MECHANISM OF SIGMA-DEPENDENT PAUSING IN REGULATION OF THE BACTERIOPHAGE LAMBDA LATE GENE OPERON

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fullfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Eric John Strobel

January 2015

MECHANISM OF SIGMA-DEPENDENT PAUSING IN REGULATION OF THE BACTERIOPHAGE LAMBDA LATE GENE OPERON

Eric John Strobel, Ph.D.

Cornell University 2015

The movement of RNA polymerase (RNAP) along a DNA template is punctuated by frequent pauses during which the elongation complex persists at a single position for an extended period of time. Such pausing events have been shown to function in the regulation of gene expression at the level of transcription. A prototypical example of regulatory pausing is provided by the lambdoid phage late gene operon, in which a DNA-encoded σ^{70}-dependent pause element halts transcription to facilitate incorporation of the antiterminator Q into the paused elongation complex. Upon binding, Q triggers rapid escape of the now terminator-resistant elongation complex from the pause site. The σ^{70}-dependent pause element consists of a hexameric "-10-like sequence", which functions by capturing RNAP through contacts with the σ^{70} subunit, and a "translocation pause element" (TPE), a pause-inducing element that underlies transcription pausing in *E. coli*. Here, we show that the TPE contributes to σ^{70}-dependent pausing by limiting the rate at which paused complexes attempt to escape the pause and we identify a novel sub-element that contributes to TPE function. Furthermore, we use a functional analogy

to show that Q drives escape from the σ^{70}-dependent pause in a process that is mechanistically related to escape from a promoter during transcription initiation.

BIOGRAPHICAL SKETCH

Eric Strobel graduated from Lynbrook High School, N.Y. in 2005 and received his B.S. in Biology from St. John's University in 2009. He entered the Field of Microbiology at Cornell University in 2009 and joined the lab of Jeffrey Roberts in 2010.

To Mom and Dad,

Thanks for everything.

ACKNOWLEDGEMENTS

I thank my advisor, Jeffrey Roberts, for making my time as a graduate student a unique and incredible experience. I thank him for allowing me the freedom to pursue my own ideas and interests, and for the guidance he provided along the way. Working alongside Jeff has been an honor and a pleasure, and I am very fortunate to have had the privilege to be his final student.

I thank Joseph Peters and John Helmann for serving as members of my committee. Their suggestions and support have been invaluable. I would also like to thank Eric Alani for his support and encouragement.

I thank all past members of the Roberts Lab. In particular, I thank Joshua Filter, Jeremy Bird, and Manhee Suh for the fun times we had while doing great science. I thank Josh for teaching me the fundamentals of *in vitro* transcription and for all of our lunch-time excursions.

I thank Jennie Sims, for making me smile while I pushed through my last few months as a graduate student.

Finally, I thank my parents for the unending love and support they have given me throughout my entire life. Without them I would not be where I am today.

TABLE OF CONTENTS

CHAPTER ONE:

INTRODUCTION

The transcription cycle comprises three stages: *initiation*, including the positioning of RNA polymerase at the start of a transcription unit, the opening of DNA to expose the template strand, and the initial synthesis of RNA; *elongation*, the synthesis of the RNA encoded in the transcription unit; and *termination*, the removal of RNA polymerase from the DNA template and release of the nascent RNA once the transcription unit has been successfully transcribed. Each of these stages is a target for transcription regulation.

RNA Polymerase

The structure and function of the cellular core DNA-dependent RNA polymerases are conserved between prokaryotes and eukaryotes (1-3). The prokaryotic RNAP core complex consists of five subunits: β, β', two α subunits, and ω. The RNAP core complex is sufficient to carry out the synthesis of RNA on DNA templates that do not require initiation from a promoter (e.g. nicked templates, templates with a pre-formed bubble), but is insufficient to initiate transcription at a promoter. Initiation from a promoter requires that RNAP core associate with a σ factor, a class of protein that both confers promoter specificity and mediates the DNA opening, to form RNAP holoenzyme ($\beta\beta'\alpha_2\omega\sigma$) (4).

Initiation

The first step of transcription initiation is the binding of RNAP to a promoter and the formation of a closed promoter complex. Each sigma factor recognizes a unique set of sequences; the *E. coli* 'housekeeping' sigma factor σ^{70} binds primarily the hexameric -10 and -35 promoter elements, named for their position relative to the transcription start site. The -10 element (consensus: 5'-TATAAT-3') is bound by σ^{70} region 2, which facilitates the melting of the AT-rich -10 element to form a transcription bubble (4,5). Some promoters contain additional elements that flank the -10 element, namely, the downstream discriminator (consensus: 5'-GGG-3') and the upstream extended -10 element (consensus 5'-TG-3'), which strengthen the interaction between σ^{70} and the promoter (6,7). The -35 element (consensus 5'-TTGACA-3') is located approximately 17 bases upstream of the -10 element and is the binding site for σ^{70} region 4 (4).

Once bound to the promoter, the DNA of the -10 element and downstream is melted from -12 to +2 to generate an open complex, placing a template DNA nucleotide in the active center (8). Upon open complex formation, RNA synthesis begins. However, RNAP must overcome two barriers in order to undergo promoter escape and transition into elongation. First, the contacts between σ^{70} and the promoter -10 and -35 elements that facilitate promoter recognition also restrain the initial transcribing complex from forward movement and establish an energetic barrier to promoter escape (9,10). Because of the interaction between σ^{70} and the -10 and -35 elements RNA polymerase cannot move forward during the early stages of RNA synthesis, and instead draws downstream DNA into itself by compressing the melted DNA strands into a scrunched

conformation, resulting in a high-energy complex (9,10). The energy accumulated during scrunching can be released in one of two ways: First, the energy stored in scrunched DNA can be channeled to break the interaction between sigma and the promoter elements by rewinding the upstream edge of the transcription bubble. The release of sigma from the promoter elements drives RNAP out of the promoter and results in promoter escape and a transition into elongation. In the second outcome, the energy stored in scrunched DNA is released, but fails to break the interaction between sigma and the promoter elements. In this case, downstream edge of the transcription bubble rewinds and the nascent transcript is released in a process called abortive initiation, after which another initiation attempt can be made.

The second barrier to promoter escape is structural element σ^{70}, the $\sigma_{3.2}$ loop, which extends into the RNAP active center and has been implicated in binding of the initiating nucleotide (11). The $\sigma_{3.2}$ loop is positioned such that it blocks the path through which the nascent RNA exits RNAP and thus acts as a physical barrier that must be displaced by emerging RNA in order for synthesis to extend beyond a few nucleotides (12). Presumably, the $\sigma_{3.2}$ loop disfavors initiation by acting as a physical barrier that interrupts the nucleotide addition cycle and thereby contributes to the tendency of RNA synthesis to abort (13).

Elongation and Pausing

Following promoter escape, RNAP undergoes a cycle of nucleotide addition as it moves along the DNA template. The RNAP active center contains two sites, designated

"i" and "i+1", which binds substrates and mediates the reaction. In order for phosphodiester bond formation to occur, the 3' terminus of the nascent RNA must be present in the i site and the incoming nucleotide must be present in the i+1 site (14). Thus, following catalysis of the nucleotidyl transferase reaction, the RNA 3' terminus will be present in the i +1 site. Such a conformation is described as "pre-translocated" as it has not yet shifted the RNA 3' terminus to the i site where it must reside for addition of the subsequent nucleotide. Upon movement of the RNA 3' terminus to the i site, RNAP enters a "post-translocated" conformation and repeats the nucleotide addition cycle (14).

Transcription Pausing During Elongation

During elongation, RNAP frequently encounters sequence elements that interrupt RNA synthesis by causing RNAP to pause at a single position on the DNA template for an extended period of time (15). Early models for transcription pausing distinguished between two classes of pause elements (16). At a class I pause (e.g. the *E. coli his* operon pause), the nascent RNA forms a hairpin structure that interacts directly with RNAP as it emerges and induces pausing by stabilizing the RNA 3' terminus in either a frayed (RNA 3' terminus has separated from the DNA) or hyptertranslocated (RNA 3' terminus has moved upstream of the RNAP active center) conformation, either of which will inhibit nucleotide addition. At a class II pause (e.g. the *E. coli ops* pause, which mediates regulation by the transcription factor RfaH), a "weak" RNA:DNA hybrid forces RNAP into a either a pre-translocated or backtracked conformation. Backtracking refers to a process in which RNAP moves backwards along the DNA template and extrudes

the RNA 3' end (17,18). The result of a backtrack is that RNAP is unable to extend the RNA chain and is trapped in an arrested state. To resume transcription, RNAP must return to an active state in which the 3'-OH of the RNA chain is positioned at the active center. This can be accomplished by either isomerization back to an active conformation or by transcript cleavage. The RNAP active center is capable of catalyzing the transcript cleavage reaction in the absence of external factors; however, this endogenous reaction is slow and inefficient (19,20). A faster reaction is catalyzed by the Gre factors, GreA and GreB, which function by positioning absolutely conserved acidic residues (D41, E44) within the RNAP active center, stimulating its capacity to perform a RNA cleavage reaction by recruiting Mg^{++} to participate in the catalytic mechanism (21). The Gre factors have also been found to act during transcription initiation to suppress the formation of abortive transcripts, although the mechanism for this activity is not entirely clear (22).

A more current model of generalized transcription pausing suggests that the vast majority of transcription pauses originate from a sequence motif that has been found to be present throughout the *E. coli* genome, and that underlies both the class I and class II pauses described above, as well as the σ^{70}-dependent pause that is the focus of this thesis (23-25). This motif, the "translocation pause element" (TPE), disfavors nucleotide addition by inhibiting the transition from a pre-translocated state to a post-translocated state. The TPE motif consists of a pyrimidine in the i+1 site, a G one position downstream of the i+1 site, and two G:C base pairs at the upstream edge of the RNA:DNA hybrid (24). These sequence elements act in concert to preserve a pre-

translocated conformation and thereby induce a transcription pause, which may be stabilized or influenced by elements like an RNA hairbin, binding of σ^{70}, or a tendency to backtrack.

Termination

Transcription termination is defined as the destabilization of an elongation complex in order to facilitate the release of a transcript and dissociation of RNAP from the DNA template. The two major pathways of transcription termination are intrinsic termination, which is mediated by an RNA-encoded terminator, and ρ-dependent termination, which is mediated by the RNA translocase ρ.

An intrinsic terminator is an RNA encoded termination signal that consists of a GC-rich RNA hairpin and a U-rich sequence that spans the RNA:DNA hybrid; the hairpin must form just at the end of the RNA:DNA hybrid. It has been demonstrated that a stalled elongation complex can be terminated by providing a DNA oligonucleotide that is complementary to the emerging segment of a mutant hairpin, suggesting that the critical activity of intrinsic termination is the engagement of the part of the transcript just upstream of the hybrid in secondary structure, and not specifically RNA hairpin formation (26). Three models describing the mechanism of intrinsic termination have been proposed: The “forward translocation model” proposes that hairpin formation facilitates rewinding of the upstream edge and unwinding of the downstream edge of the transcription bubble, thus, forcing RNAP to forward translocate without RNA synthesis (27). This would result in the removal of RNA from the active site and the dissolution of

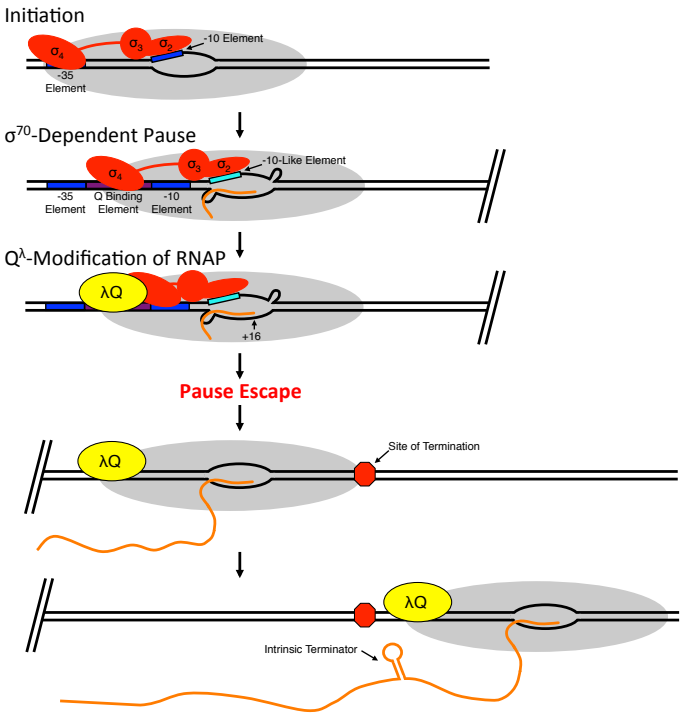
the elongation complex. The “RNA pullout” model suggests that the formation of the hairpin shortens the RNA:DNA hybrid by slippage of the homopolymeric terminal sequence against the DNA template and facilitates the release of RNA without forward translocation (28). Finally, the “allosteric” model argues that the formation of a hairpin induces a conformation change in RNAP that releases RNA. Single molecule studies have revealed that both the forward translocation and RNA pullout models do describe the function of intrinsic terminators, and suggest that the mechanism of termination at an intrinsic terminator might be largely determined by the U content of the RNA:DNA hybrid (29).

