* 27(3):406-419.

33. Silva DA*, et al.* (2014) Millisecond dynamics of RNA polymerase II translocation at atomic resolution. *Proc Natl Acad Sci U S A* 111(21):7665-7670.

34. Binshtein E & Ohi MD (2015) Cryo-electron microscopy and the amazing race to atomic resolution. *Biochemistry* 54(20):3133-3141.

35. Fischer N*, et al.* (2015) Structure of the E. coli ribosome-EF-Tu complex at <3 A resolution by Cs-corrected cryo-EM. *Nature* 520(7548):567-570.

36. Engel C, Plitzko J, & Cramer P (2016) RNA polymerase I-Rrn3 complex at 4.8 A resolution. *Nat Commun* 7:12129.

37. Pilsl M*, et al.* (2016) Structure of the initiation-competent RNA polymerase I and its implication for transcription. *Nat Commun* 7:12126.

38. Tafur L*, et al.* (2016) Molecular Structures of Transcribing RNA Polymerase I. *Mol Cell* 64(6):1135-1143.

39. Epshtein V, Dutta D, Wade J, & Nudler E (2010) An allosteric mechanism of Rho-dependent transcription termination. *Nature* 463(7278):245-249.

40. Selby CP & Sancar A (1993) Molecular mechanism of transcription-repair coupling. *Science* 260(5104):53-58.

41. Park JS & Roberts JW (2006) Role of DNA bubble rewinding in enzymatic transcription termination. *Proc Natl Acad Sci U S A* 103(13):4870-4875.

Appendix A:

Dissociation assays as a metric for intrinsic termination

Chapter contributions:

A.R.S. performed all experiments, experimental design and analysis, and prepared all text/figures; Robert Landick oversaw the experimental design and text/figure preparation.

A.1 Introduction

Transcription termination is a kinetically controlled process, where the relative rates of elongation and termination determine termination efficiency (TE). Thus, to eliminate the confounding effects of elongation, previous termination studies have utilized assays that measure dissociation from artificially stalled ECs that are not subject to variable elongation rates (1-13). In these dissociation assays, ECs are *(i)* assembled (*i.e.* reconstituted) using two complementary DNA strands, an RNA primer (together forming a scaffold) and core RNA polymerase (RNAP) near the termination site, *(ii)* "walked" to the termination site from a promoter by sequential NTP addition and wash steps, or *(iii)* physically stopped at the termination site by a protein roadblock just downstream. Using this concept, we designed a pair of dissociation assays on a scaffold-based system that allowed EC reconstitution just 2 nt upstream of the termination site to enable testing of RNAP trigger loop (TL) mutants that are severely defective for elongation.

A.2 Results

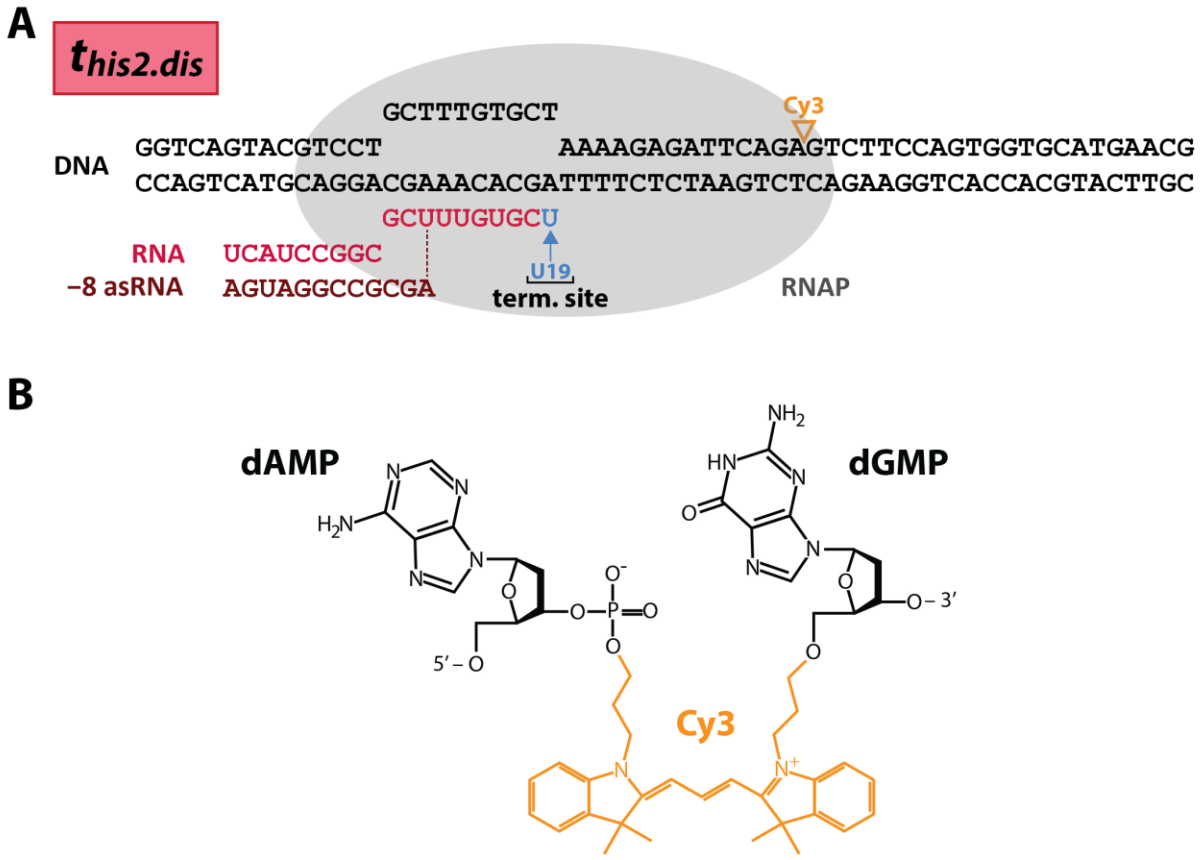
A.2.1 Two dissociation assays

Two design features were required to enable efficient reconstitution 2 nt upstream of a termination site, which overlaps the weak rU–dA bps of the terminator U-tract **(Figure A.1A)**. First, the termination hybrid was stabilized by the addition of 3 GC bp interruptions to the terminator U-tract (rather than the typical 0-2 GC bp interruptions) to prevent spontaneous dissociation of the ECs observed on other scaffold sequences (see section **A.2.6 Challenges with previously tested scaffolds**). Second, a previously designed oligonucleotide-mediated release (OMR) strategy (1) was used, in which the EC is reconstituted over a partial terminator

sequence where the upstream arm of the terminator hairpin (T_{hp}) is removed from the scaffold RNA sequence. An RNA or DNA oligonucleotide complementary to the RNA positioned in the RNAP exit channel (asRNA/asDNA) is then added *in trans* to form an artificial T_{hp} and stimulate termination. We setup two dissociation assays with this design **(Figure A.2)**: *(i)* a fluorescence anisotropy-based dissociation assay to enable the measurement of any fast dissociation kinetics of DNA, and *(ii)* a radioactive magnetic bead dissociation assay, to simultaneously measure the rates of RNA and DNA dissociation from RNAP.

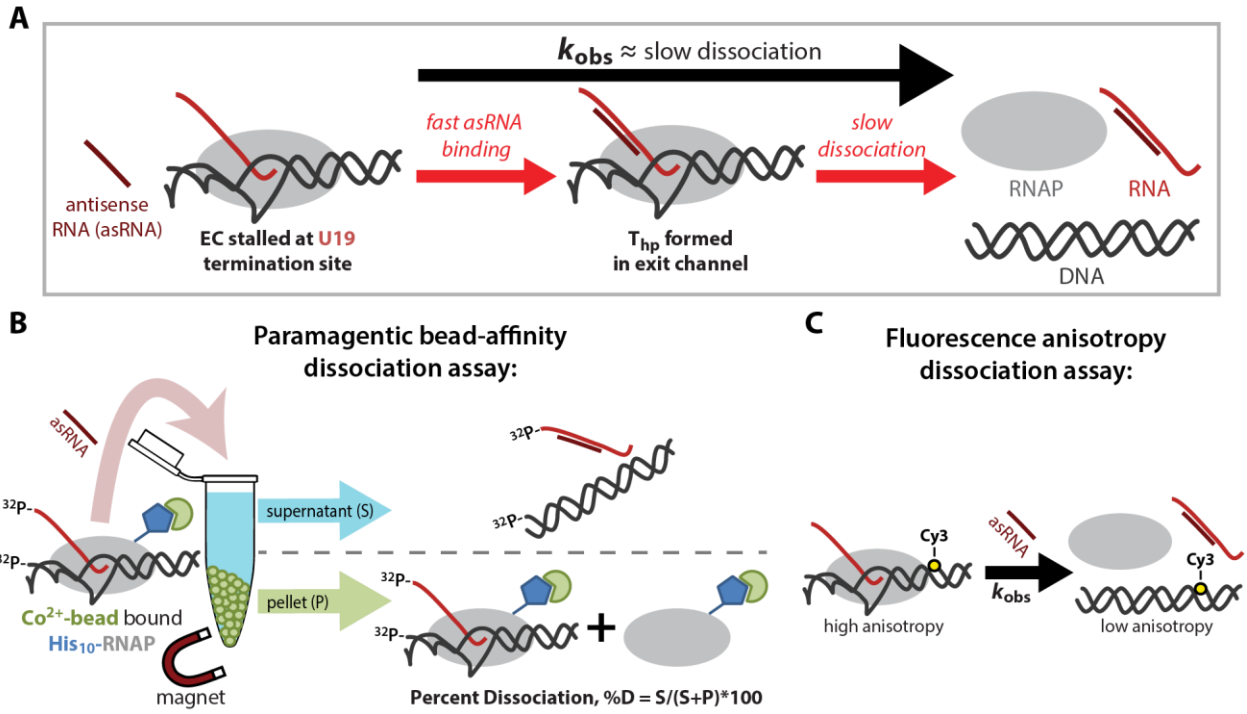
In the magnetic bead dissociation assay **(Figure A.2B)**, ECs were assembled at position C18 on the $t_{his2.dis}$ scaffold **(Figure A.1A)** with radiolabeled RNA and non-template DNA (NT-DNA) and His_{10}-tagged RNAP. The ECs were then tethered through the His-tag to paramagnetic Co^{2+}-beads, and elongated to the U19 termination position by the addition of UTP. Termination was then initiated by the addition of asRNA, and samples of the whole reaction and supernatant (containing RNA/DNA released from RNAP) were taken at various timepoints to determine the rate at which radiolabeled RNA and NT-DNA are released from RNAP.

ECs were assembled similarly for the fluorescence anisotropy assay, at position C18 on the $t_{his2.dis}$ scaffold, without radioactive labels on the RNA and NT-DNA. Rather, the doubly-tethered Cy3-fluorophore inserted in the backbone of the NT-DNA **(Figure A.1B)** enabled anisotropy measurements as the ~45 kDa fluorescent DNA is dissociated from the ~395 kDa EC **(Figure A.2C)**. The ECs were then elongated to U19 by incubation with UTP, and placed in one syringe of the Stopped-Flow apparatus; asRNA in the other syringe. Termination could then be initiated by rapid mixing of the ECs with asRNA, and monitored by the Stopped-Flow detectors in real time as a function of decrease in anisotropy.



<u>Figure A.1:</u> The $t_{his2.dis}$ scaffold design for dissociation assays.

(A) Scaffold $t_{his2.dis}$, used for both dissociation assays. The position of Cy3 fluorophore insertion in the NT-DNA backbone is indicated. Dissociation was only monitored from one site on this scaffold because complexes were extended to U19 before addition of asRNA. **(B)** Cy3 is doubly-tethered to the DNA backbone by the attachment chemistry depicted. The double tether was used to decrease free-rotation of the fluorophore, which would decrease the sensitivity of anisotropy measurements.



D

	NT release			RNA release		
	TE%*	k_{obs} (x10^{-3} s^{-1})	ratio**	TE%*	k_{obs} (x10^{-3} s^{-1})	ratio**
WT	68 ± 4	9 ± 1	1	84 ± 7	19 ± 3	1
ΔTL	75 ± 6	17 ± 3	**1.8**	83 ± 10	26 ± 3	**1.4**
ΔSI3	76 ± 6	11 ± 4	**1.2**	87 ± 8	20 ± 3	**1.0**
Ala$_6$	63 ± 7	8 ± 3	**0.9**	87 ± 22	20 ± 5	**1.0**
LTPP	74 ± 12	22 ± 1	**2.3**	84 ± 21	25 ± 8	**1.3**

*TE% (termination efficiency) = D_0; **ratio = k_{obs}(RNAP variant) / k_{obs}(WT)

E

	NT release	
	k_{obs} (x10^{-3} s^{-1})	ratio*
WT	3 ± 0.1	1
ΔTL	7 ± 1.2	**2.3**
ΔSI3	6 ± 0.5	**1.9**

*ratio = k_{obs}(RNAP variant) / k_{obs}(WT)

<u>Figure A.2:</u> Measurement of dissociation rates from stalled ECs.

(A) Schematic depicting dissociation assays using an asRNA-mediated termination strategy. Observed reaction rates (k_{obs}) are rate-limited by the slower dissociation step, resulting in an estimate for dissociation rate. **(B)** Paramagnetic bead-affinity based dissociation assay. ECs were assembled with ^{32}P-end-labeled C18 RNA and NT-DNA (radiolabel indicated by *asterisks*) and tethered to paramagnetic Co^{2+}-beads (*green circles*) using His_{10}-tagged (*blue pentagons*) RNAP (*grey ovals*). ECs were then extended to U19, and dissociation initiated by adding 100 μM –8asRNA. Percent dissociation (%D) was calculated as the proportion of RNA or NT-DNA detected in the supernatant (S) at each timepoint. Whole (W) = pellet (RNAP-bound RNA) + supernatant (S; released RNA). Data were fit to the single-exponential equation $D_t = D_0 + A(1 - e^{-k_{obs}t})$, where t = time (s), D_o = initial %D, D_t = %D at time t, A = amplitude of the fluorescence change, and k_{obs} = dissociation rate as depicted in (A). **(C)** Fluorescence anisotropy-based dissociation assay. ECs were assembled at C18 with Cy3-labeled NT-strand (8288; **Table A.S3**), and dissociation was initiated by adding 100 μM –8asRNA. Dissociation of fast-tumbling Cy3-labeled DNA (~45 kDa) from the slow-tumbling ECs (~395 kDa) was measured using KinTek stopped-flow apparatus. Anisotropy curves were fit to the single-exponential equation $r_t = r_0 + Ae^{-k_{obs}t}$, where t = time (s), r_o = initial anisotropy, r_t = anisotropy at time t, A = amplitude of the fluorescence change, and k_{obs} = dissociation rate as depicted in (A). Comparable dissociation rates are obtained from both the

(D) paramagnetic bead-affinity and **(E)** fluorescence anisotropy assays. These rates are too slow to compete with T_{hp} pause escape rates with −8asRNA as measured by our assay **(Table 2.1B)** and are inconsistent with the rates at which ECs become resistant to pause escape **(Table 2.1A)**. Therefore, these rates are unlikely to determine termination efficiency. These assays are also insensitive to mutations in the TL, contrary to our findings with actively transcribing ECs **(Table 2.1A)**. Errors represent SD from ≥3 independent experimental replicates.

A.2.2 NT-DNA dissociation rates from stalled ECs from both assays are comparable

The rate of NT-DNA dissociation for WT ECs determined by the magnetic bead-affinity ([9 ± 1] x 10^{-3} s^{-1}; **Figure A.2D**) and anisotropy ([3 ± 0.1] x 10^{-3} s^{-1}; **Figure A.2D**) dissociation assays were comparable. However, it is of note that these rates are significantly slower than the rates of EC inactivation measured on the same sequences by our termination assay ([100 ± 30] x 10^{-3} s^{-1}; see termination rate of WT ECs with −8asRNA, **Table 2.1A** and section **A.3 Discussion**).