Rho is a hexameric ATP-dependent 5’-3’ RNA translocase that induces transcript release (30). Rho displays a preference for unstructured C-rich RNA and specifically targets RNA that is not being translated, thus functioning as an agent of operon polarity (31,32). It is likely that Rho facilitates transcription termination in a manner similar to an intrinsic terminator, either by forcing RNAP forward or pulling RNAP out of the complex.

Both intrinsic terminators and Rho-dependent terminators provide a target for the regulation of gene expression. A prototypical example of such a system is found in the regulation of the lambdoid phage late gene promoter pR’ (33,34). At pR’, a promoter-proximal σ^{70}-dependent pause positions the elongation complex such that it is a substrate for the assembly of the phage antiterminator Q. Q confers the property of terminator resistance on the elongation complex, allowing it to ignore a terminator that precedes the coding region of the operon (Figure 1.1).

Figure 1.1 Assembly of Q^{λ} Into Complex with RNAP at λpR'

Upon escape from the λpR' promoter, RNAP binds to a -10-like sequence undergoes a σ^{70}-dependent pause. The paused complex is the substrate for the assembly of the antiterminator Q^{λ} into the elongation complex. Assembly of Q^{λ} drives escape from the pause and confers terminator resistance so that the Q^{λ}-modified elongation complex can "read through" a terminator the precedes the coding region of the λpR' operon.


Initiation
σ4
σ3
σ2
-10 Element
-35
Element
σ70-Dependent Pause
σ4
σ3
σ2
-10-Like Element
-35
Element
Q Binding
Element
-10
Element
Qλ-Modification of RNAP
λQ
+16
Pause Escape
λQ
Site of Termination
λQ
Intrinsic Terminator

Induction of σ^{70}-dependent pausing

σ^{70}-dependent pausing is induced by the binding of σ^{70} region 2.4 ($\sigma_{2.4}$) to the non-transcribed strand of a six base pair DNA segment that resembles a promoter -10 element (called the "-10-like sequence") (35,36). The minimum sequence required for $\sigma_{2.4}$ engagement of the -10-like sequence is [5'-NANNNT-3']. While the strength of $\sigma_{2.4}$ binding to the -10-like sequence does increase when a -10-like sequence is mutated to -10 element consensus [5'-TATAAT-3'], the adenosine at position 2 and thymidine at position 6 are absolutely required for pausing to occur, in agreement with the function of these nucleotides at promoters. In addition to the core -10-like sequence, the three nucleotides immediately downstream have been shown to augment the σ^{70}:DNA interaction at the pause site by functioning like a promoter "discriminator" and contacting σ^{70} region 1.2 (36).

Induction of a transcription pause by physically restraining RNAP from forward movement confers a distinct set of properties on the paused elongation complex. Because the contacts between σ^{70} and DNA do not specifically disfavor nucleotide addition, RNAP continues to synthesize RNA beyond the position at which it first engages the -10-like sequence. RNA synthesis under these conditions is reminiscent of initial transcription, during which the promoter bound initial transcribing complex synthesizes RNA without immediately disengaging the promoter and instead draws downstream DNA into RNAP and compresses it into a "scrunched" state (9,10). The energy stored within the scrunched DNA is then used to drive promoter escape and, upon failure, results in abortive release of the nascent transcript. Similarly, evidence

from nucleotide starvation cleavage assays imply that the σ^{70}-dependent paused complex exists in a scrunched conformation, and it is likely that energy accumulated by the process of DNA scrunching drives pause escape (37). However, in the context of σ^{70}-dependent pause, failure to escape the pause results in backtracking and arrest rather than abortive release of the RNA. The RNA cleavage reaction catalyzed by the transcription factors GreA or GreB is necessary for arrested complexes to make subsequent escape attempts (37). Because the σ^{70}:DNA contacts must be broken for pause escape to occur, the strength of the interaction between σ^{70} and the -10-like sequence also functions as an energetic barrier that determines the probability that an elongation complex will successfully escape the pause.

The function of σ^{70}-dependent pausing in the recruitment and incorporation of Q into the elongation complex is tied to the mechanism through which the pause is induced. An important consequence of σ^{70}-dependent pausing is that the σ^{70}:DNA interaction is extremely effective at preserving the upstream edge of the elongation complex at a defined position on the DNA template. The paused complex positions RNAP such that Q is able to contact both the DNA Q-binding element and RNAP (specifically the beta subunit flap) so that Q can be assembled into the elongation complex (38).

The Translocation Pause Element

In addition to the non-transcribed strand elements that define the σ^{70}:DNA interaction, the involvement of transcribed strand elements in σ^{70}-dependent pausing

was detected even before the identification of σ^{70} as an agent of transcription pausing (35). In particular, heteroduplex analysis showed that the defect in σ^{70}-dependent pausing that results from G to A mutations at positions +7 and +8 of λpR' (the first two nucleotides of the discriminator element associated with of the -10-like sequence) act through both the non-transcribed strand and the transcribed strand. As described above, the non-transcribed strand effect of this mutant is attributable to a defect in contacts with σ^{70} region 1.2 that disrupts the σ^{70}:discriminator interaction. The mutation of +7G and +8G to A was suggested also to act by facilitating easier unwinding of the upstream edge of the RNA:DNA hybrid as RNAP progresses from +16 to +17 (35). This observation was later deemed correct, not only in the case of σ^{70}-dependent pausing, but also in the case of the TPE in which the G's at positions corresponding +7 and +8 of λpR' are a critical determinant (23, 24, 25). Further, the suggestion that these mutations facilitate movement of RNAP from +16 to +17 predicted a model for σ^{70}-dependent pausing that is proposed in this work two decades later.

The involvement of transcribed strand elements in σ^{70}-dependent pausing was further understood through study of the related bacteriophage 82pR' σ^{70}-dependent pause (39). In this work, important surrounding sequences for the pause were further understood to include the entire region from the position immediately following the -10-like sequence (+7 in λpR', +15 in 82pR') through the site of the pause (+16 in λpR', +25 in 82pR'). Subsequent work found that the sequence determinants that define the contribution of this region to σ^{70}-dependent pausing are consistent with the determinants of the TPE (25). Specifically, mutation of the pause site and of the "GC-rich" segment

decreases the level of σ^{70}-dependent pausing. Further, an experiment in which the entire translocation pause element is separated from the -10-like sequence by the insertion of AT-rich sequence demonstrated that the site of σ^{70}-dependent pausing is determined by the location of the translocation pause element.

Insights into the mechanism through which the translocation pause element contributes to σ^{70}-dependent pausing are presented in Chapter 3 of this work. Here, we have studied the contribution of the internal segment of the translocation pause element to σ^{70}-dependent pausing. As a result, we have determined that the primary function of the translocation pause element in σ^{70}-dependent pausing is to modulate the escape rate of paused elongation complexes. These insights have led us to propose an updated model for σ^{70}-dependent pausing.

The Compound Nature of the σ^{70}-Dependent Pause Element

Presented in this work is a model that describes the compound nature of the unified σ^{70}-dependent pause element and attributes specific roles to the -10-like sequence and translocation pause element in the process of σ^{70}-dependent pausing. In order for a transcription pause to occur, two conditions must be satisfied. First, the pause element must capture RNAP, and second, must retain RNAP at a defined position for some period of time. An exceptional property of the σ^{70}-dependent pause is that it distributes these two functions across two distinct DNA sequence elements.

As described above, the interaction between σ^{70} and the -10-like sequence is both the initial inducer of σ^{70}-dependent pausing and the energetic barrier that must be

overcome for pause escape to occur. Failure to escape the pause results in backtracking and arrest. The frequency at which escape attempts occur is dependent on two factors: the rate at which backtracked complexes undergo GreA/GreB-mediated cleavage and the duration for which complexes persist at the pause site nucleotide prior to making an escape attempt. We show that the TPE influences the rate at which paused complexes undergo escape by determining the amount of time these complexes remain at the pause site before attempting to disengage σ^{70} from the -10-like sequence.

Function of Q in σ^{70}-Dependent Pause Escape

During early studies of Q function it was observed that in the presence of Q, σ^{70}-dependent paused complexes escape the pause at a faster rate, implying that the assembly of Q into the paused elongation complex drives pause escape (34). In this work, we study the nature of Q-dependent arrest, a phenomenon that occurs when Q-modified complexes escape the σ^{70}-dependent paused complexes. We use a functional analogy between promoter escape and σ-dependent pause escape to show that DNA scrunching is the driving force behind the escape of RNAP from the λpR' σ-dependent pause and that this process is enhanced by the activity of the Q^{λ} antiterminator. Furthermore, we show that failure of transcription complexes to escape the pause results in backtracking and arrest in a process analogous to abortive initiation and identify a sequence element that modulates both abortive synthesis and the formation of arrested elongation complexes. We establish that DNA scrunching can occur outside the context of promoter escape, specifically as a means of transcription pause escape.

References

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

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

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

4. Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J. and Young, B. (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb Symp Quant Biol*, **63**, 141-155.

5. Marr, M.T. and Roberts, J.W. (1997) Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, **276**, 1258-1260.

6. Barne, K.A., Bown, J.A., Busby, S.J. and Minchin, S.D. (1997) Region 2.5 of the Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J*, **16**, 4034-4040.

7. Travers, A.A. (1984) Conserved features of coordinately regulated E. coli promoters. *Nucleic acids research*, **12**, 2605-2618.

8. Gruber, T.M. and Bryant, D.A. (1997) Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. *J Bacteriol*, **179**, 1734-1747.

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

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

11. Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L. and Darst, S.A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell*, **9**, 527-539.

12. Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. and Darst, S.A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285-1290.

13. Murakami, K.S. and Darst, S.A. (2003) Bacterial RNA polymerases: the wholo story. *Current opinion in structural biology*, **13**, 31-39.

14. Larson, M.H., Landick, R. and Block, S.M. (2011) Single-molecule studies of RNA polymerase: one singular sensation, every little step it takes. *Mol Cell*, **41**, 249-262.

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

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

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

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

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

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

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

22. Hsu, L.M., Vo, N.V. and Chamberlin, M.J. (1995) Escherichia coli transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression in vivo and in vitro. *Proc Natl Acad Sci U S A*, **92**, 11588-11592.

23. Vvedenskaya, I.O., Vahedian-Movahed, H., Bird, J.G., Knoblauch, J.G., Goldman, S.R., Zhang, Y., Ebright, R.H. and Nickels, B.E. (2014) Transcription. Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science*, **344**, 1285-1289.

24. Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Nayak, D., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R. and Weissman, J.S. (2014) A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science*, **344**, 1042-1047.

25. Bird, J. (2013), Cornell University.

26. Yarnell, W.S. and Roberts, J.W. (1999) Mechanism of intrinsic transcription termination and antitermination. *Science*, **284**, 611-615.

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

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

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

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

31. Bear, D.G., Hicks, P.S., Escudero, K.W., Andrews, C.L., McSwiggen, J.A. and von Hippel, P.H. (1988) Interactions of Escherichia coli transcription termination factor rho with RNA. II. Electron microscopy and nuclease protection experiments. *J Mol Biol*, **199**, 623-635.

32. Adhya, S., Gottesman, M. and De Crombrugghe, B. (1974) Release of polarity in Escherichia coli by gene N of phage lambda: termination and antitermination of transcription. *Proc Natl Acad Sci U S A*, **71**, 2534-2538.

33. Grayhack, E.J. and Roberts, J.W. (1982) The phage lambda Q gene product: activity of a transcription antiterminator in vitro. *Cell*, **30**, 637-648.

34. Grayhack, E.J., Yang, X.J., Lau, L.F. and Roberts, J.W. (1985) Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell*, **42**, 259-269.

35. Ring, B.Z. and Roberts, J.W. (1994) Function of a nontranscribed DNA strand site in transcription elongation. *Cell*, **78**, 317-324.

36. Ring, B.Z., Yarnell, W.S. and Roberts, J.W. (1996) Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, **86**, 485-493.

37. Marr, M.T. and Roberts, J.W. (2000) Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol Cell*, **6**, 1275-1285.

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

39. Perdue, S.A. and Roberts, J.W. (2010) A backtrack-inducing sequence is an essential component of Escherichia coli sigma(70)-dependent promoter-proximal pausing. *Mol Microbiol*, **78**, 636-650.

CHAPTER 2:
REGULATION OF PROMOTER-PROXIMAL TRANSCRIPTION ELONGATION:
ENHANCED DNA SCRUNCHING DRIVES λQ-ANTITERMINATOR DEPENDENT ESCAPE
FROM A σ^{70}-DEPENDENT PAUSE[1]

Summary

During initial transcription, RNA polymerase remains bound at the promoter and synthesizes RNA without movement along the DNA template, drawing downstream DNA into itself in a process called scrunching and thereby storing energy to sever the bonds that hold the enzyme at the promoter. We show that DNA scrunching also is the driving force behind the escape of RNAP from a regulatory pause of the late gene operon of bacteriophage λ, and that this process is enhanced by the activity of the Q^{λ} antiterminator. Furthermore, we show that failure of transcription complexes to escape the pause results in backtracking and arrest in a process analogous to abortive initiation. We identify a sequence element that modulates both abortive synthesis and the formation of arrested elongation complexes.

Introduction

The segment of DNA immediately downstream of a transcription start site frequently is an important site of regulation of transcription elongation, where the

[1] Published in *Nucleic Acids Research (2014) 42 (8): 5097-5108. doi: 10.1093/nar/gku147*

positive energies of the polymerization reaction encounter inhibitory forces that deter elongation. In one example, all multisubunit RNA polymerases (RNAPs) tend to abort early synthesis, presumably when polymerization fails to provide the energy to escape RNAP interactions with the promoter, thereby releasing abortive RNAs typically 5-10 nt in length (1-4). Importantly, both eukaryotic and bacterial RNAPs can be restrained by specific protein interactions that make further elongation dependent on particular regulatory modifications. A prominent example is provided by eukaryotic RNAP II, which frequently stalls tens of nucleotides from the transcription start site due to the inhibitory activity of proteins like Dsif and Nelf, and then is rescued into elongation by the regulatory kinase pTefb in a reaction essential to transcription activation (5) .

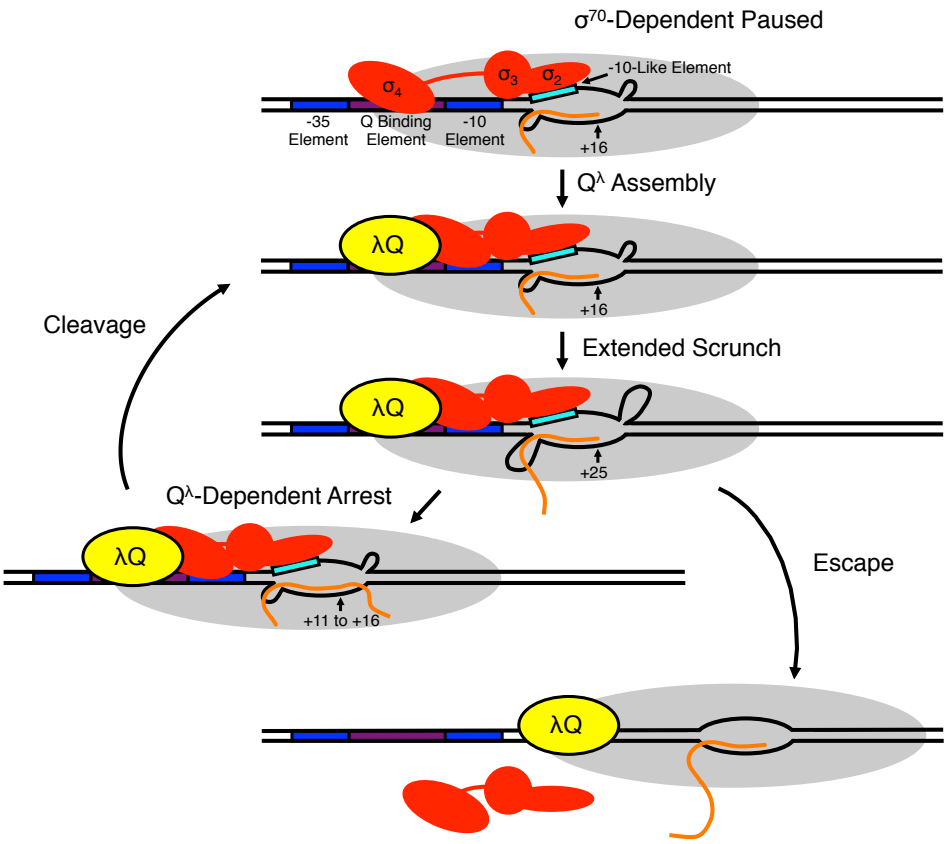
The energetic transactions in early transcription elongation are particularly exposed in the bacterial regulatory system of the bacteriophage λ gene *Q* antiterminator. Q^{λ} becomes a subunit of RNAP, allowing it to resist terminators through both antipausing and a structural rearrangement that inhibits terminator RNA function (6,7). Q^{λ} function requires an early transcription pause at +16 that is induced by a specific protein interaction, namely, binding of the σ^{70} initiation factor to a secondary -10-like sequence in the early transcribed segment (8,9). Q^{λ} helps RNAP escape this pause (as well as critical pauses at terminators downstream) (8), exhibiting an essential antipausing activity that may exemplify the way regulatory factors can overcome early transcription barriers.

The realization that initial transcribing complexes are formed through DNA 'scrunching' (10,11), namely, the unwinding of downstream DNA that then is

incorporated as single-stranded DNA into the transcription complex, provides a structural context for understanding the nature of promoter-proximal pausing (e.g. σ^{70}-induced) and the reversal of pausing by antipausing factors (like Q^{λ}). This derives directly from the scrunching model of initial transcription and abortive initiation, as follows. A consideration of energies involved in synthesis and elongation of the initial transcribing complex suggests that scrunching stores energy that is used to break the σ^{70}-DNA bonds that stabilize the open promoter complex. Further energy is required to displace a protein linker between σ^{70} domains 3 and 4 (12,13); combined, these barriers oppose the elongation reaction and can lead at a significant rate to its failure, resulting in 'abortive initiation', the loss of the initial transcript from the complex. The extent and pattern of abortive transcript release depends both on the strength of σ^{70} binding in the promoter and the nucleotide composition of the initial transcribed sequence (ITS) (3,14,15).

A reasonable inference is that σ^{70}-dependent paused complexes (Figure 2.1) have the essential features of initial transcribing complexes, including the scrunched DNA structure (16,17), with the central difference that the RNA of ≥ 16 nt is not susceptible to abortive loss because it is stabilized by the interactions that also stabilize the elongation complex, i.e. enclosure of the RNA/DNA hybrid and binding of RNA in its exit channel (18). The paused complex is anchored by σ^{70} interacting with the non-template strand of the secondary -10 sequence (9)(Figure 2.1). However, the actual site of pausing, i.e. the extent of scrunching-based elongation after σ^{70} binds DNA, is

Figure 2.1 A Model for Q^{λ}-Stimulated Escape from the λpR' σ^{70}-Dependent Pause. Upon engagement of the -10-like element by σ^{70} region 2, the transcription elongation complex undergoes a σ^{70}-dependent pause. The paused elongation complex is a substrate for modification by Q^{λ}, which engages the elongation complex through contacts with the Q binding element and with RNAP itself. Assembly of Q^{λ} into the elongation complex stimulates the resumption of transcription while σ^{70} is still bound to the -10-like element, resulting in an extended scrunch. The energy accumulated by DNA scrunching can be released in two ways: *First*, if the energy stored in the scrunch is insufficient to disrupt the interaction between σ^{70} and -10-like element, RNAP backtracks into an arrested state, termed the Q^{λ}-dependent arrested complexes (QAC), and cleavage of the backtracked RNA must occur in order to resume transcription. *Second*, if the energy stored in the scrunch is sufficient to disrupt the interaction between σ^{70} and -10-like element, the Q^{λ}-modified elongation complex disengages the pause site and resumes elongation. Presumably σ^{70} dissociates from RNAP following disengagement.


σ70-Dependent Paused
σ4
σ3
σ2
-10-Like Element
-35 Element
Q Binding Element
-10 Element
+16
Qλ Assembly
λQ
+16
Extended Scrunch
λQ
+25
Cleavage
Qλ-Dependent Arrest
λQ
+11 to +16
Escape
λQ

determined by sequence elements that also likely determine the general features of elongation such as ubiquitous pausing (19-21).

Modification of the σ^{70}-dependent paused RNAP by Q^{λ} protein, which binds to both a specific DNA Q^{λ} binding site behind RNAP (22) and to elements of RNAP [including particularly the beta flap (23)], results in a stable complex of Q^{λ} with RNAP (24) that changes the RNAP elongation properties. Binding of Q^{λ} assists RNAP in escaping the pause, and the Q^{λ}-modified complex pauses less downstream, an essential element of the antitermination mechanism (25). Just as scrunching energy is proposed to break σ^{70} bonds with promoter DNA, we propose that Q^{λ} promotes further scrunching from the σ^{70}-dependent pause that provides the energy to break σ^{70} (and presumably Q^{λ}) bonds with DNA.

We have used a distinctive property of the σ^{70}-pausing/ Q^{λ} system to provide evidence for this extended scrunching model of Q^{λ} function, and also to reveal important aspects of the sequence basis of early abortive transcript release. In addition to escaping into productive, antiterminated elongation, Q^{λ}-modified elongation complexes downstream of the σ^{70}-dependent pause site have a tendency to undergo arrest in a sequence-specific manner. We propose that this Q^{λ}-dependent arrest is analogous to a step of abortive release of early transcripts: it depends on the anchoring activity of σ^{70} (in combination with Q^{λ} in this case), and the RNAs are backtracked and sensitive to the action of GreB cleavage factor (Figure 2.1) (6), just as the formation and release of abortive initiation products are sensitive to GreB (26). The stability against dissociation

of Q^{λ}-modified complexes, which are mature elongation complexes, allows us to analyze their nature and sequence dependence. We discern important sequence elements that determine the efficiency and site of arrest of Q^{λ}-modified complexes, and we show explicitly that such sequences also determine the pattern of abortive transcript release from initial transcribing complexes. In particular, we identify a dinucleotide motif that determines a site and efficiency of major abortive RNA release. We further show that Q^{λ}-dependent arrest reflects backtracking to the site of the σ^{70}-dependent pause where Q^{λ} initially engages; this result implies that the mechanism of Q^{λ}-mediated release of RNAP from the σ^{70}-dependent pause involves further scrunching from the site of the σ^{70}-dependent pause. This result is consistent with the model that scrunching energy is stored in a stressed intermediate, which then functions in initiation to break the σ^{70}-promoter bonds, but which is proposed to act here to break both $\sigma^{70}$and Q^{λ} bonds to DNA and allow the modified complex to escape into elongation.

Materials and Methods

Plasmids

pM650 (27), p'QE-30 (28), and pET-28a- σ^{70} (29) have been described. pES3 is *GreB*-6xHis in pET-28b(+). pES7 is $GreB_{D41A,\ E44A}$-6xHis in pET-28b(+). pSAN2 (a gift from L. Hsu) contains $pN25_{Anti.}$ pVS10 (a gift from I. Artsimovitch) contains all four core RNAP subunits. All mutants were constructed using Quikchange site-directed mutagenesis.

Proteins

RNA polymerase (30), GreB and $GreB_{D41A, E44A}$ (31), σ^{70} (29), and NusA (32) were purified as described. Q^{λ} was purified as described (17) except that it is stored in 10 mM potassium phosphate pH 6.5, 50% glycerol, 200 mM potassium chloride, 1 mM ethylenediaminetetraacetic acid (EDTA), and 7mM tris(2-carboxyethyl)phosphine (TCEP) which increases the stability of Q^{λ} when stored at -20°C (T. Santangelo and J. Filter, unpublished).