A.2.3 RNA dissociates from stalled ECs more rapidly than NT-DNA

RNA and NT-DNA dissociation rates could be measured simultaneously by the magnetic bead-affinity dissociation assay. RNA dissociated ~2-fold more rapidly than NT-DNA for WT ECs ([19 ± 3] x 10^{-3} s^{-1} for RNA *vs.* [9 ± 1] x 10^{-3} s^{-1} for NT-DNA) and almost all other TL mutants tested (**Figure A.2D**; see also section **A.2.4 Dissociation assays do not detect significant effects of TL mutation**). The faster RNA dissociation rate is consistent with the idea that melting of the upstream 3-4 bps of the RNA–DNA hybrid (see section **1.4.3 Terminator hairpin extension**) sufficiently weakens the hybrid such that it can be outcompeted by reannealing of the DNA–DNA bps that must be melted to form the transcription bubble. This would result in bubble collapse and RNA exclusion from the transcription bubble, allowing dissociation of the RNA as the contacts between the exit channel and ssRNA (normally present in an EC (14, 15)) are likely already diminished upon formation of the RNA duplex in the exit channel. Contacts between the clamp and switch regions with the DNA in the main cleft and downstream, DNA channel (14, 15) likely slow DNA dissociation relative to RNA. However, this order of dissociation needs verification from actively elongating ECs (see section **A.3**

Discussion). It would also be instructive to determine dissociation rates with different scaffold sequences to determine if the order of RNA versus DNA release is obligate.

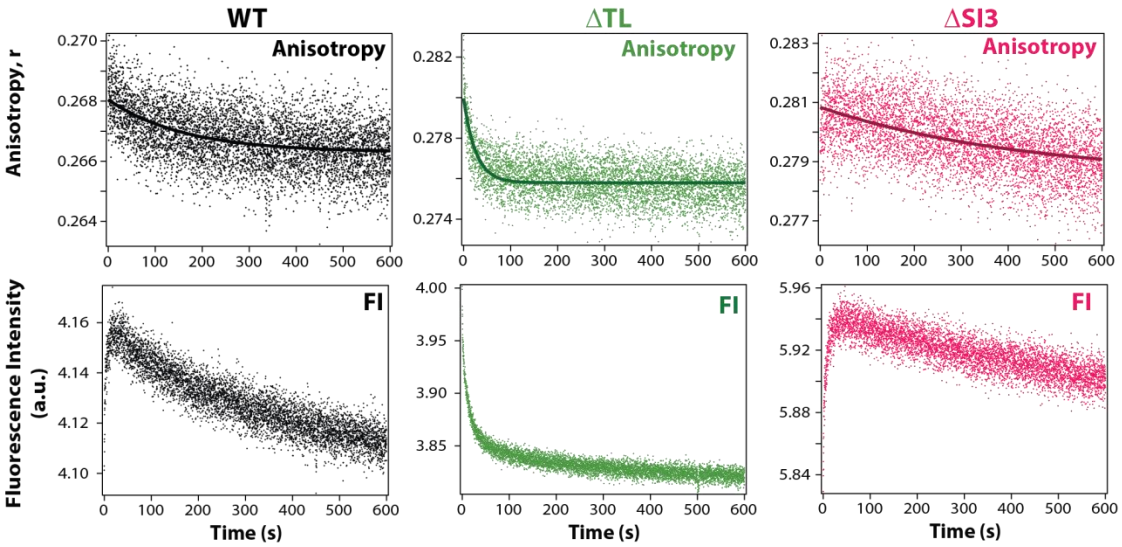
A.2.4 Dissociation assays do not detect significant effects of TL mutation

To assess the role of the TL on EC dissociation, I tested the effects of *(i)* a complete TL deletion (ΔTL RNAP), *(ii)* deletion of the sequence insertion in the TL (ΔSI3 RNAP), and *(iii)* RNAP mutants that stabilized either the folded (Ala_6 RNAP) or unfolded (LTPP RNAP) TL states on dissociation rate. In contrast to the ~75-fold defect in termination rate observed for ΔTL ECs by our termination assay **(Table 2.1A)**, there was minimal (1.4-2.3-fold) *stimulatory* effect of TL deletion on dissociation rate of DNA or RNA by both dissociation assays **(Figure A.2D-E)**. All other tested TL mutants similarly exhibited minimal differences in dissociation rate as measured by both dissociation assays, indicating that dissociation assays are not sensitive to mutations in the TL, and may not be representative of physiologically relevant effects on termination (see section **A.3 Discussion**). Interestingly, unlike all other RNAP variants, RNA and NT-DNA dissociated at similar rates for LTPP ECs, suggesting that LTPP ECs may have slightly weakened contacts to the DNA. However, this too needs to be confirmed by other methods.

A.2.5 Fluorescence intensity changes during dissociation suggest an alternative termination mechanism in TL deletion mutants

Upon inspecting the raw fluorescence intensity (FI) traces from which the fluorescence anisotropy is computed, we noticed an interesting peak in FI present early in the timecourse for WT and ΔSI3 ECs that was absent for ΔTL ECs **(Figure A.3)**. As Cy3 is an environmentally-sensitive fluorophore, the absence of a distinct peak for ΔTL ECs suggests that these complexes

may terminate by an alternative mechanism. However, I have not been able to reproduce this FI peak in subsequent tests. Thus, it will need to be determined whether this FI feature is representative of a true biological phenomenon, or an artifact.

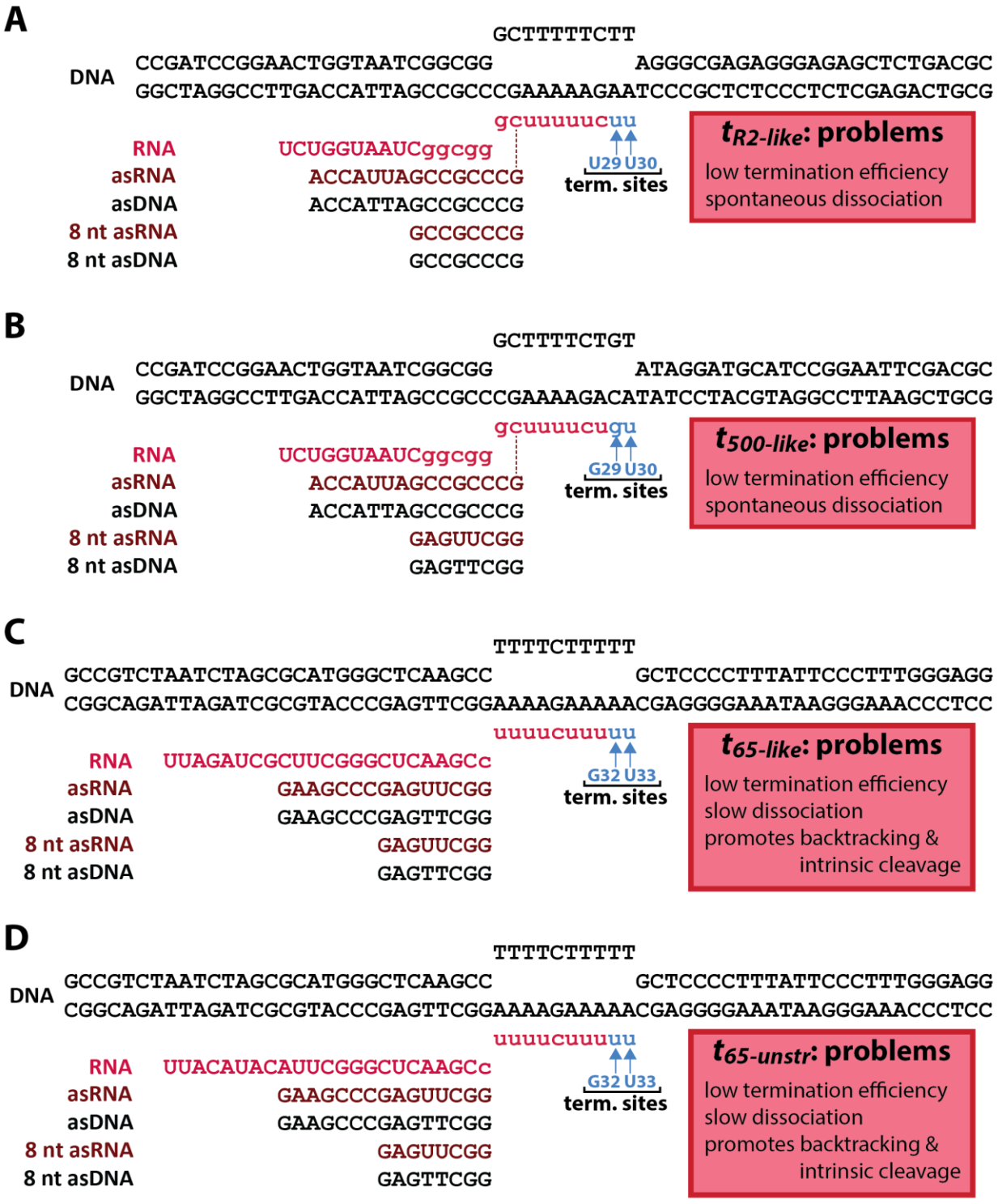


<u>Figure A.3:</u> Fluorescence intensity (FI) traces detect differences between ΔTL and WT/ ΔSI3 ECs.

Representative raw FI traces obtained from the Stopped-Flow apparatus (bottom) – from which anisotropy are computed (top) – are shown. As Cy3 is an environmentally-sensitive fluorophore, an early peak in FI seen for WT and ΔSI3 ECs is absent for ΔTL ECs suggests that ΔTL ECs may terminate by an alternative mechanism.

A.2.6 Challenges with previously tested scaffolds

Before testing the $t_{his2.dis}$ scaffold, I tested a number of other scaffold sequences that faced a subset of the following problems, making them unsuitable for termination assays, including: (*i*) spontaneous dissociation of the ECs in the absence of asRNA due to an unstable RNA–DNA hybrid; (*ii*) backtracking of ECs and intrinsic cleavage of the nascent RNA prior to addition of the antisense oligo; (*iii*) slow termination; *(iv)* low TE. These scaffolds and their associated problems are summarized in **Figure A.4**.; variants of these scaffolds were also tested (*not listed here*) that were not able to correct these problems. Other tested scaffolds are described in **Chapter 3**; of these $t_{R2.ex}$ also exhibited high levels of spontaneous dissociation **(Figures 3.5, 3.6 & 3.8)**. I also tested a range of different buffer conditions for these scaffolds, listed in **Table A.S1** (as well as Cl^- ion, Mg^{2+} ion, glycerol, and Ac-BSA titrations). However, no tested buffer condition could sufficiently correct the above issues to facilitate robust termination measurements from the previous scaffold designs.


A
GCTTTTTCTT
DNA
CCGATCCGGAACTGGTAATCGGCGG AGGGCGAGAGGGAGAGCTCTGACGC
GGCTAGGCCTTGACCATTAGCCGCCCGAAAAAGAATCCCGCTCTCCCTCTCGAGACTGCG
gcuuuuucuu
RNA UCUGGUAAUCggcgg
asRNA ACCAUUAGCCGCCCG
asDNA ACCATTAGCCGCCCG
8 nt asRNA GCCGCCCG
8 nt asDNA GCCGCCCG
U29 U30
term. sites
$t_{R2\text{-}like}$: problems
low termination efficiency
spontaneous dissociation
B
GCTTTTCTGT
DNA
CCGATCCGGAACTGGTAATCGGCGG ATAGGATGCATCCGGAATTCGACGC
GGCTAGGCCTTGACCATTAGCCGCCCGAAAAGACATATCCTACGTAGGCCTTAAGCTGCG
gcuuuucugu
RNA UCUGGUAAUCggcgg
asRNA ACCAUUAGCCGCCCG
asDNA ACCATTAGCCGCCCG
8 nt asRNA GAGUUCGG
8 nt asDNA GAGTTCGG
G29 U30
term. sites
$t_{500\text{-}like}$: problems
low termination efficiency
spontaneous dissociation
C
TTTTCTTTTT
DNA
GCCGTCTAATCTAGCGCATGGGCTCAAGCC GCTCCCCTTTATTCCCTTTGGGAGG
CGGCAGATTAGATCGCGTACCCGAGTTCGGAAAAGAAAAACGAGGGGAAATAAGGGAAACCCTCC
uuuucuuuuu
RNA UUAGAUCGCUUCGGGCUCAAGCc
asRNA GAAGCCCGAGUUCGG
asDNA GAAGCCCGAGTTCGG
8 nt asRNA GAGUUCGG
8 nt asDNA GAGTTCGG
G32 U33
term. sites
$t_{65\text{-}like}$: problems
low termination efficiency
slow dissociation
promotes backtracking &
intrinsic cleavage
D
TTTTCTTTTT
DNA
GCCGTCTAATCTAGCGCATGGGCTCAAGCC GCTCCCCTTTATTCCCTTTGGGAGG
CGGCAGATTAGATCGCGTACCCGAGTTCGGAAAAGAAAAACGAGGGGAAATAAGGGAAACCCTCC
uuuucuuuuu
RNA UUACAUACAUUCGGGCUCAAGCc
asRNA GAAGCCCGAGUUCGG
asDNA GAAGCCCGAGTTCGG
8 nt asRNA GAGUUCGG
8 nt asDNA GAGTTCGG
G32 U33
term. sites
$t_{65\text{-}unstr}$: problems
low termination efficiency
slow dissociation
promotes backtracking &
intrinsic cleavage

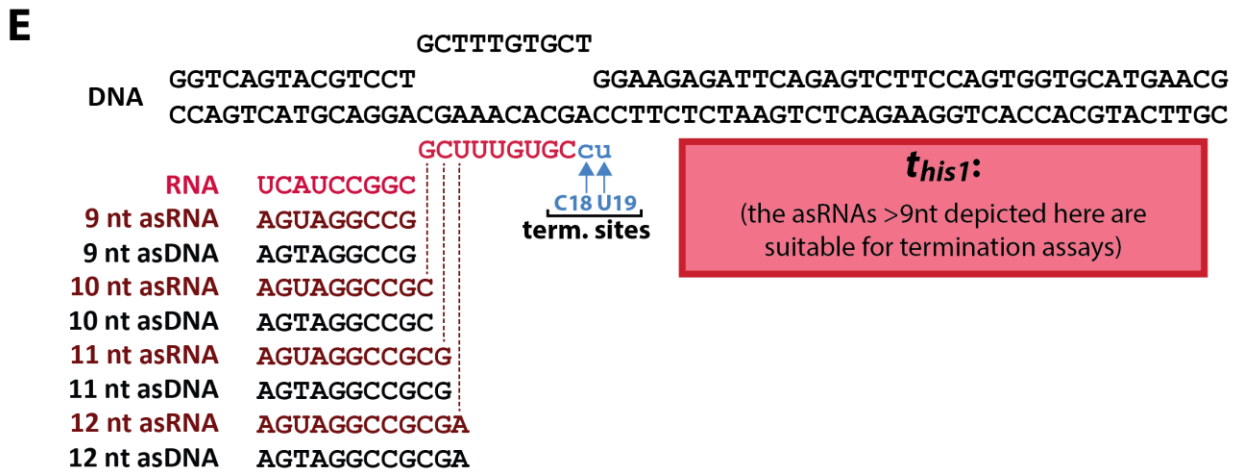


Figure A.4: Previously tested termination scaffolds and associated problems.

(A) $t_{R2\text{-}like}$, based on the t_{R2} terminator from λ phage, **(B)** $t_{500\text{-}like}$, based on the t_{500} terminator from λ phage, **(C)** $t_{65\text{-}like}$, designed previously (3) based on the t_{65} terminator from P23-45 phage, **(D)** $t_{65\text{-}unstr}$, based on the t_{65}-like scaffold sequence, with decreased RNA secondary structure formation predicted for the exiting RNA, **(E)** t_{his1}, based on the *his* pause sequence. All oligo sequences are listed 5' to 3' in **Table A.S3**.

A.3 Discussion

A.3.1 Termination efficiency and stalled EC dissociation rate measurements are not adequate to study the termination mechanism

The most commonly used method to study termination has historically been through termination efficiency measurements, which are subject to effects on both elongation and termination. However, we have shown that a comparison of the termination rates of the TL deletion mutant versus WT RNAP reveals a striking ~75-fold stimulatory role of the TL in termination – an effect that is completely masked by the ~10^4-fold effect of TL deletion on elongation (17) **(Figure 2.2D)**.

A couple of studies have attempted to counter this problem by examining dissociation of RNA from static complexes artificially stalled at a termination site (5, 7). However, the dissociation rates that arise from these experiments occur on the order of many minutes (~0.0015 s^{-1}) even with a full T_{hp} and at high salt concentrations, resulting in rates that cannot compete with physiologically relevant hairpin-stabilized pause escape rates (~0.1 s^{-1} for WT with –8asRNA; **Table 2.1B**), and are therefore not likely to be representative of a physiological termination mechanism. While we find that in our hands complexes stalled at U19 dissociate more quickly than the two previous studies (0.019 s^{-1} with –8asRNA; **Figure A.2D**), it is still ~5-fold slower than the rate at which actively transcribing ECs become transcriptionally inactive (*i.e.* functionally terminated) at the same site (0.0995 s^{-1}; **Table 2.1A**). Moreover, the dissociation rates from stalled ECs were also insensitive to TL mutations **(Figure A.2D-E)**. These findings strongly argue against the use of dissociation rates from stalled ECs as a metric for termination effects, but rather merit the study of termination occurring in competition with elongation as it occurs in the cell.

We note that we may be reporting the rate of a commitment step on the pathway to termination (18, 19) since our assay measures the rate of EC inactivation. Rigorous simultaneous measurement of dissociation rate and EC inactivation rate will be required to make this distinction, which is currently not possible for WT ECs given our assay setup. However, whether we are measuring a commitment step or the final step of termination, our method provides an easily adaptable assay platform sensitive to mutations that affect EC conformational fluctuations, indicating that it will be informative for characterizing the termination mechanism.

While existing termination assays have made great strides in elucidating the role of the terminator sequence elements and the steps of termination, our findings indicate that they are not well equipped to address questions about the termination mechanism. We suggest, instead, that methods that disentangle effects on termination from those on elongation – specifically from actively elongating complexes – are necessary to make strong arguments about factors that contribute specifically to termination. Single-molecule force-clamp studies have been successful in providing insight into the termination-specific role of the terminator U-tract and downstream DNA in directing hybrid-shearing versus hypertranslocation (19). Where such experiments are not possible, our termination assay setup described in **Chapter 2** provides an alternative method by which ensemble kinetics can be used to better understand the termination mechanism.

A.4 Materials & Methods

A.4.1 Materials

Non-fluorescent DNA and RNA oligonucleotides **(Table A.S3)** were obtained from Integrated DNA Technologies (IDT; Coralville, IA); the fluorescent Cy3-labeled oligonucleotide (8288) was obtained from TriLink BioTechnologies (San Diego, CA). All oligonucleotides were purified by 8 M urea denaturing polyacrylamide gel electrophoresis (PAGE) before use. [α-^{32}P]-CTP was obtained from PerkinElmer Life Sciences (Warwick, RI); NTPs were obtained from Promega (Madison, WI).

A.4.2 Proteins

His_{10}-tagged core wild-type and mutant *E. coli* RNAPs **(Table A.S2)** were purified from overexpression plasmids transformed in *E. coli* strain RL2657 as described previously (20).

A.4.3 Fluorescence anisotropy dissociation assay from stalled ECs

The nucleic acid scaffold was formed by mixing 7.5 μM C18 RNA and 5 μM T-DNA (8452 and 8451, respectively; **Table A.S3**) in reconstitution buffer (RB; 10 mM Tris-HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$) and heating to 95 °C for 2 min, cooling rapidly to 45 °C, then cooling to 25 °C in 2 °C increments for 2 min each, as described previously (21). ECs were assembled by incubating 2.5 μM core *E. coli* RNAPs with 1 μM nucleic acid scaffold in chloride transcription buffer (TB1.75; 20 mM Tris-Cl, pH 8.0, 75 mM NaCl, 1 mM $MgCl_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 2.5% glycerol, and 25 μg of acetylated bovine serum albumin/mL) for 15 min at 37 °C. Cy3-labeled NT-DNA (8288; **Table A.S3**) was then added at 0.5 μM and incubated for another 15 min at 37 °C. Assembled ECs were diluted 5-fold in TB1.75, and extended to U19 by incubation with 100 μM UTP at 37 °C for 5 min for WT/ΔSI3 ECs, or with 1 mM UTP at 37 °C

for 1 hr for ΔTL ECs. Reactions were supplemented at the time of NTP addition with 1 mM $MgCl_2$ to prevent Mg^{2+} depletion effects, and with 10 μM dsDNA competitor to prevent rebinding of RNAP to released NT-DNA.

ECs were then injected into one loading syringe of the stopped-flow apparatus fitted with anisotropy polarizers in the T-format (SF-300X; KinTek Corporation, Austin, TX), and 100 μM asRNA (−8asRNA; 8713, **Table A.S3**) in TB1.75 was loaded in the other syringe. Upon initiating rapid mixing at 37 °C, Cy3 fluorescence was excited at 540 nm (4.96 nm bandwidth), and emission was monitored in real time through a 550 nm long-pass filter (Edmund Optics Inc., Barrington, NJ). The dissociation rate of fast-tumbling Cy3-labeled NT-DNA (~45 kDa) from the slow-tumbling ECs (~395 kDa) was determined from n ≥3 replicates by fitting to a single exponential equation [1] and averaging rates (k_{obs}) obtained from all replicates:

$$r_t = r_0 + Ae^{-k_{obs}t} \quad [1]$$

where t = time (s); r_o = initial anisotropy; r_t = anisotropy at time t; A = amplitude of the fluorescence change; k_{obs} is the observed first-order rate constant for Cys-labeled NT-DNA dissociation.

A.4.4 Magnetic bead dissociation assay from stalled ECs

C18 RNA and NT-DNA (8452 and 8288, respectively; **Table A.S3**) were first [γ-^{32}P]-ATP-end-labeled by incubating 50 μM RNA/NT-DNA with~1 μM [γ-^{32}P]-ATP, 50 μM ATP, and T4 PNK (New England Biolabs, Ipswich, MA) for 30 min at 30 °C in T4 PNK buffer provided by the manufacturer. The nucleic acid scaffold was formed by mixing 5 μM [γ-^{32}P]-ATP-end-labeled C18 RNA and 10 μM T-DNA (8451; **Table A.S3**) in RB, annealing by heating to 95 °C and rapid incremental cooling to 25 °C, as described above. ECs were assembled by incubating 2.5 μM core

E. coli RNAPs with 0.5 μM nucleic acid scaffold in TB1.75 for 15 min at 37 °C. [γ-^{32}P]-ATP-end-labeled NT-DNA was then added at 1.5 μM and incubated for another 15 min at 37 °C.

Assembled ECs were diluted 5-fold in TB1.75 and tethered to paramagnetic Co^{2+}-beads (Dynabeads from Thermo Fisher Scientific Biosciences Inc.) through the His_{10}-tag on RNAP by incubating for 10 min at room temperature followed by 5 min at 37 °C. Beads were washed with TB1.75 to remove unbound RNAP and nucleic acids, and extended to U19 by incubation with 100 μM UTP at 37 °C for 5 min for WT/ΔSI3/Ala_6 ECs, or with 1 mM UTP at 37 °C for 1 hr for ΔTL/LTPP ECs. EC dissociation was initiated by adding an equal volume of 100 μM asRNA (−8asRNA; 8713, **Table A.S3**). Reactions were supplemented with 1 mM $MgCl_2$ at the time of NTP addition to prevent Mg^{2+} depletion effects, and with 5 μM dsDNA competitor (final concentration) at the time of asRNA addition to prevent rebinding of RNAP to released NT-DNA. Supernatant timepoints were taken by magnetic partitioning, and supernatant and whole reaction timepoints were stopped with an equal volume of 2x stop buffer. RNA products were separated by 8 M urea denaturing PAGE, and visualized by phosphorimaging, as described above.

Percent termination (PT%) was calculated for each timepoint as the percent of RNA or NT-DNA detected in the supernatant at each timepoint. The kinetics of RNA and NT-DNA dissociation were determined from n ≥3 replicates by fitting to a single exponential equation [2] and averaging rates (k_{obs}) obtained from all replicates:

$$PT_t = PT_0 + A(1 - e^{-k_{obs}t}) \quad [2]$$

where t = time (s); PT_o = initial PT%; PT_t = PT% at time t; A = amplitude of the fluorescence change; k_{obs} is the observed first-order rate constant for radiolabeled RNA or NT-DNA dissociation.

Table A.S1: Previously tested transcription buffers

Buffer Name	1x Buffer Components
TB1.75	20 mM Tris-Cl; 1 mM $MgCl_2$; 75 mM NaCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 2.5% glycerol; 100 µM/1 mM DTT
TB1.5.75	20 mM Tris-Cl; 1.5 mM $MgCl_2$; 75 mM NaCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 2.5% glycerol; 100 µM/1 mM DTT
TB2.75	20 mM Tris-Cl; 2 mM $MgCl_2$; 75 mM NaCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 2.5% glycerol; 100 µM/1 mM DTT
TB2.5.75	20 mM Tris-Cl; 2.5 mM $MgCl_2$; 75 mM NaCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 2.5% glycerol; 100 µM/1 mM DTT
TB5.75	20 mM Tris-Cl; 5 mM $MgCl_2$; 75 mM NaCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 2.5% glycerol; 100 µM/1 mM DTT
TB2.5	20 mM Tris-Cl; 2 mM $MgCl_2$; 50 mM KCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 2.5% glycerol; 100 µM/1 mM DTT
TBS	20 mM Tris-Cl; 4 mM $MgCl_2$; 250 mM NaCl; 100 µM EDTA; 50 µg Ac-BSA/mL; 100 µM/1 mM DTT
TBLM	20 mM Tris-Cl; 1 mM $MgCl_2$; 40 mM NaCl; 100 µM EDTA; 25 µg Ac-BSA/mL; 5% glycerol; 100 µM/1 mM DTT
BTB	20 mM Tris-Cl; 10 mM $MgCl_2$; 40 mM KCl; 100 µM/1 mM DTT
TB40	20 mM Tris-Cl; 10 mM $MgCl_2$; 40 mM NaCl; 100 µM EDTA; 2.5 µg Ac-BSA/mL; 100 µM/1 mM DTT

Table A.S2: RNAP variants and overexpression plasmids

Plasmid #	Plasmid Name	RNAP TL mutant	Description	Use
2956	pRM756	WT	Expresses wild-type *E. coli* RNAP (α2ββ′ω) with His_{10} tag on β′ C-terminus	(20)
2995	pRM795	WT	Expresses wild-type *E. coli* RNAP (α2ββ′ω) with His_{10} and HMK tags on β′ C-terminus	(22)
2978	pRM778	ΔTL	pRM756 derivative expressing mutant *E. coli* RNAP (α2ββ′ω) containing β' Δ931-1137 with SNAP-tag on β′ C-terminus	this work & (20)
2959	pRM759	ΔSI3	pRM756 derivative expressing mutant *E. coli* RNAP (α2ββ′ω) containing β' Δ943-1130	(20)
4288	pMTF025	Ala_6	pRM756 derivative expressing mutant *E. coli* RNAP (α2ββ′ω) containing β' Ala_6 (924A, 926A, 927A, 11336A, 1137A, 1139A)	(20)
2961	pRM761	LTPP	pRM756 derivative expressing mutant E. coli RNAP (α2ββ′ω) containing β' L930P, T931P	(20)

Table A.S3: Oligonucleotides used in this study.

Oligo #	RNA / DNA	Sequence (5' to 3')	Use
8288	DNA	GGTCAGTACGTCCTGCTTTGTGCTAAAAGAGATTCAG**AG**TCTTCCAGTGGTGCATGAACG	Cy3 labeled NT-DNA for $t_{his2.dis}$ reconstitution; Cy3 is doubly-tethered to the phosphate backbone between the 2 indicated nucleotides (9).
8451	DNA	CGTTCATGCACCACTGGAAGACTCTGAATCTCTTTTAGCACAAAGCAGGACGTACTGACC	T-DNA for $t_{his2.dis}$ reconstitution
8452	RNA	UCAUCCGGCGCUUUGUGC	RNA strand for $t_{his2.dis}$ & t_{his1} reconstitution
8713	RNA	AGCGCCGGAUGA	$t_{his2.dis}$ −8asRNA/ t_{his1} 12 nt asRNA
7516	DNA	CCGATCCGGAACTGGTAATCGGCGGGCTTTTTCTTAGGGCGAGAGGGAGAGCTCTGACGC	NT-DNA for $t_{R2\text{-}like}$ reconstitution
7517	DNA	GCGTCAGAGCTCTCCCTCTCGCCCTAAGAAAAAGCCCGCCGATTACCAGTTCCGGATCGG	T-DNA for $t_{R2\text{-}like}$ reconstitution
7518	DNA	CCGATCCGGAACTGGTAATCGGCGGGCTTTTCTGTAtAGGATGCATCCGGAATTCGACGC	NT-DNA for $t_{500\text{-}like}$ reconstitution
7519	DNA	GCGTCGAATTCCGGATGCATCCTATACAGAAAAGCCCGCCGATTACCAGTTCCGGATCGG	T-DNA for $t_{500\text{-}like}$ reconstitution
7522	RNA	UCUGGUAAUC	RNA strand for $t_{R2\text{-}like}$ & $t_{500\text{-}like}$ reconstitution
7520	DNA	GCCCGCCGATTACCA	$t_{R2\text{-}like}$ & $t_{500\text{-}like}$ asDNA
7524	RNA	GCCCGCCGAUUACCA	$t_{R2\text{-}like}$ & $t_{500\text{-}like}$ asRNA
8376	DNA	GCCCGCCG	$t_{R2\text{-}like}$ & $t_{500\text{-}like}$ 8 nt asDNA
8377	RNA	GCCCGCCG	$t_{R2\text{-}like}$ & $t_{500\text{-}like}$ 8 nt asRNA
7816	DNA	GCCGTCTAATCTAGCGCATGGGCTCAAGCCTTTTCTTTTTGCTCCCCTTTATTCCCTTTGGGAGG	NT-DNA for $t_{65\text{-}like}$ & $t_{65\text{-}unstr}$ reconstitution
7817	DNA	CCTCCCAAAGGGAATAAAGGGGAGCAAAAAGAAAAGGCTTGAGCCCATGCGCTAGATTAGACGGC	T-DNA for $t_{65\text{-}like}$ & $t_{65\text{-}unstr}$ reconstitution
7819	RNA	UUAGAUCGCUUCGGGCUCAAGC	RNA strand for $t_{65\text{-}like}$ reconstitution

Oligo #	RNA / DNA	Sequence (5' to 3')	Use
7820	DNA	GGCTTGAGCCCGAAG	$t_{65\text{-}like}$ & $t_{65\text{-}unstr}$ asDNA
8062	RNA	GGCUUGAGCCCGAAG	$t_{65\text{-}like}$ & $t_{65\text{-}unstr}$ asRNA
8209	DNA	GGCTTGAG	$t_{65\text{-}like}$ & $t_{65\text{-}unstr}$ 8nt asDNA
8210	RNA	GGCUUGAG	$t_{65\text{-}like}$ & $t_{65\text{-}unstr}$ 8nt asRNA
8286	RNA	UUACAUACAUUCGGGCUCAAGC	RNA strand for $t_{65\text{-}unstr}$ reconstitution
8448	DNA	GGTCAGTACGTCCTGCTTTGTGCTGGAAGAGATTCAGAGTCTTCCAGTGGTGCATGAACG	NT-DNA for t_{his1} reconstitution
8449	DNA	CGTTCATGCACCACTGGAAGACTCTGAATCTCTTCCAGCACAAAGCAGGACGTACTGACC	T-DNA for t_{his1} reconstitution
8453	DNA	GCCGGATGA	t_{his1} 9 nt asDNA
10427	RNA	GCCGGAUGA	t_{his1} 9 nt asRNA
8454	DNA	CGCCGGATGA	t_{his1} 10 nt asDNA
10428	RNA	CGCCGGAUGA	t_{his1} 10 nt asRNA
8455	DNA	GCGCCGGATGA	t_{his1} 11 nt asDNA
8714	RNA	GCGCCGGAUGA	t_{his1} 11 nt asRNA
8703	DNA	AGCGCCGGATGA	t_{his1} 12 nt asDNA

A.5 REFERENCES

1. Yarnell WS & Roberts JW (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* 284(5414):611-615.

2. Santangelo TJ & Roberts JW (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Molecular cell* 14(1):117-126.

3. Berdygulova Z*, et al.* (2012) A novel phage-encoded transcription antiterminator acts by suppressing bacterial RNA polymerase pausing. *Nucleic Acids Res* 40(9):4052-4063.