In Vitro transcription

Reaction mixtures containing 2 nM template and 20 nM RNAP (20nM core reconstituted with 100 nM σ^{70}) were incubated in transcription buffer (20 mM Tris-HCL pH 8.0, 0.1 mM EDTA, 1 mM dithiothreitol and 50 mM KCl), 0.1 mg/ml bovine serum albumin, and 200 µM ATP, GTP, CTP and 50 µM UTP (containing 0.5 µCi/µl α-^{32}P-UTP) and 150 nM NusA for 15 min at 37°C to form open complex. Figure 2.6B used a nucleotide mixture containing 200 µM UTP, GTP, CTP and 50 µM ATP (containing 1.0 µCi/µl γ-^{32}P-ATP) and 10 nM template. Single-round transcription reactions were initiated by addition of magnesium chloride to 5 mM and rifampicin to 10 µg/ml. Total reaction volume was 25 µl. For abortive initiation assays (except that of Figure 2.10B), open complex was formed as above, except that the NTP concentration was 200 µM ATP, UTP, GTP and 50 µM CTP (or a dilution of this NTP mix) and contained 0.5 µCi/µl α-^{32}P-CTP). Figures 2.8D and 2.11B were repeated using three alternative NTP mixes (with ATP, UTP, or GTP 50 µM, and the remaining NTPs at 200 µM) to ensure that the NTP mix did not bias the abortive pattern. Figure 2.10B was performed using 20 µM UTP, GTP, CTP

and 5 µM ATP (containing 0.5 µCi/µl α-^{32}P-ATP). Multi-round transcription was initiated by the addition of magnesium chloride to 5 mM. When present, GreB was added to 100 nM before formation of open complex. In the experiments of Figure 2.2, 2.3A, and 2.3B containing $GreB_{D41A, E44A}$, both GreB and $GreB_{D41A, E44A}$ were added to a concentration of 1 µM in their respective reactions. When present, Q^{λ} was added to a concentration of 250 nM after open complex formation and incubated at 37°C for 30 seconds before initiation. Reactions were stopped by adding 125 µl of stop solution (0.6 M Tris pH 8.0, 12 mM EDTA, 0.16 mg/ml transfer RNA).

Cleavage assays

Biotinylated template was bound to Promega Streptavidin MagneSphere Paramagnetic Particles. Transcription was initiated in the presence of Q^{λ} and allowed to proceed for 2 min to form the Q^{λ}-dependent arrested complexes (QAC), at which point the reactions were washed three times with T buffer + 0.1 mg/ml bovine serum albumin to remove nucleotides and magnesium chloride, halting the reaction. For nucleotide starvation cleavage, reactions were then resuspended in T buffer + 100 nM GreB, incubated at 37°C, and initiated by addition of magnesium chloride to 5 mM. To cleave backtracked RNAs and chase the 5' cleavage products, reactions were resuspended in T buffer, 100 nM GreB, and 200 µM NTPs, incubated at 37°C, and initiated by addition of magnesium chloride to 5 mM and rifampicin to 10 µg/ml. Reactions were stopped by adding 125 µl of stop solution.

Heteroduplex templates

Heteroduplex templates were constructed as described previously (33).

Purification, fractionation, and analysis of transcription reactions

Stopped transcription reactions were phenol extracted by addition of 150 µl of phenol/chloroform/isoamyl alcohol (25:24:1), vortexing, centrifugation, and collection of the aqueous phase. Ethanol precipitation of RNA was performed by adding 450 µl of 100% ethanol to each reaction, followed by storage at -20 °C overnight. Precipitated RNA was resuspended in transcription loading dye (1x T buffer, 80% formamide, 0.05% bromophenol blue and xylene cyanol). Reactions were fractionated by electrophoresis using 12 or 15% denaturing polyacrylamide gels containing 6M urea. Reactive bases were detected by an Amersham Biosciences Typhoonn 9400 Variable Mode Imager. Quantitation was performed using ImageQuant. Bands were normalized based on the number of labeled bases in the transcript.'%Q^{λ}-modified' refers to the percentage of a subset of Q^{λ}-modified complexes out of all Q^{λ}-modified complexes (QAC + readthrough). '%Non-Abortive Transcripts' refers to the percentage of a particular transcript length out of all non-abortive transcripts, thus excluding abortive products.

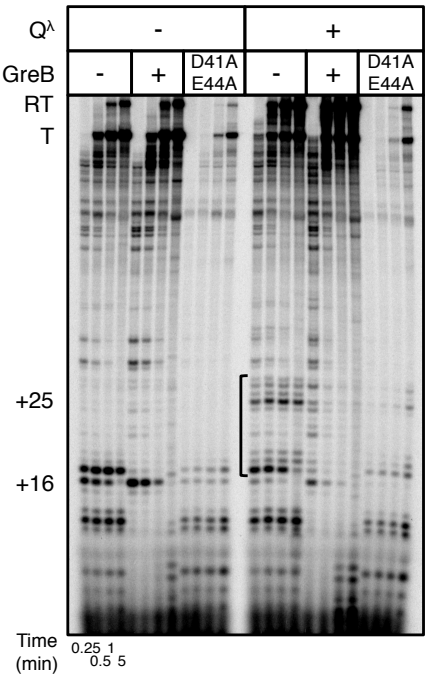
Results

Q^{λ} induces formation of a backtracked early transcription arrest

Inclusion of Q^{λ} protein during *in vitro* transcription from λpR' causes modification by Q^{λ} of 70% of the transcription complexes (Figure 2.2); of these, 52% appear as

Figure 2.2 Q^{λ}-Dependent Arrest at λpR'.

Single round *in vitro* transcription from λpR' in the presence and absence of Q^{λ}, GreB, and $GreB_{D41A, E44A}$. Readthrough (RT), terminated (T), Q^{λ}-dependent arrested complexes (QAC), and paused transcripts (+16, +17) are shown. The QAC are designated by a bracket. +17 is included in the QAC because complexes trapped at +17 chase by 5 min in the presence of Q^{λ}, suggesting that they are distinct from the complexes trapped at +17 in the absence of Q^{λ}. Percent readthrough was determined by normalizing the intensity of the RT and T bands for U content and dividing the normalized RT by the sum of the normalized RT and T.


Q^λ
-
+
GreB
-
+
D41A E44A
-
+
D41A E44A
RT
T
+25
+16
Time (min)
0.25 1
0.5 5

terminator readthrough owing to the antitermination property of Q^{λ}, and the remaining 48% are trapped in a backtracked state – termed the Q^{λ}-dependent arrested complexes (QAC). The QAC include species of 17-19 and 24-28 nt (Figures 2.1 and 2.2, bracket); they are eliminated by addition of the transcription cleavage factor GreB, resulting in all Q^{λ}-modified complexes appearing as readthrough (Figures 2.1 and 2.2). Although the QAC appear largely stable over 5 min (Figure 2.2), they disappear on longer incubation (Figure 2.3A). We earlier surmised that the QAC are paused (16), implying an ability to resume elongation, but it is more likely that their disappearance results from intrinsic transcript cleavage (or possibly trace contamination by GreA or GreB) that rescues them into productive elongation. As further evidence that the QAC are backtracked and rescued by transcript cleavage, we used a catalytic mutant of *greB* that changes both essential carboxylates (D41 and E44) to alanine (34), resulting in a mutationally altered protein that inhibits the active center-dependent intrinsic cleavage reaction (Figure 2.3B). Inclusion of $GreB_{D41A, E44A}$ slows the elongation rate substantially, but also traps 90% of the Q^{λ}-modified complexes as QAC (Figures 2.2 and 2.3A). Interestingly, $GreB_{D41A, E44A}$ also causes 97% of the total transcript to appear as abortive products (Figures 2.4A and 2.4B).

Q^{λ} drives escape from the λpR' σ^{70}-dependent pause by stimulating DNA scrunching

The identity of the first 20 nt following a transcription start site (the ITS) is a major determinant of the pattern and quantity of aborted transcripts at a promoter (15).

Figure 2.3 GreB$_{D41A, E44A}$ Enhances Formation and Prevents Cleavage of the QAC.

(A) Single round *in vitro* transcription from λpR' in the presence of Q^{λ}, and presence and absence of GreB$_{D41A, E44A}$.

(B) Nucleotide Starvation Cleavage of the λpR' QAC in the presence and absence of GreB, and GreB$_{D41A, E44A}$. In the absence of GreB most of the QAC were cleaved by 5 min, either from intrinsic cleavage or GreA/GreB contamination. In the presence of GreB the QAC were cleaved within 30 seconds. In the presence of GreB$_{D41A, E44A}$, cleavage was dramatically decreased and the QAC persisted to the final 25 minute time point.

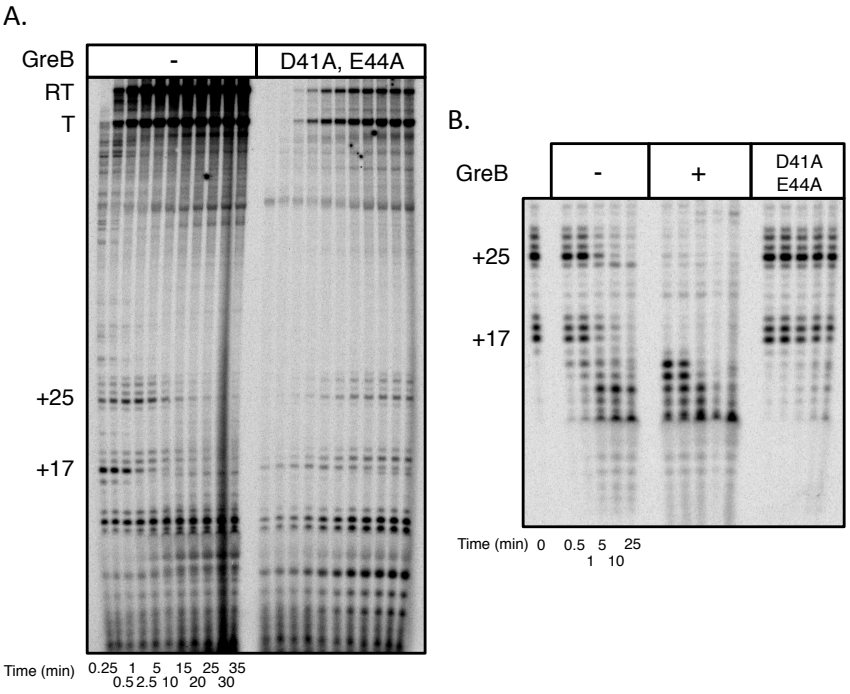
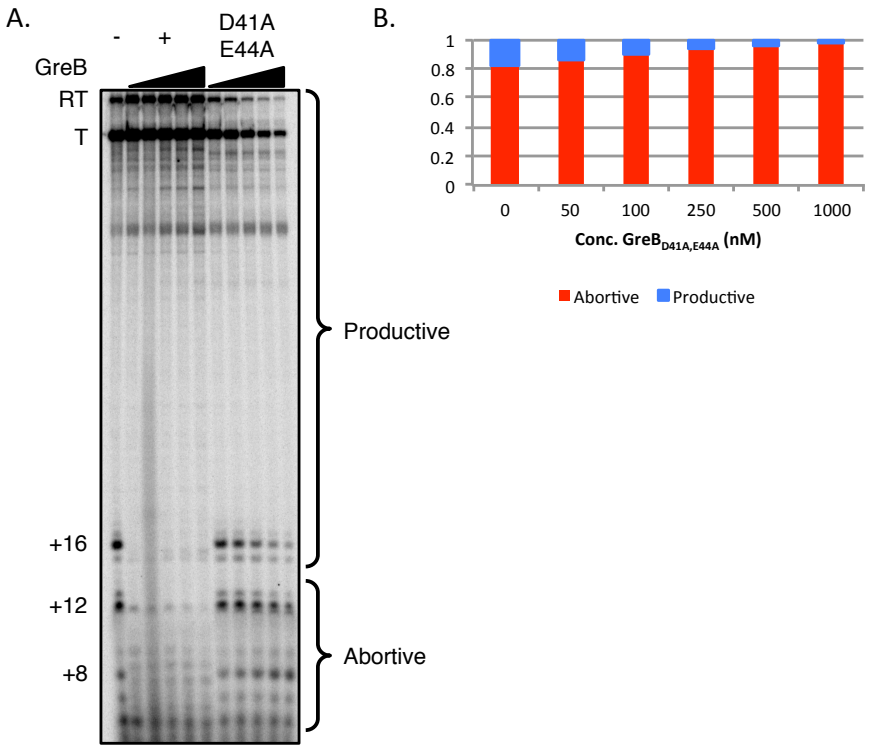

A.
GreB
-
D41A, E44A
RT
T
+25
+17
Time (min)
0.25 1 5 15 25 35
0.5 2.5 10 20 30
B.
GreB
-
+
D41A
E44A
+25
+17
Time (min)
0 0.5 5 25
1 10

Figure 2.4 GreB$_{D41A, E44A}$ Inhibits Promoter Escape.

(A) Single round *in vitro* transcription from λpR' with increasing concentrations of GreB and GreB$_{D41A, E44A}$.

(B) Quantitation of (A). The intensity of the bands was normalized for U content and calculations were performed as described in the Experimental Procedures.


A.
-
+
D41A
E44A
GreB
RT
T
+16
+12
+8
Productive
Abortive
B.
1
0.8
0.6
0.4
0.2
0
0
50
100
250
500
1000
Conc. GreB D41A,E44A (nM)
Abortive
Productive

Because abortive initiation occurs from a scrunched complex, the ITS presumably affects the stability of the scrunched complexes as transcription proceeds. Similarly, if Q^{λ}-dependent arrest is analogous to abortive initiation, the sequence downstream of the σ^{70}-dependent pause should modulate the length and distribution of the bound at the σ^{70}-dependent pause, just as an initial transcribing complex would remain bound at the promoter following a failed attempt at promoter escape.

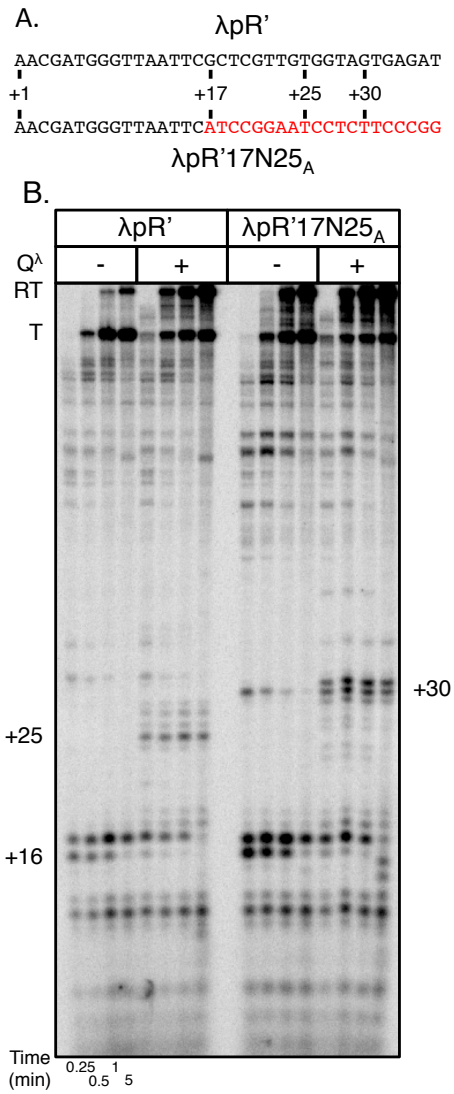
To determine the role of sequences downstream of the σ^{70}-dependent pause in Q^{λ}-dependent arrest, we replaced positions +17 to +36 of λpR' with the ITS of promoter $N25_{Anti}$ to create a template termed λpR'17$N25_{A}$ (Figure 2.5A). $N25_{Anti}$ is a strongly abortive promoter known to form abortive transcripts up to 15 nt in length (26,35), making it an ideal model for studying our analogy. If the sequence downstream of the σ^{70}-dependent pause affects Q^{λ}-dependent arrest in the same way that the ITS affects abortive initiation, then substitution of this downstream sequence with an ITS that generates long abortive products should result in the formation of QAC that contain longer RNAs as a result of transcription further downstream before collapse into a backtracked state.

Figure 2.5B depicts single-round transcription reactions from both λpR' and λpR'17$N25_{A}$ in the presence and absence of Q^{λ}. In the presence of Q^{λ}, elongation complexes transcribing from λpR' undergo Q^{λ}-dependent arrest after transcribing to a cluster of positions focused at +25, whereas elongation complexes transcribing from λpR'17$N25_{A}$ undergo Q^{λ}-dependent arrest after transcription to a cluster focused at +30 and +31. The 30-nt RNA present in these QAC is exactly 14 nt longer than the σ^{70}-

Figure 2.5 Sequence Dependence of the QAC

(A) Sequence comparison of λpR' and λpR'17N25$_A$ from positions +1 to +36

(B) Single round *in vitro* transcription from λpR' and λpR'17N25$_A$ in the presence and absence of Q^{λ}.


A.
λpR'
AACGATGGGTTAATTCGCTCGTTGTGGTAGTGAGAT
+1
+17
+25
+30
AACGATGGGTTAATTCATCCGGAATCCTCTTCCCGG
λpR'17N25A
B.
λpR'
λpR'17N25A
Qλ
-
+
-
+
RT
T
+30
+25
+16
Time
(min)
0.25
0.5
1
5

dependent pause product and thus corresponds in endpoint to a major 14-nt long abortive product of the $N25_{Anti}$ promoter. This result supports the interpretation that the $N25_{Anti}$ ITS alters Q^{λ}-dependent arrest in a way that reflects its abortive properties.

Complexes that have undergone Q^{λ}-dependent arrest are backtracked, and thus are sensitive to the cleavage factor GreB (Figure 2.2). As expected, the λpR‘17$N25_A$ QAC are GreB-sensitive, indicating that they are backtracked (Figure 2.6). It is noteworthy that in addition to the effects on Q^{λ}-dependent arrest, complexes transcribing λpR‘17$N25_A$ undergo a pause at +30 in the absence of Q^{λ} that is insensitive to GreB. The appearance of such a pause suggests that the $N25_{Anti}$ ITS has an innate preference for pausing at this position, which may influence its abortive properties and its behavior in the fusion λpR‘17$N25_A$.

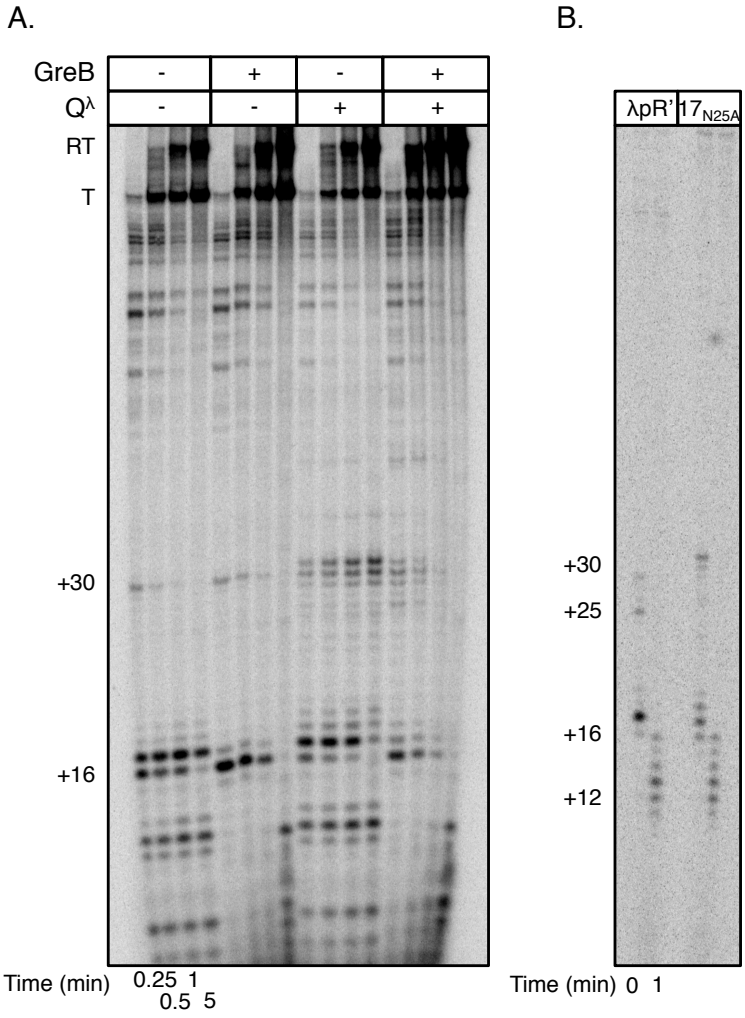
DNA scrunching requires that RNAP remain bound at a defined position while synthesizing RNA without translocation. In the context of initial transcription, failure to escape the promoter causes abortive release of the nascent transcript through a process thought to be similar to backtracking (15,26). In our model, the increase in DNA scrunching that results from modification of the σ^{70}-dependent paused complex by Q^{λ} occurs while σ^{70} is bound to the pause-inducing sequence; thus, collapse during further scrunching should result in a Q^{λ}-modified elongation complex still bound through σ^{70} interaction at the pause-inducing sequence, and arrested in a backtracked conformation. To determine the position to which the QAC of λpR’ and λpR’17$N25_A$ backtrack, we first performed GreB-mediated cleavage in the absence of NTPs (Figure 2.7A). We generated QAC on template bound to magnetic beads, washed to remove

Figure 2.6 GreB Sensitivity of the λpR'17N25$_A$ QAC

(A) The λpR'17N25$_A$ QAC Are Sensitive to GreB Stimulated Cleavage.

Single round *in vitro* transcription from λpR' was performed in the presence and absence of Q^{λ} and GreB. Readthrough (RT), terminated (T), Q^{λ}-dependent arrested complexes (QAC), and paused transcripts(+16, +17) are shown.

(B) Nucleotide starvation cleavage on the QAC of λpR' and λpR'17N25$_A$. RNAs were end-labeled using γ-^{32}P-ATP so that only 5' cleavage products are visible.


A.
GreB
-
+
-
+
Q^λ
-
-
+
+
RT
T
+30
+16
Time (min)
0.25 1
0.5 5
B.
λpR'17N25A
+30
+25
+16
+12
Time (min) 0 1

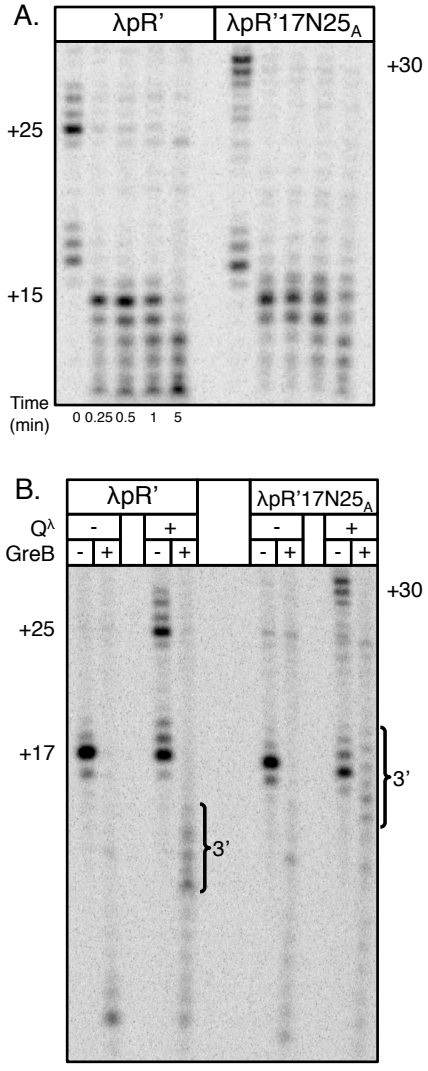
NTPs, and induced cleavage by incubating with GreB. In addition to an experiment in which RNAs were internally labeled, we performed a second nucleotide starvation cleavage experiment in which RNAs were end-labeled to allow for clear visualization of the 5' cleavage products (Figure 2.6B). Strikingly, both λpR' and λpR'17N25$_A$ QAC backtrack to a range of positions from +11 to +16, which is consistent with the previously established range of positions to which σ^{70}-dependent paused complexes backtrack on λpR' (16). Thus, the observation that the λpR' and λpR'17N25$_A$ QAC backtrack to the same positions indicates that they remain bound at the σ^{70}-dependent pause regardless of the position to which they have transcribed.