4. Shankar S, Hatoum A, & Roberts JW (2007) A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. *Molecular cell* 27(6):914-927.

5. Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S, & Nudler E (2007) An allosteric path to transcription termination. *Molecular cell* 28(6):991-1001.

6. Epshtein V, Dutta D, Wade J, & Nudler E (2010) An allosteric mechanism of Rho-dependent transcription termination. *Nature* 463(7278):245-249.

7. Komissarova N, Becker J, Solter S, Kireeva M, & Kashlev M (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Molecular cell* 10(5):1151-1162.

8. Artsimovitch I & 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(19):3110-3122.

9. Yager TD & von Hippel PH (1991) A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. *Biochemistry* 30(4):1097-1118.

10. Wilson KS & von Hippel PH (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A* 92(19):8793-8797.

11. Reynolds R, Bermudez-Cruz RM, & Chamberlin MJ (1992) Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators. *J Mol Biol* 224(1):31-51.

12. Lubkowska L, Maharjan AS, & Komissarova N (2011) RNA folding in transcription elongation complex: implication for transcription termination. *J Biol Chem* 286(36):31576-31585.

13. Chan CL, Wang D, & Landick R (1997) Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *J Mol Biol* 268(1):54-68.

14. Korzheva N*, et al.* (2000) A structural model of transcription elongation. *Science* 289(5479):619-625.

15. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, & Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448(7150):157-162.

16. Bar-Nahum G*, et al.* (2005) A ratchet mechanism of transcription elongation and its control. *Cell* 120(2):183-193.

17. Toulokhonov I, Zhang J, Palangat M, & Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Molecular cell* 27(3):406-419.

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

19. Larson MH, Greenleaf WJ, Landick R, & Block SM (2008) Applied force reveals mechanistic and energetic details of transcription termination. *Cell* 132(6):971-982.

20. Windgassen TA, *et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

21. Kyzer S, Ha KS, Landick R, & Palangat M (2007) Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J Biol Chem* 282(26):19020-19028.

22. Nayak D, Voss M, Windgassen T, Mooney RA, & Landick R (2013) Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Molecular cell* 50(6):882-893.

Appendix B:

Effects of microcin J25 on intrinsic termination

Chapter contributions:

A.R.S. performed all experiments, experimental design and analysis, and prepared all text/figures; A.R.S. and Alex Harwig purified Microcin J25.

B.1 INTRODUCTION

As we describe in **Chapter 2**, our measurements of termination rate for various RNA polymerase (RNAP) trigger loop (TL) mutants suggest that the TL requires conformational flexibility rather than a particular TL conformational state to stimulate the rate-limiting step of the termination mechanism. One prediction of this hypothesis is that constraining the TL to a single conformation (or a very limited set of conformations) would decrease termination. An endogenous bacterial transcriptional inhibitor, microcin J25 (MccJ25), has been shown to bind the TL and trap it in a folded conformation (Darst and co-workers, *unpublished findings*), and was thus a good candidate for testing our hypothesis.

MccJ25 is a 21-aa ribosomally-translated antimicrobial peptide produced and secreted under limiting nutrient conditions by Gram-negative bacteria containing the plasmid-encoded MccJ25 expression, maturation and export systems (1-3). Maturation of the peptide forms a lariat structure with the tail threaded through the lariat ring by formation of an internal lactam linkage (4), where distinct regions of the structure are required for the inhibitory activity and export of MccJ25 (5). Genetic and biochemical studies have demonstrated that sufficiently high (micromolar) concentrations of MccJ25 slow the production of abortive products during initiation, slow elongation, and increase pausing in an NTP concentration-dependent manner (5-7). Moreover, mutations in the TL and BH have been shown to confer resistance of RNAP to the inhibitory activity of MccJ25 (8). These findings combined with the crystal structure of MccJ25-bound RNAP from the Darst group thus provide strong evidence that MccJ25 limits TL-facilitated activities by trapping the TL in a folded conformation.

Nudler and co-workers had shown previously that a monoclonal antibody (mAb) that constrains the TL in an unfolded conformation decreases dissociation of elongation complexes (ECs) stalled at a termination site (9). It was this dissociation defect for ECs with an unfolded

TL, combined with observations of a crosslink formed between the TL and terminator hairpin (T_{hp}) that led Nudler and co-workers to suggest that a folded TL is required for termination. The use of MccJ25 to test effects on termination thus provides the opposite test, in which the TL is trapped in a folded conformation rather than an unfolded conformation.

B.2 Results & Discussion

B.2.1 Microcin J25 arrests U19 ECs but prevents RNA release

To test whether trapping the TL in a single (folded) conformation (or a very limited set of conformations) has an effect on termination rate, I added high concentrations of MccJ25 to my termination assay and determined its effects on the termination and RNA release efficiencies (TE and RE, respectively).

ECs were assembled at position G17 on the t_{his2} scaffold **(Figure B.1A)** with His_{10}-tagged RNAP and tethered through the His-tag to paramagnetic Co^{2+}-beads. The G17 ECs were radiolabeled at C18 by incorporation of [α-^{32}P]-CMP, and incubated with MccJ25 or an equivalent volume of water for ~2 min. ECs were then elongated through U19 in the absence or presence of asRNA. Samples of the whole reaction and supernatant (containing RNA/DNA released from RNAP) were taken after 32 min to determine the effects of MccJ25 on TE and RE. Freshly prepared MccJ25 was compared to a previous purification of MccJ25 to confirm activity of the new MccJ25 stock. The two MccJ25 preps behaved nearly identically **(Figure B.1)**.

A comparison of the percent occupancy of ECs at each RNA position (C18, U19 or ≥A20) at the end of the reaction **(Figure B.1B)**, and the TE calculated based on this EC occupancy **(Figure B.1C)** would suggest that the presence of MccJ25 *increase* termination at C18 and U19 in the presence and absence of asRNA. However, the relative REs in the presence and absence of

MccJ25 **(Figure B.1D)** indicate that MccJ25 actually *inhibits* release of ECs at U19 without asRNA and with –10asRNA, and thus increases arrest of these complexes at the termination sites without allowing complete termination. This observation is consistent with the folded TL-trapping activity of MccJ25, which would be expected to prevent translocation, and effectively arrest complexes. At the high concentrations used here, MccJ25 is likely always associated with the TL (despite its relatively low K_d for EC binding (6, 8)), resulting in a stronger arrest efficiency than observed previously (6). Notably, MccJ25 did not affect RE at C18, or at U19 with –8asRNA **(Figure B.1D)**, consistent with the synergistic effects of the TL and T_{hp} extension (described in **Chapter 2**) due to which ECs are likely able to terminate even in the absence of TL motion when T_{hp} extension is possible, but not when T_{hp} extension is limited. Interestingly, as a consequence of the MccJ25-mediated enrichment of ECs at the C18 and U19 positions, the overall productive TE is increased for both asRNAs despite the RE defect for U19 ECs **(Figure B.1E)**. The increase in productive TE is particularly pronounced for the –8asRNA, for which RE is not diminished by MccJ25 binding.

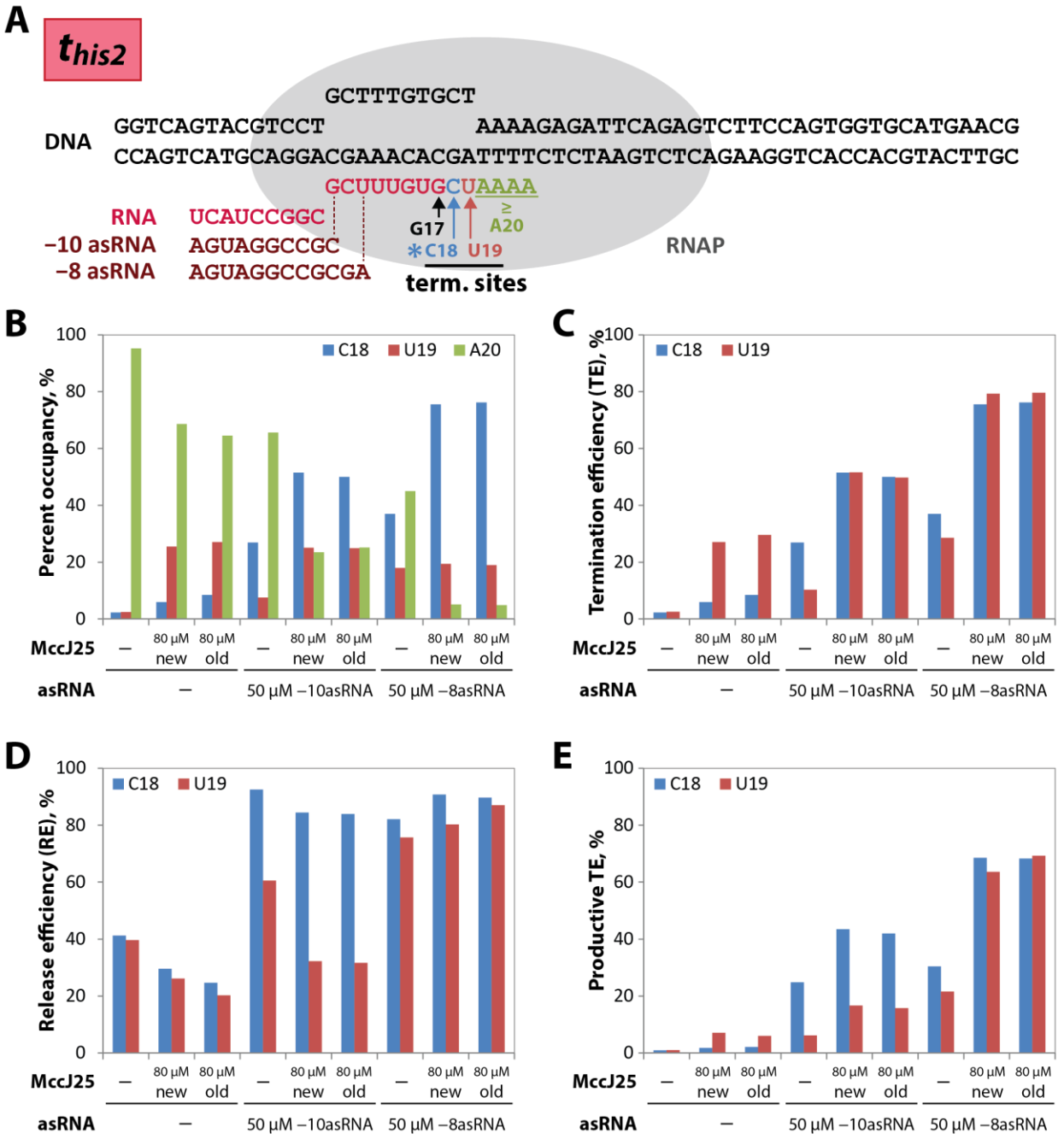


Figure B.1: MccJ25 arrests ECs at C18 & U19 and inhibits release from weak terminators.

(A) Scaffold t_{his2} depicted with two asRNA alternatives. **(B-E)** ECs were assembled with His-tagged WT RNAP at position G17, tethered to paramagnetic Co^{2+}-beads, and radiolabeled at C18 by the addition of [α-^{32}P]-CTP. C18 ECs were then incubated with ~80 µM new or old MccJ25 (see **B.3 Methods**) or an equivalent volume of water, and transcription restarted by the addition of 100 µM UTP and 10 µM ATP in the absence or presence of 50 µM −10asRNA or −8asRNA. Whole reaction (W; all RNA) and supernatant (S; released RNA) samples were taken after 32 min. The data were quantified in a few ways: **(B)** Percent occupancy, where occupancy at position n = band intensity(n) / (total band intensity) * 100;

(C) Termination efficiency (TE), where TE(C18) = C18 / (C18+U19+A20) * 100;
TE(U19) = U19 / (U19+A20) * 100; **(D)** Release efficiency (RE), where
RE at position n = S band intensity(n) / W band intensity(n) * 100; **(E)** Productive TE, where
productive TE at position n = TE(n) * (RE(n) / 100).

B.3 Materials & Methods

B.3.1 Materials

DNA and RNA oligonucleotides **(Table B.S2)** were obtained from Integrated DNA Technologies (IDT; Coralville, IA). All oligonucleotides were purified by 8 M urea denaturing polyacrylamide gel electrophoresis (PAGE) before use. [α-^{32}P]-CTP was obtained from PerkinElmer Life Sciences (Warwick, RI); NTPs were obtained from Promega (Madison, WI).

B.3.2 Proteins

His_{10}-tagged core wild-type *E. coli* RNAP **(Table B.S1)** was purified from an overexpression plasmid transformed in *E. coli* strain RL2657 as described previously (10).

Microcin J25 (MccJ25) was purified from the pJP73 overexpression plasmid, a kind gift from A. James Link (Princeton University, NJ) by a protocol adapted from (11). Briefly, DH10B cells containing the pJP73 plasmid were grown up in LB overnight at 37 °C, shaking at 250 rpm. Cultures were diluted to a starting OD_{600} of ~0.02 in M9 salts supplemented with thiamine, glucose and all 20 amino acids, and grown at 37 °C with shaking at 250 rpm up to a final OD_{600} of 0.5. The mcjABCD operon was then induced with 1 mM IPTG and grown overnight for ~20 hrs at 37 °C, shaking at 250 rpm. Cells were then pelleted, and MccJ25-containing supernatant collected. MccJ25 was purified from the supernatant on BondElut C8 cartridges (from Varian, now acquired by Agilent Technologies, Santa Clara, CA) following manufacturer instructions. The elutions were dried by rotovap, followed by speedvac, and stored as both a powder, and a solution in water. Purity was checked by resolving products on a homogenous 20% denaturing PhastGel (GE Healthcare), and Coomassie brilliant blue R-250 staining; MccJ25 was the only detectable product.

B.3.3 *In vitro* EC reconstitution

The nucleic acid scaffold for EC reconstitution was formed by mixing 5 μM G17 RNA and 10 μM T-DNA (10002 and 8451, respectively; **Table B.S2**) in reconstitution buffer (RB; 10 mM Tris-HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$) and heating to 95 °C for 2 min, cooling rapidly to 45 °C, then cooling to 25 °C in 2 °C increments for 2 min each, as described previously (12). The 17 nt RNA was designed to have 8 nt complementarity to the T-DNA at the site of reconstitution to prevent base pairing of upstream T-DNA and RNA, further preventing backtracking off of the termination site.

ECs were reconstituted by incubating 2.5 μM core *E. coli* RNAPs with 0.5 μM nucleic acid scaffold in transcription buffer (TB; 20 mM Tris-OAc, pH 8.0, 75 mM NaOAc, 1 mM $Mg(OAc)_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 2.5% glycerol, and 25 μg of acetylated bovine serum albumin/mL) for 15 min at 37 °C. NT-DNA (8450; **Table B.S2**) was then added at 1.5 μM and incubated for another 15 min at 37 °C. Fully complementary NT-DNA was used to allow the energy of DNA reannealing upon termination and bubble collapse to contribute to the termination energetics.