To confirm our interpretation of the nucleotide starvation cleavage results, we examined the 3' cleavage products that are generated when the QAC are treated with GreB: If λpR' and λpR'17N25$_A$ QAC backtrack to positions from +11 to +16, then treatment with GreB should yield 3' cleavage products of 9 to 14 and 14 to 20 nts, respectively. We generated QAC on template bound to magnetic beads, washed to remove NTPs, and incubated in the presence of GreB and unlabeled NTPs to stimulate cleavage and chase the 5' cleavage products into longer RNAs. As expected, cleavage of the QAC on λpR' yields RNA of ~9 – 14 nt in length, and cleavage of the QAC on λpR'17N25$_A$ yields RNA of ~14 – 20 nt in length (Figure 2.7B). It is important to note that the long 3' cleavage products are Q^{λ}-dependent, indicating that they originate from the QAC that contain long RNAs, thus, confirming the interpretation that the QAC backtrack to a cluster of positions from +11 to +16 regardless of the position to which they transcribe.

Figure 2.7 QAC Backtrack to +15 Regardless of the Positions to Which They Have Transcribed

(A) Cleavage of the λpR' and λpR'17N25$_A$ QAC. After formation of the QAC, NTPs were removed and GreB was added. Samples were incubated at 37°C and removed at the indicated times. 5' and 3' cleavage products were differentiated by the assay shown in (B) and Figure S3B.

(B) 3' Cleavage Products of the λpR' and λpR'17N25$_A$ QAC. As in (A), except transcription was resumed in the presence of GreB and NTPs. The exact sizes of 3' cleavage products are uncertain within about a nucleotide because they have 5'-monophosphate rather than 5'-triphosphate ends like the other transcripts. However, the approximate sizes of the λpR' and λpR'17N25$_A$ 3' cleavage products are both distinct and within the expected ranges.


A.
λpR'
λpR'17N25A
+30
+25
+15
Time (min)
0 0.25 0.5 1 5
B.
λpR'
λpR'17N25A
Qλ
GreB
-
+
+30
+25
+17
3'
3'

Together, the Q^{λ}-dependent nature of these complexes and the position to which they backtrack indicate that they result from Q^{λ}-modified elongation complexes that attempt and fail to escape from the σ^{70}-dependent pause. The observation that the λpR‘ and λpR‘17$N25_{A}$ QAC originate from the σ^{70}-dependent pause, yet transcribe to different positions before backtracking, establishes that the sequence composition immediately following the σ^{70}-dependent pause is a critical determinant of Q^{λ}-dependent arrest and is consistent with a mechanism of DNA scrunching.

A systematic mutant scan reveals a sequence element responsible for the long abortive products of the $N25_{Anti}$ promoter and the extended Q-dependent arrest

The ability of the ITS to change the position to which Q^{λ}-modified elongation complexes transcribe before backtracking and arrest indicates that a sequence element within the $N25_{Anti}$ ITS is responsible for determining the extent of scrunching that occurs during escape from the σ^{70}-dependent pause and the $N25_{Anti}$ promoter. To reveal any such elements we performed a systematic mutant scan that covered the positions +17 to +31 of the λpR‘17$N25_{A}$ template (Figure 2.8A). The scan consisted of overlapping 3 bp substitutions in which we made A/C and G/T transversions. We then performed single-round *in vitro* transcription to assay for altered Q^{λ}-dependent arrest. Strikingly, two overlapping mutants showed a decrease in the amount of complexes arrested at +30 and the appearance of complexes at +24, 1 nt shorter than the wild-type position of +25 (Figure 2.8B). These mutants shared G to T mutations in positions +21 and +22 (positions +5 and +6 of the $N25_{Anti}$ ITS). To establish that the +21 and +22 nt are

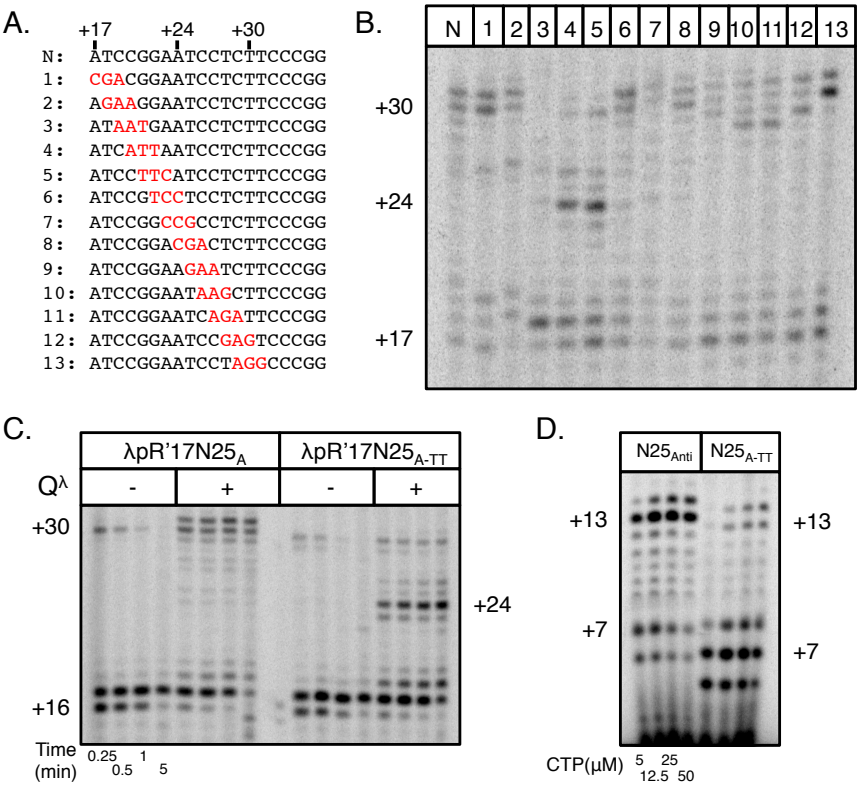
Figure 2.8 A Dinucleotide Motif is Required for the Abortive and Arrest Properties of the $N25_{Anti}$ ITS.

(A) Mutants used in the scan of λpR'17$N25_{A.}$

(B) Single round *in vitro* transcription from the mutants in (A). Mutant 3 displays increased arrest at +18, which may account for the decreased downstream QAC formation. Mutants 4 and 5 display an increased tendency to form QAC at +24

(C) Single round *in vitro* transcription from λpR'17$N25_{A}$ and λpR'17$N25_{A\text{-}TT}$ in the presence and absence of Q^{λ}.

(D) Multi-round *in vitro* transcription from p$N25_{Anti}$ and p$N25_{A\text{-}TT}$.


A.
+17
+24
+30
N: ATCCGGAATCCTCTTCCCGG
1: CGACGGAATCCTCTTCCCGG
2: AGAAGGAATCCTCTTCCCGG
3: ATAATGAATCCTCTTCCCGG
4: ATCATTAATCCTCTTCCCGG
5: ATCCTTCATCCTCTTCCCGG
6: ATCCGTCCTCCTCTTCCCGG
7: ATCCGGCCGCCTCTTCCCGG
8: ATCCGGACGACTCTTCCCGG
9: ATCCGGAAGAATCTTCCCGG
10: ATCCGGAATAAGCTTCCCGG
11: ATCCGGAATCAGATTCCCGG
12: ATCCGGAATCCGAGTCCCGG
13: ATCCGGAATCCTAGGCCCGG
B.
N 1 2 3 4 5 6 7 8 9 10 11 12 13
+30
+24
+17
C.
λpR'17N25A
λpR'17N25A-TT
Qλ
- + - +
+30
+24
+16
Time (min) 0.25 0.5 1 5
D.
N25Anti
N25A-TT
+13
+7
+13
+7
CTP(μM) 5 12.5 25 50

entirely responsible for the restoration of a shorter Q^{λ}-dependent arrest, we constructed λpR'17N25$_{A\text{-}TT}$, a λpR'17N25$_{A}$ mutant containing the +21 and +22 G to T mutations. As expected, in the presence of Q^{λ}, the majority of QAC clustered around position +24 rather than +30 (Figure 2.8C).

We then performed the mutant scan described above on λpR' from positions +16 to +25 to determine whether a similar sequence element existed within the wild-type λpR' sequence (Figure 2.9A). The introduction of a G-rich region at positions near +21 and +22 shifted the primary position of arrest from +25 to +27, with some complexes appearing at +31 and +32 (Figure 2.9B). Thus, introducing the GG sequence element from N25$_{Anti}$ is sufficient to make the arrest pattern of λpR' similar to that of λpR'17N25$_{A}$; in fact, no other portion of the N25$_{Anti}$ sequence has significant effect on the arrest, although the exact pattern of arrest sites varies slightly. (Figures 2.9C and 2.9D).

If there is a relationship between the effect of the N25$_{Anti}$ ITS on Q^{λ}-dependent arrest and the abortive properties of the N25$_{Anti}$ promoter, mutations that disrupt the extended Q^{λ}-dependent arrest of λpR'17N25$_{A}$ should affect the long abortive transcripts of the N25$_{Anti}$ promoter in a similar manner. To test this, we introduced G to T mutations at positions +5 and +6 of the N25$_{Anti}$ promoter. Multi-round *in vitro* transcription revealed that the N25$_{A\text{-}TT}$ mutant primarily accumulates abortive products of 6, 7, and 8 nt and displays a dramatic reduction of 13 and 14 nt aborted transcripts (Figure 2.8D). The effect of the +5, +6 G to T mutations is enhanced at low nucleotide concentrations, but persists up to the standard concentration of our transcription reactions. This shift is analogous to the shift observed in the QAC of the λpR'17N25$_{A\text{-}TT}$ mutant, further

Figure 2.9 A Mutant Scan of λpR'

(A) λpR' mutants used in (B). Overlapping three bp A/C and G/T transversions were made to cover positions +16 to +25.

(B) Single round *in vitro* transcription from the λpR' mutants in (A). Mutants 7 and 8 move the QAC from the wild-type cluster at +25 to clusters at +27 and +31. These mutants share a pair of G to T mutations at +22 and +23.

(C) λpR' mutants used in (D). Mutant 1 contains G to T mutations at +22 and +23 that create a G-rich region from +21 to +24. Mutants 2 through 6 contain the mutation in 1 and also introduce sequence from the $N25_{Anti}$ ITS.

(D) Single round *in vitro* transcription from λpR'. The G-rich region in mutant 1 is sufficient to move the QAC from +25 to +31 and +32. The $N25_{Anti}$ sequence introduced in mutants 2 - 6 does not substantially change the length of the RNA in the QAC.

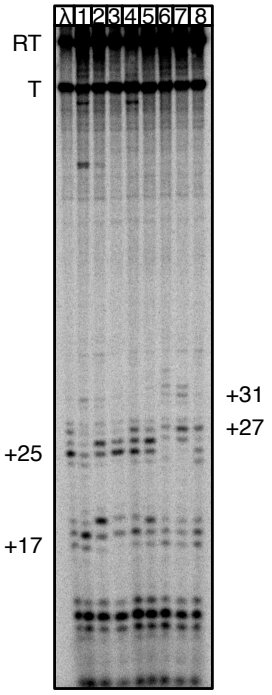
A.

```
          +17   +22+25
λ:  CGCTCGTTGTG
1:  ATATCGTTGTG
2:  CTAGCGTTGTG
3:  CGAGAGTTGTG
4:  CGCGATTTGTG
5:  CGCTATGTGTG
6:  CGCTCTGGGTG
7:  CGCTCGGGTTG
8:  CGCTCGTGTGG
```

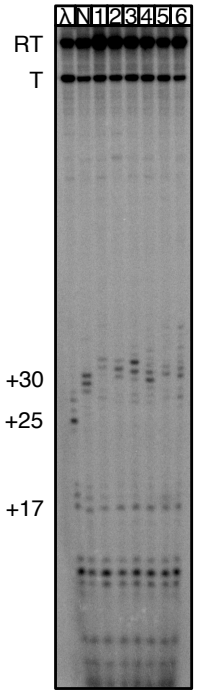
C.

```
          +17   +22+25   +30
λ:  GCTCGTTGTGGTAGTGAGAT
N:  ATCCGGAATCCTCTTCCCGG
1:  GCTCGGGGTGGTAGTGAGAT
2:  GCTCGGGATCGTAGTGAGAT
3:  GCTCGGGATCCTCGTGAGAT
4:  GCTCGGGATCCTCTTCAGAT
5:  GCTCGGGGTGGTATTCAGAT
6:  GCTCGGGGTGCTCTTCAGAT
```

B.



D.



indicating that the abortive properties of the $N25_{Anti}$ ITS are reflected in its effects on Q^{λ}-dependent arrest and thus supporting the model that scrunching is the mechanism by which Q^{λ}-modified elongation complexes escape the σ^{70}-dependent pause. In fact, a scan by overlapping 3-nt substitutions in the $N25_{Anti}$ ITS confirms that only the +5 and +6 doublet is important to the abortive pattern (Figures 2.10A and 2.10B). [We ignore the transcripts of mutants 1,2, and 3, which are obscured by stuttering due to the T run in the beginning of the transcript (36)].

The +21/+22 mutation acts primarily through the template strand

The template and non-template strands of a DNA duplex have distinct functions and contacts within a transcribing complex. Such asymmetry means that the effect of a mutation can be assigned to the template, non-template, or both strands of DNA. It is possible to differentiate the activity of each strand by analyzing the behavior of heteroduplex templates. The availability of a 2-bp mutation that generates arrest (in the context of $\lambda pR'17N25_{A}$) and abortive (in the context of $N25_{Anti}$) patterns that are easily distinguishable from those of the original $N25_{Anti}$ sequence allowed us to create heteroduplex templates and evaluate the strandedness of the $N25_{Anti}$ +5T, +6T mutant effects.

Single-round *in vitro* transcription was performed on $\lambda pR'17N25_{A}$, $\lambda pR'17N25_{A\text{-}TT}$, and the corresponding heteroduplexes. Figure 2.11A shows that the template strand dictates the position of Q^{λ}-dependent arrest: when $\lambda pR'17N25_{A}$ is in the template strand Q^{λ}-dependent arrest is identical to that of the $\lambda pR'17N25_{A}$ homoduplex, whereas when

Figure 2.10 Mutant Scan of pN25$_{Anti}$.

(A) Initial transcribed sequences of pN25$_{Anti}$ mutants that introduce the sequence "TTT" at all positions from +1 to +12.

(B) Multi-round *in vitro* transcription from the pN25$_{Anti}$ mutants in (A). Transcription was performed at 20 μM UTP, 20 μM GTP, 20 μM CTP, 5 μM ATP and labeled internally with α-^{32}P-ATP. The ladders generated by mutants 1, 2, and 3 are the result of a phenomenon known as stuttering that occurs as a result of transcription from a promoter with a T-rich start site (36) and no inference can be made about these. Mutants 4 and 5 contain the overlapping G to T mutations at positions +5 and +6, and are prone to aborting synthesis at positions 6 and 7, resulting in decrease in longer transcripts. Mutants 6 through 12 do not display significant deviation from the abortive pattern of N25$_{Anti}$.

A.

```
          +1     +7     +13
N:  ATCCGGAATCCTCTTCCCGG
1:  TTTCGGAATCCTCTTCCCGG
2:  ATTTGGAATCCTCTTCCCGG
3:  ATTTTGAATCCTCTTCCCGG
4:  ATCTTTAATCCTCTTCCCGG
5:  ATCCTTTATCCTCTTCCCGG
6:  ATCCGTTTCCTCTTCCCGG
7:  ATCCGGTTTCCTCTTCCCGG
8:  ATCCGGATTTCTCTTCCCGG
9:  ATCCGGAATTTCTTCCCGG
10: ATCCGGAATTTCTTCCCGG
11: ATCCGGAATCTTTTCCCGG
12: ATCCGGAATCCTTTCCCGG
```

B.

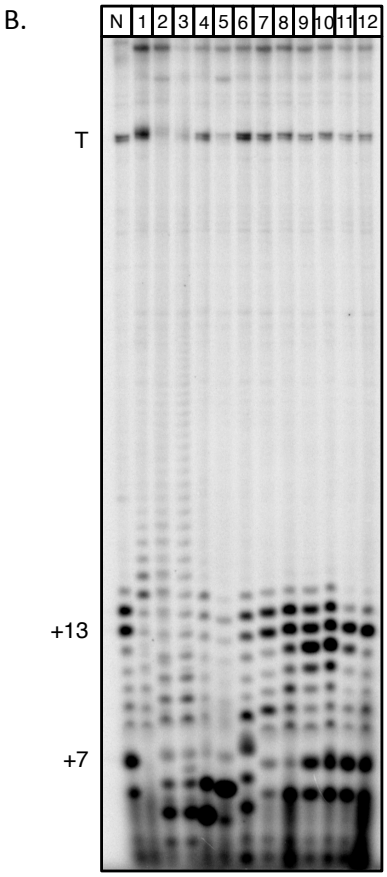
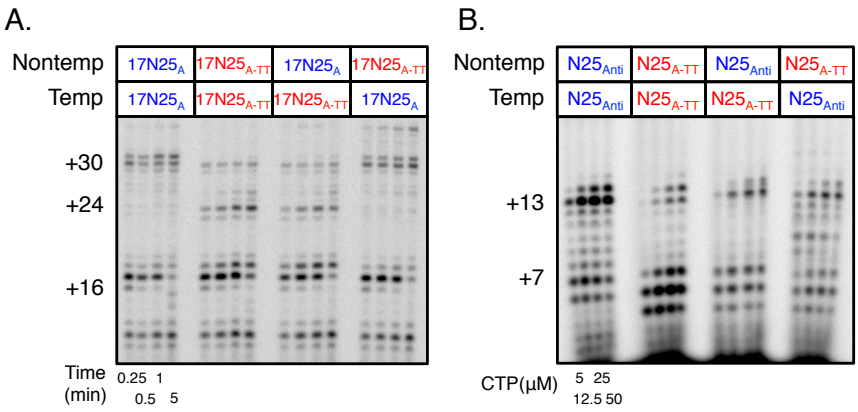


Figure 2.11 The +5, +6 G to T Mutations Act Primarily Through the Template Strand

(A) Single round *in vitro* transcription from λpR'17N25$_{A}$, λpR'17N25$_{A\text{-}TT}$, and the corresponding heteroduplexes.

(B) Single round *in vitro* transcription from pN25$_{Anti}$, pN25$_{Anti\text{-}TT}$, and the corresponding heteroduplexes.


A.
Nontemp
Temp
17N25A
17N25A-TT
17N25A
17N25A-TT
17N25A
17N25A-TT
17N25A-TT
17N25A
+30
+24
+16
Time (min)
0.25 1
0.5 5
B.
Nontemp
Temp
N25Anti
N25A-TT
N25Anti
N25A-TT
N25Anti
N25A-TT
N25A-TT
N25Anti
+13
+7
CTP(μM)
5 25
12.5 50

λpR'17N25$_{A\text{-}TT}$ is in the template strand, Q^{λ}-dependent arrest is identical to that of the λpR'17N25$_{A\text{-}TT}$ homoduplex.

We then performed a heteroduplex analysis of N25$_{Anti}$ and N25$_{A\text{-}TT}$ to determine the strandedness of the +5T, +6T mutations in the context of abortive initiation (Figure 2.11B). Neither heteroduplex produces an abortive profile like that of the parental homoduplexes. Moreover, heteroduplexes containing a mutant non-template strand form a unique abortive profile, indicating that the ITS non-template strand is also a determinant of abortive transcription. The contribution of the non-template strand in the context of abortive initiation and the absence of such an effect in the context of Q^{λ}-dependent arrest suggests that while the underlying mechanism is shared, differences between a promoter-bound initial transcribing complex and a Q^{λ}-modified σ^{70}-dependent paused complex affect the ways in which ITS composition modulates escape.

Strengthening the interaction between σ^{70} and the -10 like element increases the quantity and alters the distribution of both Q^{λ}-dependent arrest and σ^{70}-dependent pausing.

In addition to the ITS, the strength of the interaction between RNAP and the promoter elements plays a role in the formation of abortive products, with stronger promoters tending to yield more aborted transcripts (14). If the mechanism of Q^{λ}-dependent arrest is in fact analogous to abortive initiation, strengthening the interaction between σ^{70} and the -10-like element should increase the likelihood of Q^{λ}-modified

transcription elongation complexes undergoing Q^{λ}-dependent arrest. To test this, we created a λpR' mutant, λpR'-10Like$_{Cons}$, that contains the -10 element consensus sequence 'TATAAT' in place of the wild-type 'AACGAT' in positions +1 to +6 (Figure 2.12A). In λpR', substitution of a consensus -10-like element in place of the wild-type sequence creates an especially strong interaction with σ^{70} due to the presence of a GGG 'discriminator-like' sequence (37). Single-round *in vitro* transcription reveals that a consensus -10-like element increases the quantity and alters the distribution of both Q^{λ}-dependent arrest and σ^{70} pausing (in addition to shifting the ladder down 1 nt because initiation occurs at the A of position 2 relative to wild-type) (Figures 2.12B and 2.13A).

The effect of the λpR'-10Like$_{Cons}$ on Q^{λ}-modified elongation complexes is to broaden the range of species captured during Q^{λ}-dependent arrest from 17-19 and 24-28 so that QAC appear at all positions from 17-28 (Figures 2.12B and 2.13A). We propose that the expanded range of QAC on λpR'-10Like$_{Cons}$ is the result of the strengthened interaction between σ^{70} and the -10-like element trapping complexes that would normally bypass Q^{λ}-dependent arrest and causing them to collapse into a backtracked state. We constructed a -10-like consensus mutant of λpR'17N25$_A$, termed λpR'17N25$_A$-10Like$_{Cons}$, to test whether strengthening the interaction between σ^{70} and the -10-like element would have similar effects on mutant QAC (Figure 2.12C). As expected, λpR'17N25$_A$-10Like$_{Cons}$ expands the QAC from 17-19 and 29-31 (on λpR'17N25$_A$) to 17-34 (Figure 2.12D). It is striking that the pattern of Q-dependent arrest that λpR'17N25$_A$-10Like$_{Cons}$ displays from +17 to +30 (Figure 2.12D) is very similar to the abortive release pattern of N25$_{Anti}$ from +1 to +14 (Figure 2.8D), which is

Figure 2.12 A Consensus -10-Like Element Expands the Formation of QAC.

(A) Sequence comparison of λpR' and λpR'-10Like$_{Cons}$ from positions +1 to +28. The λpR'-10Like$_{Cons}$ alternate transcription start site is indicated by an asterisk.

(B) Single round *in vitro* transcription from λpR' and λpR'-10Like$_{Cons}$ in the presence and absence of Q^{λ}. Transcripts were labeled internally with α-^{32}P-UTP and removed at the indicated times. Positions for λpR' are indicated on the left, and positions for λpR'-10Like$_{Cons}$ are indicated on the right and are designated in reference to the wild-type start site.

(C) Sequence comparison of λpR'17N25$_A$ and λpR'17N25$_A$ -10Like$_{Cons}$ from positions +1 to +36. λpR'17N25$_A$ -10Like$_{Cons}$ alternate transcription start site is indicated by an asterisk.

(D) Single round *in vitro* transcription from λpR'17N25$_A$ and λpR'17N25$_A$ -10Like$_{Cons}$ in the presence of Q^{λ}. Transcripts were labeled internally with α-^{32}P-UTP and removed at the indicated times. Positions for λpR'17N25$_A$ are indicated on the left, and positions for λpR'17N25$_A$ -10Like$_{Cons}$ are indicated on the right and are designated in reference to the wild-type start site.