B.3.4 *In vitro* transcription

ECs were diluted 5-fold in TB to 100 nM ECs and tethered to paramagnetic Co^{2+}-beads (Dynabeads from Thermo Fisher Scientific Biosciences Inc.) through the His_{10}-tag on RNAP by incubating for 10 min at room temperature followed by 10 min at 37 °C. Bead-bound ECs were washed with TB to remove unbound RNAP and nucleic acids, and radiolabeled by incorporation labeling with 0.1 μM [α-^{32}P]-CTP and 0.9 μM CTP for 5 min to extend ECs to C18. C18 ECs were incubated with ~80 μM MccJ25 prepped in this work (new) or previously (old), or with an equivalent volume of water for ~2 min at 37 °C. Transcription was restarted by addition of diluted ECs to an equal volume of TB containing 100 μM UTP and 10 μM ATP, in the presence

or absence of 50 μM asRNA. The NTP mix was supplemented with 110 μM $Mg(OAc)_2$ to prevent Mg^{2+} depletion effects. Supernatant timepoints were taken by magnetic partitioning, and supernatant and whole reaction timepoints were stopped with an equal volume of 2x stop buffer (10 M urea, 50 mM EDTA, 90 mM Tris-borate buffer, pH 8.0, 0.02% bromophenol blue and 0.02% xylene cyanol). Timepoints were heated to 95 °C for 5 min to denature RNAP, and RNA products resolved by 8 M urea denaturing PAGE. Gels were exposed to phosphorimager screens, scanned using the Typhoon PhosphorImager and quantified using ImageQuant software (GE Healthcare).

Table B.S1: RNAP variant and overexpression plasmid

Plasmid #	Plasmid Name	RNAP variant	Description	Use
2956	pRM756	WT	Expresses wild-type *E. coli* RNAP (α2ββ´ω) with His_{10} tag on β´ C-terminus	(10)

Table B.S2: Oligonucleotides used in this study.

Oligo #	RNA / DNA	Sequence (5' to 3')	Use
8450	DNA	GGTCAGTACGTCCTGCTTTGTGCTAAAAGAGATTCAGAGTCTTCCAGTGGTGCATGAACG	NT-strand for t_{his2} reconstitution
8451	DNA	CGTTCATGCACCACTGGAAGACTCTGAATCTCTTTTAGCACAAAGCAGGACGTACTGACC	T-strand for t_{his2} reconstitution
10002	RNA	UCAUCCGGCGCUUUGUG	G17 RNA strand for t_{his2} reconstitution
8713	RNA	AGCGCCGGAUGA	−8asRNA
10428	RNA	CGCCGGAUGA	−10asRNA

B.4 REFERENCES

1. Salomon RA & Farias RN (1992) Microcin 25, a novel antimicrobial peptide produced by Escherichia coli. *J Bacteriol* 174(22):7428-7435.

2. Solbiati JO, Ciaccio M, Farias RN, & Salomon RA (1996) Genetic analysis of plasmid determinants for microcin J25 production and immunity. *J Bacteriol* 178(12):3661-3663.

3. Chiuchiolo MJ, Delgado MA, Farias RN, & Salomon RA (2001) Growth-phase-dependent expression of the cyclopeptide antibiotic microcin J25. *J Bacteriol* 183(5):1755-1764.

4. Wilson KA*, et al.* (2003) Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. *J Am Chem Soc* 125(41):12475-12483.

5. Semenova E, Yuzenkova Y, Peduzzi J, Rebuffat S, & Severinov K (2005) Structure-activity analysis of microcinJ25: distinct parts of the threaded lasso molecule are responsible for interaction with bacterial RNA polymerase. *J Bacteriol* 187(11):3859-3863.

6. Mukhopadhyay J, Sineva E, Knight J, Levy RM, & Ebright RH (2004) Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. *Mol Cell* 14(6):739-751.

7. Pavlova O, Mukhopadhyay J, Sineva E, Ebright RH, & Severinov K (2008) Systematic structure-activity analysis of microcin J25. *J Biol Chem* 283(37):25589-25595.

8. Yuzenkova J*, et al.* (2002) Mutations of bacterial RNA polymerase leading to resistance to microcin j25. *J Biol Chem* 277(52):50867-50875.

9. Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S, & Nudler E (2007) An allosteric path to transcription termination. *Mol Cell* 28(6):991-1001.

10. Windgassen TA*, et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

11. Pan SJ, Cheung WL, & Link AJ (2010) Engineered gene clusters for the production of the antimicrobial peptide microcin J25. *Protein Expr Purif* 71(2):200-206.

12. Kyzer S, Ha KS, Landick R, & Palangat M (2007) Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J Biol Chem* 282(26):19020-19028.

Appendix C:

Potential calculations for genome-scale analysis of NusA effects on intrinsic termination in *Bacillus subtilis*

For alternative analysis of RNA-seq data collected in:

Smarajit Mondal, Alexander V. Yakhnin, Aswathy Sebastian, Istvan Albert, and Paul Babitzke. (2016) NusA-dependent transcription termination prevents misregulation of global gene expression. *Nature Microbiology*. Article number: 15007; doi:10.1038/nmicrobiol.2015.7

Chapter contributions:

A.R.S. designed all planned experiments.

C.1 TERMINATION EFFICIENCY, NUSA EFFECT & PROMOTER STRENGTH CALCULATIONS

Define terminator.

Typically a 2-3nt region; most frequently spans 2nt. (Mondal *et al.* used a 3nt region.) Terminators assignments defined in the Babitzke paper (called POT, or point of termination) will be used. (See NOTE 1 below for verification method.) We will identify the terminator in a similar way, but we are also interested to know the TE at each of these positions, and relate them to terminator properties and NusA effect.

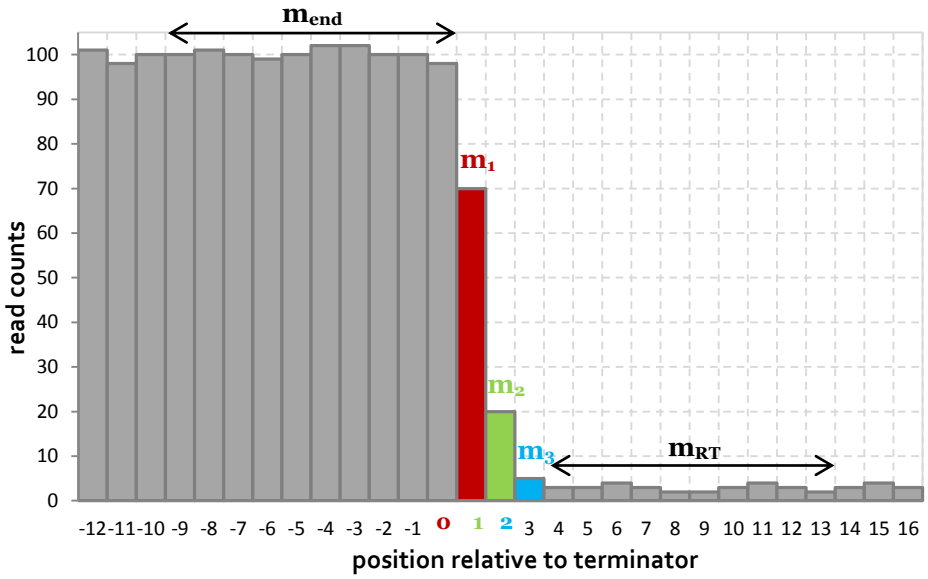
Set "m_{end}" as mean read counts for a window 1nt upstream of terminator.

Set window size. Mondal *et al.* used a 10nt window. We will currently use the same.

Set "m_1", "m_2", and "m_3" as the read counts at each position following the 3 terminator positions 0, 1 and 2. *See diagram below.*

Set "m_{RT}" as mean read counts for a window 1nt downstream of terminator.

Mondal *et al.* used a 10nt window. We will currently use the same.



Set "TE" as the overall termination efficiency, where

$TE = (m_{end} - m_{RT}) / m_{end}$

TE ranges from 0 to 1.

TE will also be computed at each termination position 0, 1 & 2.

(Mechanistically relevant: *to determine the strength of each termination position*.)

$TE_0 = (m_{end} - m_1) / m_{end}$; $TE_1 = (m_1 - m_2) / m_1$; $TE_2 = (m_2 - m_3) / m_2$

(Biologically relevant: *to determine the proportion of complexes terminated at each position*.)

$fTE_0 = TE_0 = (m_{end} - m_1) / m_{end}$; $fTE_1 = (m_1 - m_2) / m_{end}$; $fTE_2 = (m_2 - m_3) / m_{end}$

NOTE 1: The 2 equations (1) $TE \approx fTE_0 + fTE_1 + fTE_2$, and (2) $m_3 \sim m_{RT}$, should be true if the positioning of the 3nt window for the terminator has been correctly assigned.

Assuming the simplest model, where elongation and termination compete directly, **set "k_T" and "k_E" as rate constants**,

where k_E is set to be an arbitrary constant, since it will not affect calculations of NusA effect (see below).

$$term.EC_n \xleftarrow{k_T} EC_n \xrightarrow{k_E} EC_{n+1}$$

Here, TE can also be represented as:

$TE = k_T / (k_T + k_E)$

or, $k_T = (TE * k_E) / (1 - TE)$

NOTE 2: Here, k_E was set to be 25 s^{-1}; this constant gets canceled out in the δ_{NusA} calculation, below.

Set "δ_{NusA}" as fold effect of NusA addition on rate of termination, ie k_T(+NusA) versus k_T(−NusA).

$\delta_{NusA} = k_T(+NusA) / k_T(-NusA)$

δ_{NusA} ranges from 0 to ∞.

Set "m_{TU}" as mean read counts across the transcription unit.

As a readout of expression level of the transcription unit. Only for genes with identified intrinsic terminators.

C.2 SORTING TERMINATORS INTO CLASSES

NOTE 3: For any classes with division into sub-categories, each analysis should be done both with the entire class, as well as with the sub-categories.

NusA-affected.

$\delta_{NusA} > 1$

Divide into *(1) weak, (2) moderate, and (3) strong NusA effect* based on δ_{NusA}.

Classify terminators that are **NusA-dependent.**

$\delta_{NusA} > 1$ & TE(–NusA) ~ 0.

Also divide into *(1) weak, (2) moderate, and (3) strong terminators in the absence of NusA* based on TE(–NusA).

Will first look at distributions, and determine how to classify these ranges

(e.g. 1 standard deviation vs. % cutoff vs. top 100).

NusA-independent.

$\delta_{NusA} \sim 1$

Divide into *(1) weak, (2) moderate, and (3) strong terminators in the absence of NusA* based on TE(–NusA).

Will first look at distributions, and determine how to classify these ranges

(e.g. 1 standard deviation vs. % cutoff vs. top 100).

NusA-inhibited.

$\delta_{NusA} < 1$

NOTE 4: will need to decide on a reasonable cutoff to allow for small deviations in TE that are NusA-independent, but result in a $\delta_{NusA} < 1$.

Terminators with 1 or 2 nt deviations from a perfect U-tract.

Divide into those that have an interruption *(1) in first 3 nt of U-tract, or (2) after first 3 nt of U-tract*.

Terminators without 3 consecutive G–C bps in the base of the stem.

High/low gene expression.

Divide into **(1) low, (2) moderate, and (3) high expression +/–NusA** based on m_{TU}(+NusA) and m_{TU}(–NusA).

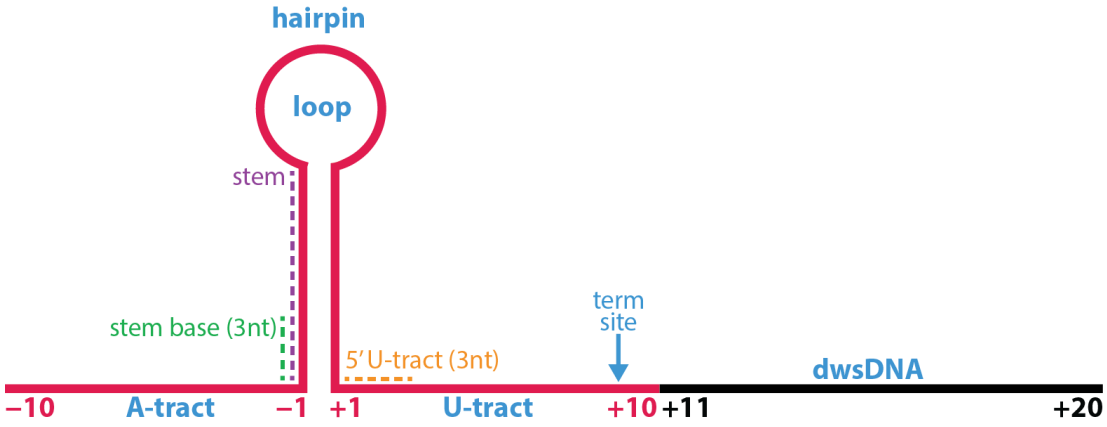
Will first look at distributions, and determine how to classify these ranges

(e.g. 1 standard deviation vs. % cutoff vs. top 100).

Next compare expression level classes with terminator classifications with respect to NusA effect.

C.3 Analyses & equations

The terminators can be split into a few different segments, highlighted below.



All analyses will need to be done for each class of terminators (outlined above). Comparison of the different properties of each terminator element for these classes (and their sub-categories) may help identify the differences between terminators that are aided by NusA versus those that are not.

U-tract.

(1) Number of nucleotides in the U-tract: $\mathbf{n_U}$

(2) Probability of U at each position of the U-tract: $\mathbf{P(U_n)} = \frac{\#\text{ terminators with U at position } n}{\text{total \# terminators in category}}$

NOTE 5: can also be used to determine the probability of non-Us at each position of the U-tract ($\mathbf{1-P(U_{+1})}$) for those terminators that have 1 or 2 nt interruptions in a perfect U-tract.

(3) Binding energy of U-tract and template DNA: $\mathbf{\Delta G_U} = \Delta G_{\text{R:D initiation}} + \sum_{i=1}^{nu-1} \Delta G_{\text{R:D}}(n_i, n_{i+1})$**

(4) Binding energy of 3 nt at 5′ end of U-tract: $\mathbf{\Delta G_{5'U}} = \Delta G_{\text{R:D initiation}} + \sum_{i=1}^{nu-1} \Delta G_{\text{R:D}}(n_i, n_{i+1})$**

** using RNA–DNA hybridization and stacking energies from Sugimoto_Sasaki 1995 *Biochem* (see value table in section **C.4 RNA–DNA Hybridization & Stacking Energies**), OR using Vienna RNA/DNA preliminary folding package

Hairpin stem.

(1) Number of paired nucleotides in left arm of stem: $\mathbf{n_s}$

(2) Probability of G–C bp at each position in left arm of stem: $\mathbf{P(G{-}C_n)} = \frac{\text{\# terminators with G or C at position } n}{\text{total \# terminators in category}}$

(3) Folding energy of base of hairpin: $\mathbf{\Delta G_{HB}}$ = (use ViennaRNA folding package)

Hairpin loop.

(1) Number of nucleotides in hairpin loop: $\mathbf{n_L}$

(2) Folding energy of loop closure: $\mathbf{\Delta G_L}$ = (use ViennaRNA folding package)

Hairpin.

(1) Folding energy of the terminator hairpin: $\mathbf{\Delta G_H}$ = (use ViennaRNA folding package)

(2) Height of hairpin: $\mathbf{n_H} = n_S + f(n_L)$

A-tract.

(1) Probability of A at each position of the A-tract: $\mathbf{P(A_n)} = \frac{\text{\# terminators with A at position } n}{\text{total \# terminators in category}}$

(2) Folding energy of A-tract-extended hairpin: $\mathbf{\Delta G_{HA}}$ = (use ViennaRNA folding package)

(3) Folding energy of A-tract: $\mathbf{\Delta G_A} = \mathbf{\Delta G_{HA}} - \mathbf{\Delta G_H}$

Downstream DNA (dwsDNA).

(1) Probability of G–C bp at each position in dwsDNA:

$$\mathbf{P(dwsG{-}C_n)} = \frac{\text{\# terminators with G or C at position } n}{\text{total \# terminators in category}}$$

Competing secondary structural effects.

(1) Use KineFold to simulate kinetic folding of RNA as it is transcribed by RNAP, **OR**

(2) look for binding potential of the upstream arm of the terminator hairpin with different windows of RNA upstream of the terminator.

NOTE:

Also use the below reference to inform plans for analyses:

Chen, Y J *et al.* (1995) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature Methods*. **10**(7)**:**659-664.

C.4 RNA–DNA Hybridization & Stacking Energies

** Thermodynamic parameters for hybrid duplex initiation and propagation in 1 M NaCl buffer			
RNA Sequence	**ΔH°, kcal mol^{-1}**	**ΔS°, kcal mol^{-1} K^{-1}**	**ΔG$_{37}$°, kcal mol^{-1}**
AA	−7.8	−21.9	−1.0
AC	−5.9	−12.3	−2.1
AG	−9.1	−23.5	−1.8
AU	−8.3	−23.9	−0.9
CA	−9.0	−26.1	−0.9
CC	−9.3	−23.2	−2.1
CG	−16.3	−47.1	−1.7
CU	−7.0	−19.7	−0.9
GA	−5.5	−13.5	−1.3
GC	−8.0	−17.1	−2.7
GG	−12.8	−31.9	−2.9
GU	−7.8	−21.6	−1.1
UA	−7.8	−23.2	−0.6
UC	−8.6	−22.9	−1.5
UG	−10.4	−28.4	−1.6
UU	−11.5	−36.4	−0.2
initiation	1.9	−3.9	3.1
Estimated errors in ΔH°, ΔS°, and ΔG$_{37}$° are ± 0.3 kcal mol^{-1}, ± 1.3 kcal mol^{-1} K^{-1}, and ± 0.1 kcal mol^{-1}, respectively.			

Above table adapted from Table 3 of:

Sugimoto, N. *et al.* (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry*. **34:**11211-11216.

Appendix D:

Contributions of the trigger loop to CBR antimicrobial-mediated transcription inhibition

Adapted from:

Brian Bae, Dhananjaya Nayak, Ananya Ray, Arkady Mustaev, Robert Landick, and Seth A. Darst. (2015) CBR antimicrobials inhibit RNA polymerase via at least two bridge-helix cap-mediated effects on nucleotide addition. *PNAS*. **112**(31): E4178-E4187.

A.R.S.' contribution:

Experiments to determine the effect of CBR9379 on transcription elongation by WT and ΔTL RNAP, Figure D.3B and Figure D.4A

D.1 Abstract

RNA polymerase inhibitors like the CBR class that target the enzyme's complex catalytic center are attractive leads for new antimicrobials. Catalysis by RNA polymerase involves multiple rearrangements of bridge helix, trigger loop, and active-center side chains that isomerize the triphosphate of bound NTP and two Mg^{2+} ions from a preinsertion state to a reactive configuration. CBR inhibitors target a crevice between the N-terminal portion of the bridge helix and a surrounding cap region within which the bridge helix is thought to rearrange during the nucleotide addition cycle. We report crystal structures of CBR inhibitor/ *Escherichia coli* RNA polymerase complexes as well as biochemical tests that establish two distinct effects of the inhibitors on the RNA polymerase catalytic site. One effect involves inhibition of trigger-loop folding via the F loop in the cap, which affects both nucleotide addition and hydrolysis of 3′-terminal dinucleotides in certain backtracked complexes. The second effect is trigger-loop independent, affects only nucleotide addition and pyrophosphorolysis, and may involve inhibition of bridge-helix movements that facilitate reactive triphosphate alignment.

D.2 Introduction

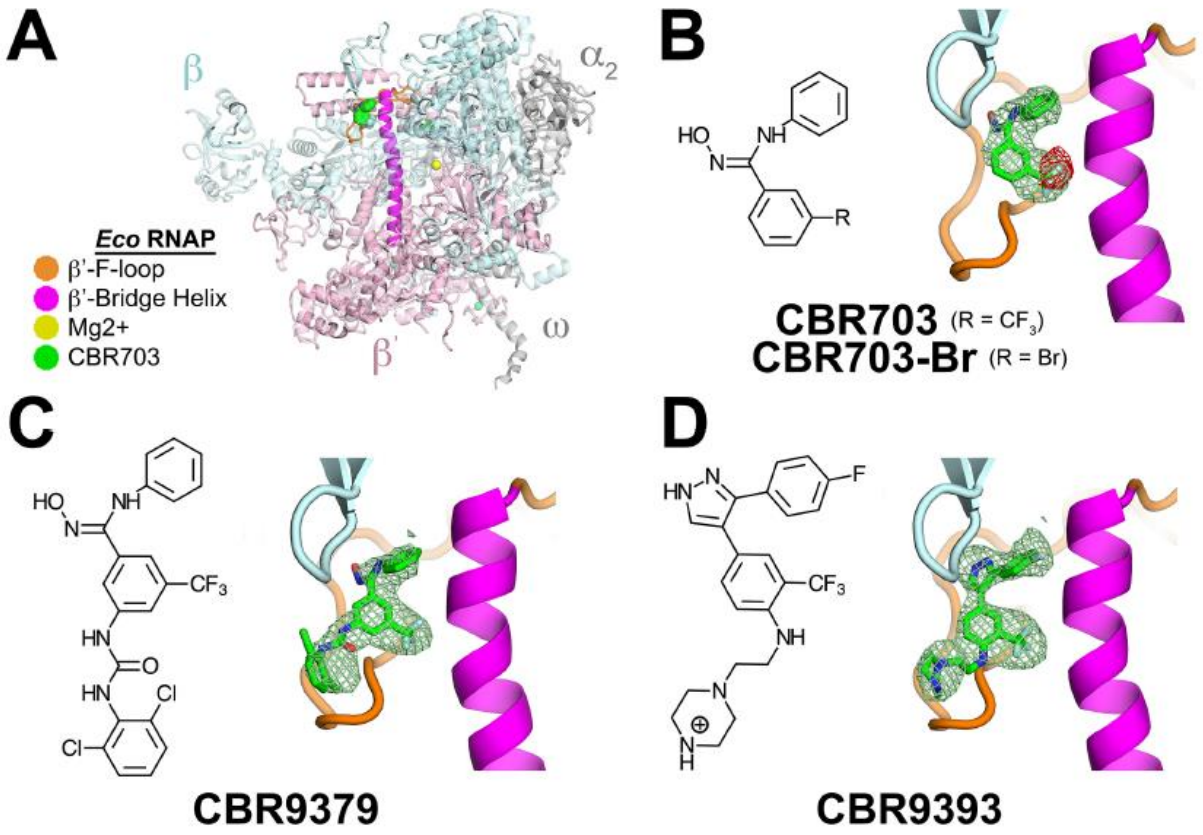
Bacteria harbor a single RNA polymerase (RNAP) enzyme (subunit composition α2ββ′ω, ~400 kDa) that performs all transcription in the cell. The bacterial RNAP is a proven target for antimicrobials. The rifamycins (Rifs), which bind to and inhibit the bacterial RNAP (1-3), are potent, broad-spectrum antimicrobials and are the lynchpin of current tuberculosis therapy (4). Nevertheless, the emergence of multidrug-resistant pathogens highlights the importance of discovering novel antimicrobials (5). New RNAP inhibitors can also be used as tools to probe transcriptional mechanisms.

Artsimovitch et al. (6) described a new class of bacterial RNAP inhibitors, the CBR compounds. Single amino acid substitutions in the RNAP β and β′ subunits that conferred resistance to CBR compounds (CBR^R) defined a pocket near the RNAP surface bounded by β′ residues of the F loop and the N-terminal part of the bridge helix (BH), and β residues neighboring fork-loop 2 (FL2) and the βDII loop, all RNAP structural elements shown to play important roles in the enzyme's nucleotide addition cycle (NAC) (7-20). The genetically defined CBR site is ~31 Å from the RNAP active site Mg^{2+} ion and is also distinct from the binding sites of other well-characterized bacterial RNAP inhibitors. Inhibition of RNAP by the CBR compounds appears to be mechanistically distinct as well; the CBR compounds allosterically inhibit known catalytic activities of the RNAP active site preferentially at pause sites but have lesser to no effects on translocation (6, 21).

The prototype CBR compound [3-(trifluoromethyl)-N-phenylbenzamidoxime, designated CBR703] **(Figure D.1 A-B)**, identified by screening a library of compounds for inhibition of *Escherichia coli* RNAP, yielded 50% inhibitory concentrations (IC50s) in the 10–20 μM range in in vitro assays. More potent variants, CBR9379 (IC_{50} of 0.3 μM) **(Figure D.1C)** and CBR9393 (IC_{50} of 2.5 μM) **(Figure D.1D)** were also synthesized and characterized (6). In addition to inhibiting RNAP from Gram-negative *E. coli*, the CBR compounds inhibit the RNAP from a Gram-positive organism (*Staphylococcus aureus*) (22, 23). The CBR compounds do not significantly inhibit human RNAP II (IC_{50} > 200 μM), *Thermus aquaticus*, or *Thermus thermophilus* RNAPs. Further characterization of the CBR compound inhibition mechanism and improvement of CBR compound potency has been hampered by a lack of structural information on CBR/RNAP complexes.

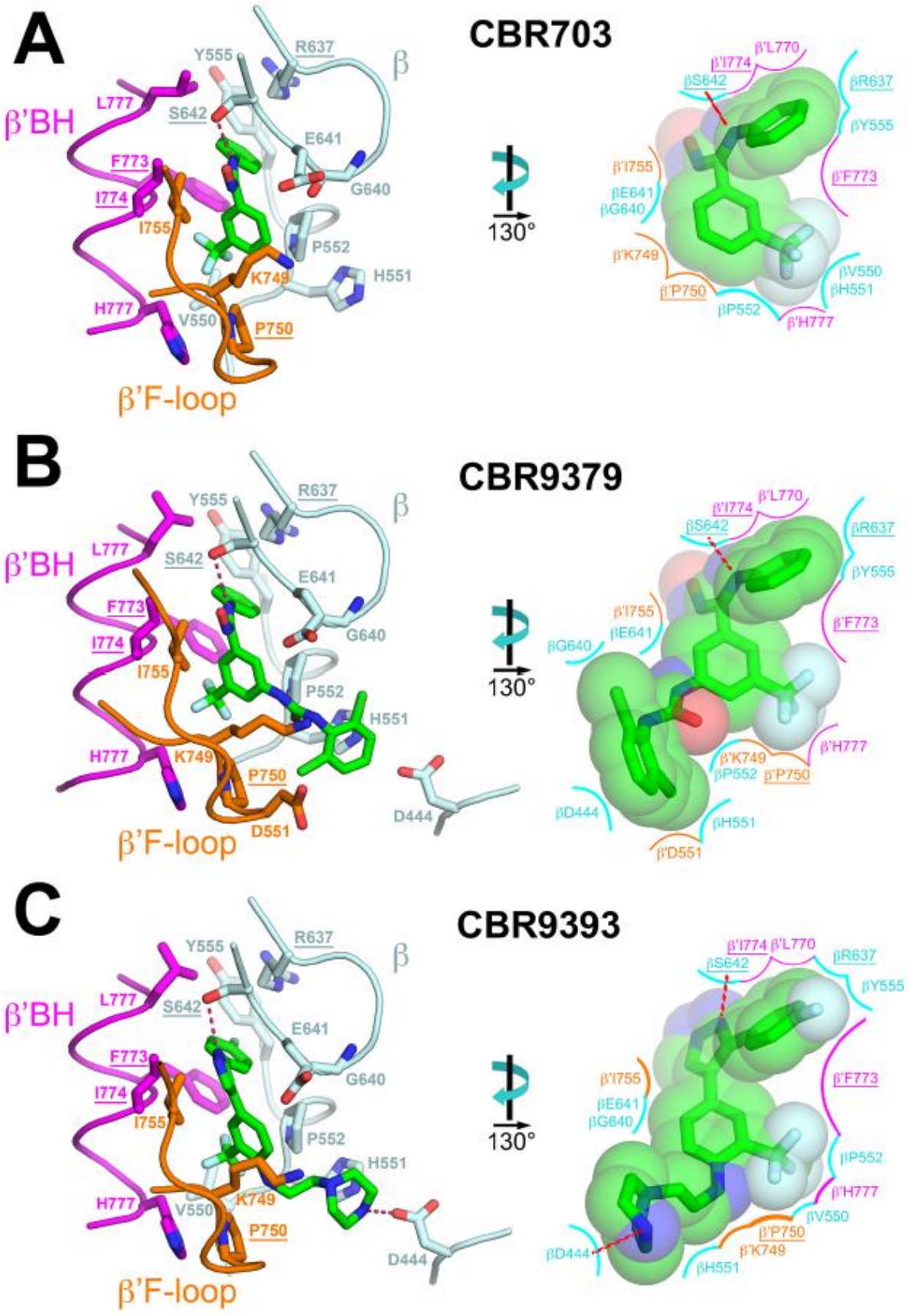
In this work, we determine crystal structures that detail CBR/*E. coli* RNAP interactions. The structures confirm the CBR binding site defined by CBR^R mutants, provide a structural framework to investigate details of the inhibition mechanism, and set the stage for structure-

based design of CBR analogs with improved properties. CBR effects on the reactions of nucleotidyl transfer and hydrolytic transcript cleavage by the RNAP catalytic site provide insights into the enzyme's catalytic mechanism.



<u>Figure D.1:</u> Crystal structures of *E. coli* RNAP–holoenzyme with CBR compounds.

(A) Structure of *E. coli* RNAP–holoenzyme in complex with CBR703. The *E. coli* RNAP core is shown as an *α-carbon backbone ribbon*: α and ω subunits, *gray*; β, *light cyan*; β′, *light pink* except the F loop is shown in *orange* and the BH is shown in *magenta*. The active-site Mg^{2+} is shown as a *yellow sphere*. CBR703 is shown with *green carbon atoms*. Some obscuring parts of β are rendered transparent. The σ^{70} subunit (present in the structure) is also not shown for clarity. **(B)** (Left) Chemical structure of CBR703 (6) and CBR703-Br. (Right) Zoom-in of CBR binding pocket between elements of β (*light cyan*), the β′-F loop (*orange*), and the β′-BH (*magenta*). The atomic model for CBR703 is shown (*green carbon atoms*) along with the unbiased $F_o - F_c$ map (*green mesh*, contoured at 3σ) and the Br anomalous Fourier difference peak (*red mesh*, contoured at 4.5σ). **(C)** (Left) Chemical structure of CBR9379 (6). (Right) Zoom-in of CBR binding pocket. The atomic model for CBR9379 is shown (*green carbon atoms*) along with the unbiased $F_o - F_c$ map (*green mesh*, contoured at 3σ). CBR9379S is a more easily synthesized, close analog of CBR9379 that was used in some experiments but not for crystal structure determination. **(D)** (Left) Chemical structure of CBR9393 (6). (Right) Zoom-in of CBR binding pocket. The atomic model for CBR9393 is shown (*green carbon atoms*) along with the unbiased $F_o - F_c$ map (*green mesh*, contoured at 3σ).


A
CBR703
β'BH
β'F-loop
Y555
R637
β
L777
S642
E641
F773
G640
I774
I755
P552
K749
H551
H777
V550
P750
130°
β'I774
β'L770
βS642
βR637
βY555
β'I755
βE641
βG640
β'F773
β'K749
βV550
β'P750
βH551
βP552
β'H777
B
CBR9379
β'BH
β'F-loop
Y555
R637
β
L777
S642
E641
F773
G640
I774
I755
P552
H551
H777
K749
P750
D551
D444
130°
β'I774
β'L770
βS642
βR637
βY555
β'I755
βE641
βG640
β'F773
β'H777
β'K749
βP552
β'P750
βD444
βH551
β'D551
C
CBR9393
β'BH
β'F-loop
Y555
R637
β
L777
S642
E641
F773
G640
I774
I755
P552
H551
H777
V550
K749
P750
D444
130°
β'I774
β'L770
βS642
βR637
βY555
β'I755
βE641
βG640
β'F773
βP552
β'H777
βV550
β'P750
β'K749
βD444
βH551

<u>Figure D.2:</u> Contacts between RNAP and CBR compounds.

A schematic summary of the CBR compound/RNAP contacts is shown on the left. The *arcs* denote nonpolar contacts. Polar contacts are denoted by *red lines*. Sites of RNAP substitutions that confer CBR^R are *underlined*. A structural view of the CBR binding pocket is shown on the right. The RNAP is shown as an *α-carbon backbone worm* (β subunit, *light cyan*; β′F loop, *orange*; β′BH, *magenta*). Side chains that make direct contact with the CBR compound are shown in *stick representation*. The CBR compound is also shown in *stick representation* with *green carbon atoms*. Polar interactions are denoted by *dashed red lines*. **(A)** CBR703; **(B)** CBR9379; **(C)** CBR9393.

D.3 Results

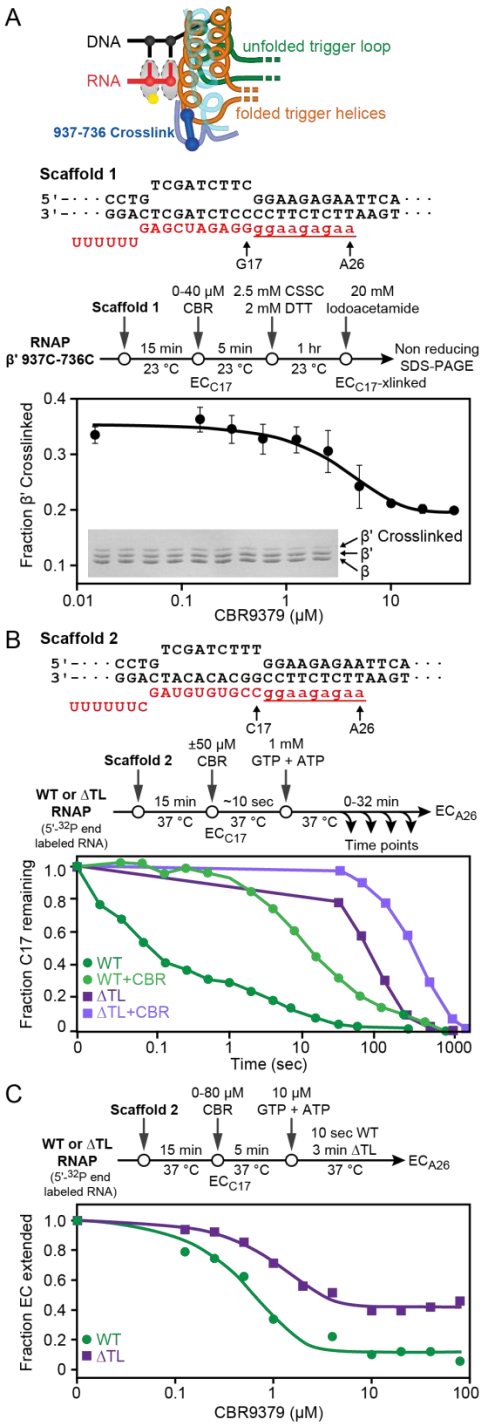
D.3.1 CBR inhibition of RNAP nucleotide addition is partially independent of the trigger loop

The RNAP nucleotide addition cycle (NAC) is governed by alternate opening and closure of the active site by a mobile structural element of the β′ subunit called the trigger loop (TL). Translocation of the elongation complex (EC) along the DNA template, as well as entry and binding of the NTP substrate in the active site, is facilitated by an open active site with an unfolded TL (24). TL contacts with the correct NTP substrate stabilize the folded TL, which closes the active site. TL contacts to NTP and RNAP then participate in an active-site rearrangement that shifts the NTP triphosphate into reactive alignment and stimulates catalysis by ~10^4 (25-27).

The tip of the folded TL is accommodated in a pocket surrounded by conserved structural elements of the RNAP, including the F loop that immediately precedes the BH in the β′ sequence. Alterations to the F loop have profound effects on the RNAP NAC, likely through effects on TL folding (8-10).

The CBR compounds make direct interactions with residues of the F loop (β′K749, β′P750, β′D751, and β′I755) **(Figure D.2)**, so we hypothesized that the binding of the CBR compounds may influence TL folding. To test this idea, we monitored TL folding by the formation of a disulfide bond in a Cys-pair reporter engineered to report the folded TL conformation (β′937–736 cross-link) (28). Addition of CBR9379 significantly decreased the formation of the 937–736 cross-link in a previously used model EC, with an apparent K_i of ~5 μM **(Figure D.3A)**. We conclude that CBR9379, and most likely CBR703 and CBR9393, inhibits TL folding.

Mutants of RNAP with a deleted TL (ΔTL RNAP) have a severely compromised NAC but are nevertheless transcriptionally active **(Figure D.3B and Figure D.4)** (27, 29). We surmised that if the inhibition of TL folding by the CBR compounds was the sole mechanism for the inhibition of RNAP, then the CBR compounds would not inhibit ΔTL RNAP. However, the TL is not required for all transcription inhibition by CBR9379 in an EC that enables measurement of nucleotide addition rates at saturating NTP **(Figure D.3B and Figure D.4A)**; of ~120-fold CBR inhibition in WT RNAP, ~5-fold remained in ΔTL RNAP. Moreover, the concentration dependence of CBR9379-mediated inhibition of WT and ΔTL RNAPs was similar **(Figure D.3C and Figure D.4B)**. We conclude that although the CBR compounds interfere with TL folding, which is obligatory for efficient catalysis of phosphodiester bond formation, the inhibition of TL folding is not the sole mechanism through which the CBR compounds inhibit the RNAP NAC. The CBR compounds also either inhibit a step distinct from TL action that becomes rate-limiting in ΔTL RNAP or, more likely, inhibit an active-site rearrangement linked to the BH that acts in concert with the TL to aid catalysis (see section **D.4 Discussion**).



<u>Figure D.3:</u> CBR9379 inhibits TL folding in an EC but still inhibits ΔTL RNAP.

(A) CBR9379 inhibits formation of the folded TL in an EC. (Top) Diagram illustrating the location of cysteine substitutions at position 937 and 736 on the β′ subunit that cross-link in the folded TL conformation (β′937C–736C) and the model EC scaffold used in prior studies (28). (Middle) The experimental protocol. (Bottom) Plot showing the disulfide cross-linking profile as a function of CBR9379 concentration in the (β′937C–736C)-RNAP. The inset shows the nonreducing SDS/PAGE gel that detects

a decrease in formation of the cross-linked β′ as the CBR9379 concentration was increased from 0-40 μM. **(B)** CBR9379 inhibits transcript elongation by the WT and ΔTL RNAP. (Top) The EC scaffold used in the elongation assay. The 5′-[γ-^{32}P]-labeled RNA (C17, shown in *red capital letters*) can be extended 9 nt (to A26; lowercase, underlined text) upon addition of GTP and ATP. (Middle) Diagram illustrating the experimental protocol. (Bottom) Plot showing the effect of CBR9379 on the elongation profiles of WT and ΔTL RNAP. With or without 50 μM CBR9379, elongation of C17 was single-exponential for ΔTL RNAP (~0.002 s^{-1} and ~0.009 s^{-1}, respectively) but was complex for WT RNAP (on average, ~0.05 s^{-1} and ~6 s^{-1}, respectively). **(C)** Effect of CBR9379 titration on the elongation profiles of WT and ΔTL RNAPs. The EC scaffold used in the elongation assay is the same as in **(B)**. (Top) Diagram illustrating the experimental protocol. (Bottom) Plot showing fraction EC (>RNA17) extended versus CBR9379 concentration. Note that some residual activity remains even at the highest CBR9379 concentrations for the ΔTL RNAP.

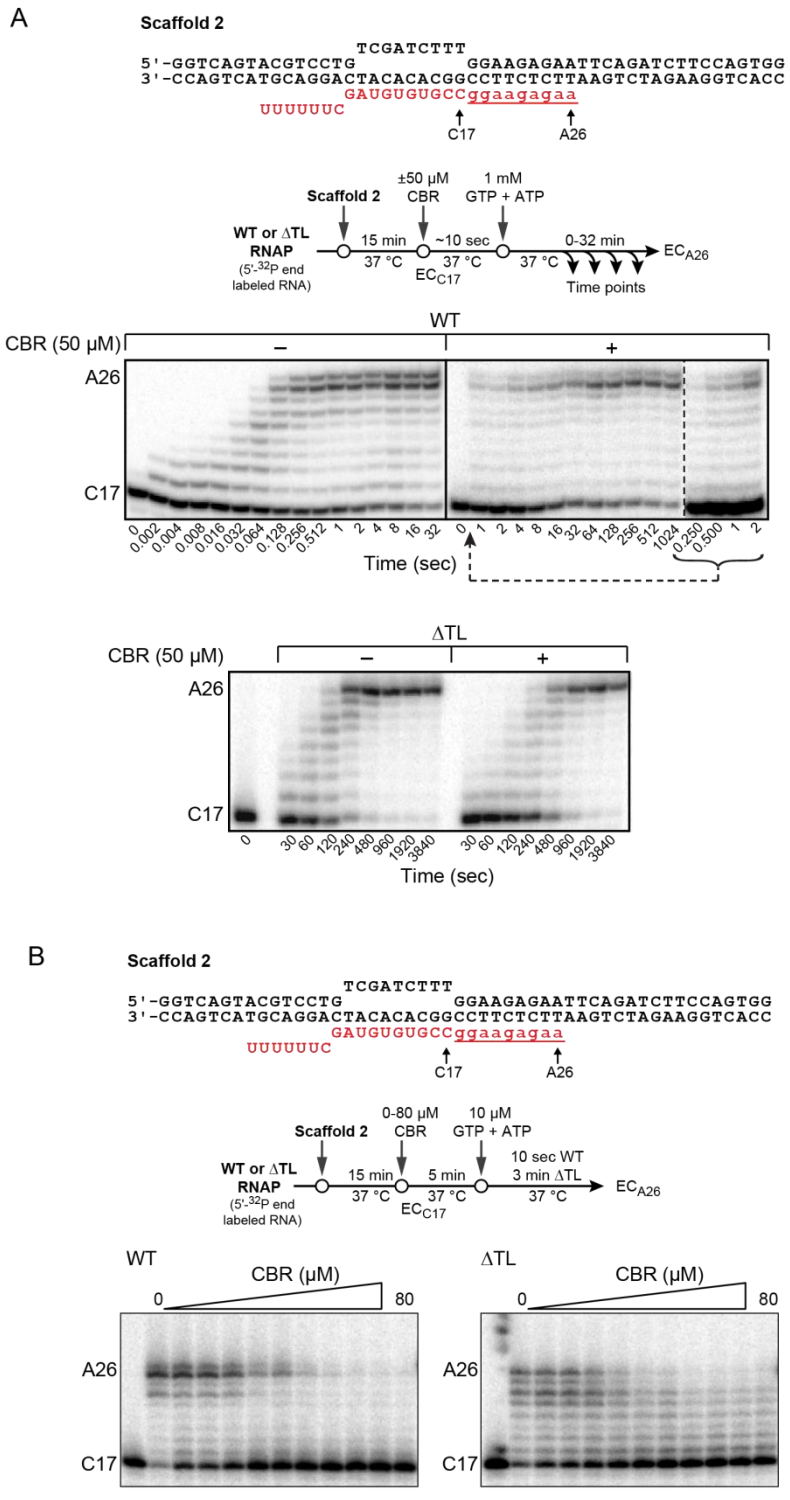


Figure D.4: CBR9379 inhibits WT and ΔTL RNAPs.

(A) (Top) The EC scaffold used in the elongation assay. The 5′-[γ-^{32}P]-labeled RNA (C17, shown in *red capital letters*) can be extended 9 nt (to A26; *lower case, underlined text*) upon addition of GTP and ATP. (Middle) Diagram illustrating the experimental protocol. (Bottom) Denaturing PAGE gels showing

extension of the 5′-[γ-^{32}P]-labeled RNA17 (C17) by 9 nt as a function of time (0-32 s or 0-1024 s for WT RNAP -/+ CBR9379, 0-3840 s for ΔTL RNAP). **(B)** (Top) The EC scaffold used in the elongation assay. The 5′-[γ-^{32}P]-labeled RNA (C17, shown in *red capital letters*) can be extended 9 nt (to A26; *lower case, underlined text*) upon addition of GTP and ATP. (Middle) Diagram illustrating the experimental protocol. (Bottom) Denaturing PAGE showing extension of the 5′-[γ-^{32}P]-labeled RNA17 (C17) by 9 nt as a function of increasing concentration of CBR9379 (0-80 μM). NTP concentration and reaction time were optimized both for WT and ΔTL RNAP, so that it will span the spectrum of no inhibition to maximal CBR inhibition. For WT, reactions were quenched 10 s after addition of 10 μM GTP and ATP and for ΔTL RNAP, 3 min after addition of 1 mM each of GTP and ATP (see reaction schemes above the gel).

D.4 Discussion

The multisubunit cellular RNAPs are complex molecular machines that control catalysis through concerted conformational changes of structural modules surrounding the active site. These structural modules include the F loop, the TL, and the BH. The nature of these conformational changes, their coordination, and their detailed roles in the different steps of the reactions of RNA synthesis and degradation catalyzed by the RNAP active site (30) remain central issues in understanding the structural basis of RNAP catalytic activity.

The BH, a long α-helix of the β′ subunit that traverses across the RNAP active site, separating the RNAP nucleic acid binding cleft from the secondary channel, plays a central role by buttressing the F loop and TL. The BH is one of the most highly conserved elements of the RNAP, highlighting its key role in RNAP catalysis. Because the BH has been observed in either straight or kinked conformations in different RNAP crystal structures, it has been proposed that cycling of the BH between these two conformations is an integral part of the RNAP NAC (7, 15, 17, 31, 32).

Transcription inhibitors can be powerful tools to study RNAP conformational dynamics. Some RNAP inhibitors, such as streptolydigin (Stl), are thought to function by interfering with TL and/or BH conformational changes (27, 33). Other inhibitors have been proposed to function by interfering with conformational changes of the RNAP clamp domain (34, 35).

Our structural results establish that the CBR compounds bind a small, preformed hydrophobic pocket in the RNAP structure surrounded by structural elements known to be critical for RNAP function **(Figure D.2)**, including elements of the β subunit that link FL2 and βDII, and the β′ subunit F loop and BH. The mobile RNAP active-center structural modules are either contacted directly by the bound CBR compounds (F loop and BH), or CBR binding affects their dynamics (TL). The CBR binding pocket is far from the RNAP active site (31 Å), but CBR

compound binding effectively inhibits all known catalytic activities of the RNAP active site (6), including intrinsic transcript cleavage (*this work, not shown here*). The CBR compounds do not compete with substrate binding and also do not principally affect translocation (6, 21), reflecting an allosteric inhibitory mechanism.

The results of the present study provide insights into the catalytic mechanism of RNAP. By virtue of specific contacts with the F loop, the CBR compounds inhibit TL folding **(Figure D.3A)**. However, the TL is not required for CBR-mediated suppression of nucleotide addition **(Figure D.3B-C)**, pointing to an additional mechanism for inhibition. Similarly, CBR compounds inhibited pyrophosphorolysis even when the TL was deleted (*this work, not shown here*). Thus, although interference with TL folding may contribute to CBR-mediated inhibition, a key basis for CBR inhibition of nucleotide addition involves an aspect of RNAP catalysis that is TL-independent. This feature of RNAP catalysis likely involves a conformational rearrangement that is blocked by CBR compound binding and therefore would likely involve rearrangements of RNAP elements directly contacted by the CBR compounds (the N-terminal segment of the BH, the F loop, and nearby elements of the β subunit) **(Figure D.2)**.

Intrinsic cleavage is promoted by the TL, at least for the 1-nt A-backtracked EC (*this work, not shown here*), but intrinsic cleavage by the ΔTL RNAP is not further inhibited by the presence of CBR9379 (*this work, not shown here*), indicating that the putative conformational rearrangement inhibited by the CBR compounds, while being important for nucleotide addition, is not important for intrinsic cleavage.

Essential conformational changes around the N-terminal part of the BH (where CBRs bind) **(Figure D.2)** have previously been proposed (36). The most compelling evidence for conformational changes in the N-terminal region of the BH comes from Pro-scanning mutagenesis of the entire BH sequence of *Methanocaldococcus jannaschii* RNAP; Pro

substitutions at every position of the BH drastically lower transcription activity, except at two positions, corresponding to *E. coli* β′H777P (defining the BH N-terminal hinge; BH-H_N) and β′S793P (defining BH C-terminal hinge; BH-H_C), where Pro substitutions strikingly increase activity (36).

The finding that Pro substitutions at two specific positions increase transcription activity implies that the helical structure of the BH may normally (in WT RNAP) be transiently disrupted (possibly kinked) at these sites during the NAC and that the kinked conformations aid one or more steps of the NAC. Kinking has been directly observed for the BH-H_C: Crystal structures of bacterial RNAP generally show two distinct conformations of the BH, either straight or kinked. In structures with a kinked BH, the kink occurs exactly at the position defined as the BH-H_C. No such BH conformational changes have been observed at the BH-H_N. However, crystal structures trap static, thermodynamically stable structures. It is possible the putative alternative conformation of the BH kinked at BH-H_N is unstable and occurs only transiently during the NAC.

Our results support the hypothesis that the CBRs inhibit phosphoryl transfer reactions catalyzed by RNAP in part by preventing the kinked BH conformation at the BH-H_N. Another recently described inhibitor, salinamide, also binds near BH-H_N, although to a distinct site (37). It seems likely that salinamide acts by a mechanism similar to CBR inhibitors by inhibiting TL folding and also by affecting BH-H_N kinking.

Given that the TL-independent inhibition of catalysis by CBR compounds is observed only for reactions involving triphosphate (nucleotide addition and pyrophosphorolysis) and the observation that NTPs bind in the RNAP active site in both active (insertion) and inactive (preinsertion) configurations of the triphosphate (25), it is tempting to speculate that the triphosphate rearrangement may be facilitated by linked conformational changes in the active

site and BH-H_N. If such motions are required to move the triphosphate into reactive position relative to the catalytic Mg^{2+} ions (or to move pyrophosphate in the case of pyrophosphorolysis), then the TL-independent effect of CBR compounds could be accounted for by inhibition of the BH movements or movements of surrounding parts of RNAP coupled to the BH movements. Such an effect of the BH could stabilize the transition state additively with nucleotide positioning by TL folding and thus could contribute to catalysis in WT RNAP, although we cannot currently rule out the possibility that a slow BH-dependent effect on catalysis is unmasked in ΔTL RNAP. It also is possible that a TL- or BH-dependent effect contributes on the relevant time scale at some template positions more than others, which could explain why CBR compounds affect WT RNAP at some template positions like pause sites more than others (6, 21). Indeed, we observed more uniform effects of CBR compounds in ΔTL RNAP than WT RNAP (e.g., **Figure D.4B**). Further experimentation will be necessary to test these ideas.

D.5 Materials & Methods

D.5.1 Materials

DNA and RNA oligonucleotides **(Table D.S2)** were obtained from Integrated DNA Technologies (IDT; Coralville, IA). All oligonucleotides were purified by 8 M urea denaturing polyacrylamide gel electrophoresis (PAGE) before use. [γ-^{32}P]-ATP was obtained from PerkinElmer Life Sciences (Warwick, RI); NTPs were obtained from Promega (Madison, WI).

D.5.2 Proteins

Core wild-type and mutant *E. coli* RNAPs **(Table D.S1)** were purified from overexpression plasmids transformed in *E. coli* strain RL2657 as described previously (38).

D.5.3 CBR compounds

CBR703 **(Figure D.1B)**, CBR9379 **(Figure D.1C)**, and CBR9393 **(Figure D.1D)** were synthesized as described previously (6, 22, 23). CBR703-Br was synthesized as described for 4-bromo-N-phenylbenzamidoxime (39) substituting 3-bromo-N-phenylbenzamide for 4-bromo-N-phenylbenzamide. CBR9379S was synthesized as described for CBR9379 but substituting thiophosgene for phosgene.

D.5.4 *In vitro* transcription elongation assay

RNAs was 5′-[γ-^{32}P]-ATP-end-labeled by incubating 50 μM RNA (8858; **Table D.S2**) with ~ 1 μM [γ-^{32}P]-ATP, 50 μM ATP and T4 PNK (New England Biolabs, Ipswich, MA) for 30 min at 30 °C in T4 PNK buffer provided by the manufacturer. Scaffold was prepared by mixing 500 nM [γ-^{32}P]-ATP-end-labeled RNA, 1000 nM template DNA, and 2000 nM nontemplate- DNA (8854 and 8853, respectively; **Table D.S2**) in reconstitution buffer (RB; 20 mM Tris-Cl, pH 8.0,

20 mM NaCl, 0.1 mM EDTA), heated to 95 °C for 5 min, followed by rapid cooling to 45 °C, and then slow cooling to RT in 2 °C/2 min steps (28). ECs were reconstituted by incubating 50 nM scaffold either with 150 nM WT RNAP or ΔTL RNAP in transcription buffer (TB.CBR; 50 mM Tris-Cl, pH 8.0, 20 mM NaCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 2.5 μg acetylated BSA/mL) for 15 min at 37 °C. Following reconstitution, ECs were incubated without or with 50 μM CBR9379 at 37 °C for 5 min before extending by 9 nt by adjusting to 1 mM each GTP and ATP. Samples were taken for times of 5 s or less using a KinTek quench-flow apparatus (KinTek Corporation, Austin, TX), quenched with 2 M HCl (40 μL), and immediately neutralized to pH 8.0 by addition of 3 M Tris base (40 μL). The RNA was phenol extracted, ethanol precipitated, and suspended in 1x stop buffer, made with equal volumes of 2x stop buffer (8 M urea, 50 mM EDTA, 90 mM Tris-borate buffer, pH 8.0, 0.02% bromophenol blue, 0.02% xylene cyanol) and 1x TB.CBR. Samples for 5 s or longer were mixed with pipettes and quenched by addition of an equal volume of 2x stop buffer. Transcription products were resolved by denaturing 15% polyacrylamide gel electrophoresis. Gels were exposed to phosphor screens and quantified using a Typhoon PhosphorImager and ImageQuant software (GE Healthcare). The remaining RNA-17 fraction in each lane was calculated and plotted versus time for the WT and ΔTL RNAPs both in the absence and presence of CBR9379. Fraction EC extended (>RNA-17) was calculated in each lane and plotted versus the CBR concentration **(Figure D.3B)**.

Table D.S1: RNAP variants and overexpression plasmids

Plasmid #	Plasmid Name	RNAP TL mutant	Description	Use
2956	pRM756	WT	Expresses wild-type *E. coli* RNAP (α2ββ´ω) with His_{10} tag on β´ C-terminus	(38)
2978	pRM778	ΔTL	pRM756 derivative expressing mutant *E. coli* RNAP (α2ββ´ω) containing β' Δ931-1137 with SNAP-tag on β´ C-terminus	this work & (38)

Table D.S2: Oligonucleotides used in this study.

Oligo #	RNA / DNA	Sequence (5' to 3')	Use
8853	DNA	GGTCAGTACGTCCTGTCGATCTTTGGAAGAGAATTCAGATCTTCCAGTGG	NT-strand for scaffold reconstitution
8854	DNA	CCACTGGAAGATCTGAATTCTCTTCCGGCACACATCAGGACGTACTGACC	T-strand for scaffold reconstitution
8858	RNA	UUUUUUCGAUGUGUGCC	RNA strand for scaffold reconstitution

D.6 REFERENCES

1. Campbell EA*, et al.* (2001) Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell* 104(6):901-912.

2. Feklistov A*, et al.* (2008) Rifamycins do not function by allosteric modulation of binding of Mg2+ to the RNA polymerase active center. *Proc Natl Acad Sci U S A* 105(39):14820-14825.

3. Hartmann G, Honikel KO, Knusel F, & Nuesch J (1967) The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochim Biophys Acta* 145(3):843-844.

4. Floss HG & Yu TW (2005) Rifamycin-mode of action, resistance, and biosynthesis. *Chem Rev* 105(2):621-632.

5. Bush K*, et al.* (2011) Tackling antibiotic resistance. *Nat Rev Microbiol* 9(12):894-896.

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

7. Lane WJ & Darst SA (2010) Molecular evolution of multisubunit RNA polymerases: structural analysis. *J Mol Biol* 395(4):686-704.

8. Miropolskaya N, Artsimovitch I, Klimasauskas S, Nikiforov V, & Kulbachinskiy A (2009) Allosteric control of catalysis by the F loop of RNA polymerase. *Proc Natl Acad Sci U S A* 106(45):18942-18947.

9. Miropolskaya N*, et al.* (2014) Interplay between the trigger loop and the F loop during RNA polymerase catalysis. *Nucleic Acids Res* 42(1):544-552.

10. Miropolskaya N, Nikiforov V, Klimasauskas S, Artsimovitch I, & Kulbachinskiy A (2010) Modulation of RNA polymerase activity through the trigger loop folding. *Transcription* 1(2):89-94.

11. Tan L, Wiesler S, Trzaska D, Carney HC, & Weinzierl RO (2008) Bridge helix and trigger loop perturbations generate superactive RNA polymerases. *J Biol* 7(10):40.

12. Jovanovic M*, et al.* (2011) Activity map of the Escherichia coli RNA polymerase bridge helix. *J Biol Chem* 286(16):14469-14479.

13. Hein PP & Landick R (2010) The bridge helix coordinates movements of modules in RNA polymerase. *BMC Biol* 8:141.

14. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, & Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448(7150):157-162.

15. Gnatt AL, Cramer P, Fu J, Bushnell DA, & Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* 292(5523):1876-1882.

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

17. Korzheva N*, et al.* (2000) A structural model of transcription elongation. *Science* 289(5479):619-625.

18. Naji S, Bertero MG, Spitalny P, Cramer P, & Thomm M (2008) Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res* 36(2):676-687.

19. Thomm M, Reich C, Grunberg S, & Naji S (2009) Mutational studies of archaeal RNA polymerase and analysis of hybrid RNA polymerases. *Biochem Soc Trans* 37(Pt 1):18-22.

20. Kireeva ML, Domecq C, Coulombe B, Burton ZF, & Kashlev M (2011) Interaction of RNA polymerase II fork loop 2 with downstream non-template DNA regulates transcription elongation. *J Biol Chem* 286(35):30898-30910.

21. Malinen AM*, et al.* (2014) CBR antimicrobials alter coupling between the bridge helix and the beta subunit in RNA polymerase. *Nat Commun* 5:3408.

22. Li LC, X.; Mihalic, J. T.; Cutler, S. (2004) *Pyrazole Antimicrobial Agents*.US Patent 6,673,923.

23. Li LC, X.; Fan. P.; Mihalic, J. T.; Cutler, S. (2006) *Hydroxyamidines and Related Compounds Are Provided Which Are Suitable as Antibacterial Agents*.US Patent 7,148,259.

24. Malinen AM*, et al.* (2012) Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Res* 40(15):7442-7451.

25. Vassylyev DG*, et al.* (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448(7150):163-168.

26. Wang D, Bushnell DA, Westover KD, Kaplan CD, & Kornberg RD (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127(5):941-954.

27. Temiakov D*, et al.* (2005) Structural basis of transcription inhibition by antibiotic streptolydigin. *Mol Cell* 19(5):655-666.

28. Nayak D, Voss M, Windgassen T, Mooney RA, & Landick R (2013) Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol Cell* 50(6):882-893.

29. Toulokhonov I, Zhang J, Palangat M, & Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* 27(3):406-419.

30. Sosunov V*, et al.* (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J* 22(9):2234-2244.

31. Zhang G*, et al.* (1999) Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. *Cell* 98(6):811-824.

32. Epshtein V*, et al.* (2002) Swing-gate model of nucleotide entry into the RNA polymerase active center. *Mol Cell* 10(3):623-634.

33. Tuske S*, et al.* (2005) Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell* 122(4):541-552.

34. Mukhopadhyay J*, et al.* (2008) The RNA polymerase "switch region" is a target for inhibitors. *Cell* 135(2):295-307.

35. Belogurov GA*, et al.* (2009) Transcription inactivation through local refolding of the RNA polymerase structure. *Nature* 457(7227):332-335.

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

37. Degen D, *et al.* (2014) Transcription inhibition by the depsipeptide antibiotic salinamide A. *Elife* 3:e02451.

38. Windgassen TA, *et al.* (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res* 42(20):12707-12721.

39. Cibian M, Ferreira JG, & Hanan GS (2009) 4-Bromo-N-phenyl-benzamidoxime. *Acta Crystallogr Sect E Struct Rep Online* 65(Pt 11):o2820.

ABBREVIATIONS

2° channel: secondary channel of RNAP

Ala: alanine

Arg: arginine

asDNA: antisense DNA oligonucleotide, complementary to the RNA in the exit channel of RNA polymerase

asRNA: antisense RNA oligonucleotide, complementary to the RNA in the exit channel of RNA polymerase

ATP: adenosine triphosphate

AT-rich: adenine- and thymine-rich

BH: bridge helix

BH-H_C: bridge helix C-terminal hinge

BH-H_N: bridge helix N-terminal hinge

bp(s): base pair(s)

CBR^R: CBR resistant

cryo-EM: cryo electron microscopy

CTD: C-terminal domain

CTP: cytidine triphosphate

dA: deoxyadenosine monophosphate

dsDNA: double-stranded DNA

dwsDNA: downstream DNA

EC: elongation complex

FL2: fork-loop 2

FL-NusA: full-length dissociable transcription factor NusA

GC-rich: guanosine- and cytosine-rich

GTP: guanosine triphosphate

His: histidine

mAb: monoclonal antibody

MccJ25: microcin J25

NA: nucleic acids

NAC: nucleotide addition cyle

nasDNA: non-antisense DNA oligonucleotide, non-complementary to the RNA in the exit channel of RNA polymerase

nasRNA: non-antisense RNA oligonucleotide, non-complementary to the RNA in the exit channel of RNA polymerase

nt: nucleotide(s)

NTD: N-terminal domain

NT-DNA: non-template strand

NTP: nucleoside triphosphate

NusA-NTD: N-terminal domain of the dissociable transcription factor NusA

pC: pyrrolo-C

Rif(s): rifamycin(s)

RNAP: RNA polymerase

RNAP-FT: tip of the flap module of RNA polymerase

RNAP-αCTD: C-terminal domain of 1 of the 2 RNA polymerase alpha subunits

rU: uridine monophosphate

S: supernatant fraction of magnetic-partitioning experiment

SBHM: sandwich-barrel hybrid motif

SI3: sequence insertion 3

ssRNA: single-stranded RNA

Stl: streptolydigin

Sw1-5: switches 1-5

T-DNA: template strand

TE: termination efficiency

Tgt: tagetotoxin

TH: trigger helices, the folded form of the trigger loop

T_{hp}: terminator hairpin

TL: trigger loop

U: uridine monophosphate

UTP: uridine triphosphate

U-tract: uridine monophosphate-rich RNA tract

Val: valine

W: full sample from magnetic-partitioning experiment (supernatant + pellet)

WT: wild-type