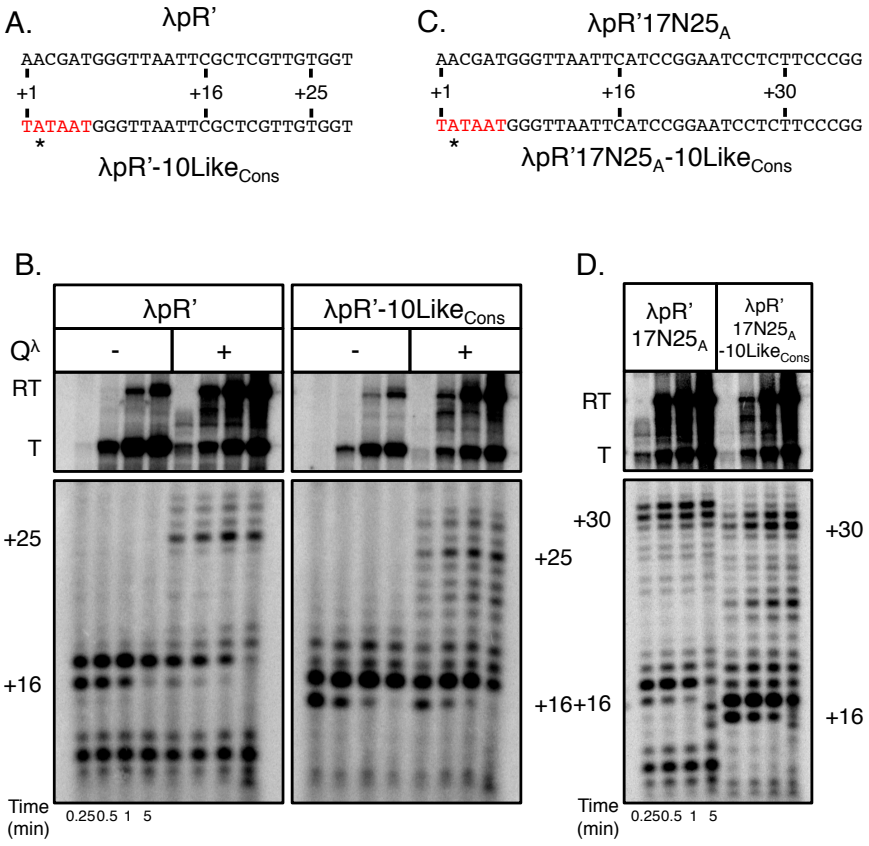
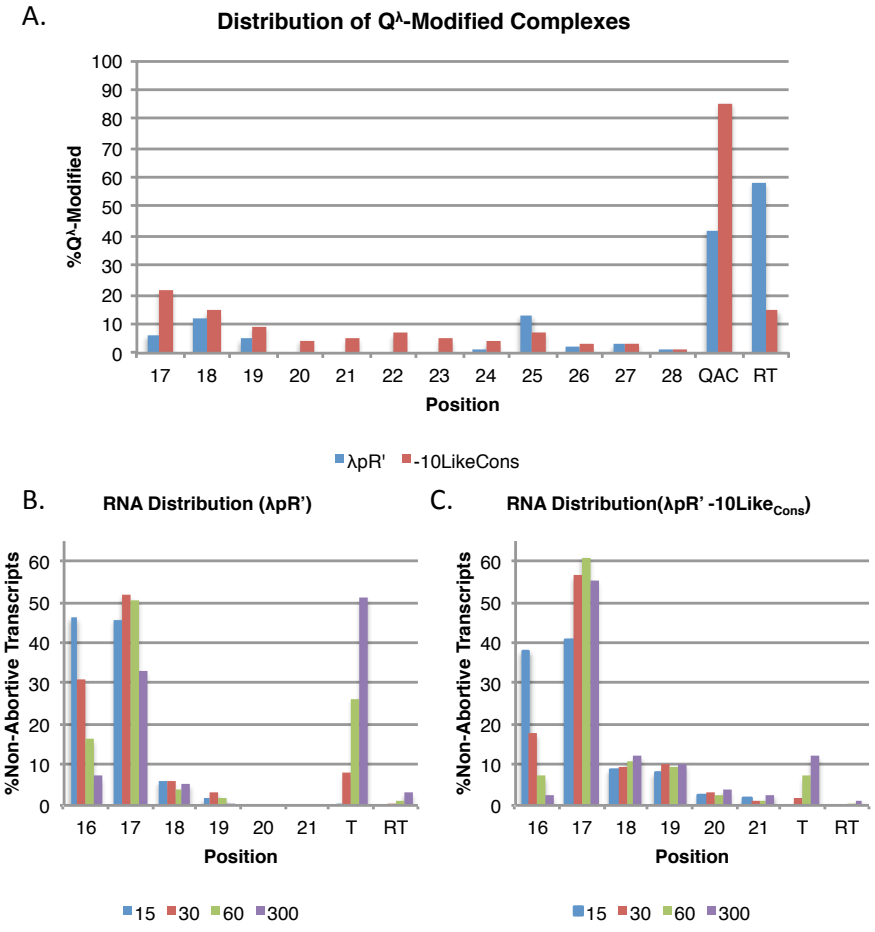

A.
λpR'
AACGATGGGTTAATTCGCTCGTTGTGGT
+1
+16
+25
TATAATGGGTTAATTCGCTCGTTGTGGT
*
λpR'-10Like$_{Cons}$
C.
λpR'17N25$_A$
AACGATGGGTTAATTCATCCGGAATCCTCTTCCCGG
+1
+16
+30
TATAATGGGTTAATTCATCCGGAATCCTCTTCCCGG
*
λpR'17N25$_A$-10Like$_{Cons}$
B.
λpR'
λpR'-10Like$_{Cons}$
Q$^{\lambda}$
-
+
-
+
RT
T
+25
+16
Time (min)
0.25 0.5 1 5
D.
λpR' 17N25$_A$
λpR' 17N25$_A$ -10Like$_{Cons}$
RT
T
+30
+25
+16+16
+30
+16
Time (min)
0.25 0.5 1 5

Figure 2.13 A Consensus -10-Like Element Alters the Distribution of Q-modified and σ^{70}-Dependent Paused Complexes

(A) Distribution of Q^{λ}-Modified Complexes on λpR' and λpR'-10Like$_{Cons}$. Quantification was performed on the 5 minute time point of Fig. 7B for λpR' and λpR'-10Like$_{Cons}$ in the presence of Q^{λ}.

(B) Distribution of Transcripts on λpR' in the absence of Q^{λ}.

(C) Distribution of Transcripts on λpR'-10Like$_{Cons}$ in the absence of Q^{λ}.


A.
Distribution of Q^λ-Modified Complexes
%Q^λ-Modified
100
90
80
70
60
50
40
30
20
10
0
17 18 19 20 21 22 23 24 25 26 27 28 QAC RT
Position
λpR' -10LikeCons
B.
RNA Distribution (λpR')
%Non-Abortive Transcripts
60
50
40
30
20
10
0
16 17 18 19 20 21 T RT
Position
15 30 60 300
C.
RNA Distribution(λpR' -10Like_Cons)
%Non-Abortive Transcripts
60
50
40
30
20
10
0
16 17 18 19 20 21 T RT
Position
15 30 60 300

the same sequence of DNA; note particularly the peak of arrest at +23 of λpR'17N25$_A$-10Like$_{Cons}$, which corresponds to +7 in N25$_{Anti}$.

In the absence of Q$^{\lambda}$, elongation complexes undergo σ^{70}-dependent pausing on both λpR' and the λpR'-10Like$_{Cons}$ mutant, as shown by bands at +16 and +17. λpR'-10Like$_{Cons}$ captures complexes with greater efficiency, as would be expected from a mutation that increases affinity for σ^{70} (Figures 2.13B and 2.13C). In addition to an increase in pausing, λpR'-10Like$_{Cons}$ captures several-fold more complexes that have transcribed to +18 and +19 as well. This effect appears to be similar, albeit less dramatic, to the broadening of Q$^{\lambda}$-dependent pausing in the same mutant.

Discussion

We have shown that two determinants of the efficiency and pattern of promoter escape and abortive initiation, namely the nucleotide composition of the ITS and the strength of the interaction between σ^{70} and DNA, also underlie Q$^{\lambda}$-dependent escape from the σ^{70}-dependent pause at λpR' and the formation of the QAC. We conclude that both Q$^{\lambda}$-dependent escape from the σ^{70}-dependent pause and escape of initial transcribing complexes from the promoter share a common mechanism of DNA scrunching. Thus, we have shown that DNA scrunching occurs in a context outside of promoter escape and, furthermore, can be modulated by a trans-acting factor, in this case the Q$^{\lambda}$-antiterminator. In addition, we identify a sequence element that is an important determinant of the pattern of abortive release of RNAs.

Mechanism of Q^{λ}–dependent release of paused complexes into elongation

At λpR', substitution of the $N25_{Anti}$ ITS in place of the WT sequence from +17 to +36 changes the position to which Q^{λ}-modified complexes transcribe before undergoing Q^{λ}-dependent arrest, from a cluster focused at +25 to a cluster focused at +30 and +31. However, on both templates RNAP remains bound to the -10-like element because the active center is found at +15 in both the WT and mutant QAC. Thus the composition of the 20 bp immediately downstream of the σ^{70}-dependent pause site dictates the position of the 3' end of the RNA bound within the QAC, but does not affect the position to which RNAP backtracks. These characteristics are consistent with a mechanism of scrunching, in which RNA is synthesized without movement of RNAP along DNA, but instead DNA is melted and drawn into the enzyme.

The role of the ITS in abortive synthesis is well defined and it is clear that the base composition of positions +1 to +20 of transcribed DNA dictates the positions at which abortive transcripts are released. Similarly, sequences downstream of the σ^{70}-dependent pause determine where Q^{λ}-dependent arrest occurs; we suggest that in both processes DNA sequence sets the transcription pattern by determining the stability of scrunched complexes. The relationship between the effects of the $N25_{Anti}$ ITS on abortive synthesis and Q^{λ}-dependent arrest supports the interpretation that Q^{λ} promotes escape of RNAP from the σ^{70}-dependent pause as an advancing scrunched complex, accumulating energy of scrunching that is used to break σ^{70} or σ^{70} and Q^{λ} interactions with DNA and release modified RNAP into downstream elongation.

Increasing the strength of the interaction between RNAP and the -10-like element of the λpR' σ^{70}-dependent pause, by substituting the wild-type sequence AACGAT with the consensus -10-element TATAAT, expands the region in which elongation complexes are captured, resulting in Q^{λ}-dependent arrest from +17 to +28 instead of the wild-type positions of +24 to +28. Similarly, when a consensus -10-like element is introduced into λpR'17N25$_A$, Q^{λ}-dependent arrest occurs at all positions between +17 and +34, as opposed to arrest primarily at +17 to +19 and a cluster around +30 in the presence of the wild-type λpR' -10-like element. The changed pattern of arrest between wild-type and consensus -10-like elements gives insight into the natural process of Q^{λ}-dependent escape from the σ^{70}-dependent pause. We propose the following interpretation.

First, we note that the consensus -10-like element increases Q^{λ}-dependent arrest relative to the wild-type -10-like element in regions near the +16 pause, namely +17, +18, and +19, and also invokes novel arrest sites in the region +20 to +23. What is the origin of this enhanced arrest? It must represent complexes that with the wild-type pause-inducing sequence would either continue scrunching or break the σ^{70} and Q^{λ} bonds and continue into productive Q^{λ}-modified elongation. However, it is implausible that the consensus -10-like element would disfavor continuing the scrunch, because it should only increase the strength of the σ^{70}–DNA bond that must be maintained during the scrunch. Thus, it follows that the increase in arrest due to the consensus element must reflect complexes that, when bound to the wild-type -10-like element, have sufficient scrunching energy to release σ^{70} and Q^{λ} bonds with DNA and proceed into

productive Q^{λ}-modified elongation. In the presence of a consensus -10-like element, the scrunching energy required for pause escape is increased so that these complexes no longer have sufficient energy to break the interaction between σ^{70} and the DNA, and instead, backtrack and arrest. Thus the sites of enhanced arrest with the consensus -10-like element mark natural sites where σ^{70} dissociates from the -10-like element and where productive Q^{λ}-modified elongation begins.

On exceeding some limit of stability, the scrunched complex becomes prone to collapse and backtracks with high probability, whether the wild-type or consensus -10-like element is present, resulting in Q^{λ}-dependent arrest at RNA lengths 24-28. As evidenced above, this limit is sensitive to the nature of the 20 bp of sequence immediately following the site of the σ^{70}-dependent pause; for example, introduction of the $N25_{Anti}$ ITS changes the site of arrest to lengths around +30. Thus, these results reveal the natural process of Q^{λ}-dependent escape into elongation.

Nature of abortive initiation

In addition to the similarities between an initial-transcribing complex and a σ^{70}-dependent paused complex, there are also discrete differences, most notably the contacts between σ^{70} and the core, the presence of a mature RNA, and, when present, the contacts with Q^{λ}. The availability of a second system besides promoter initiation in which RNAP must break an interaction between σ^{70} and DNA to proceed forward provides a structurally distinct context in which to examine the elements that contribute

to promoter escape. Such an analysis can clarify the role of these elements and may lead to a more general understanding of abortive synthesis.

The existence of abortive initiation has been attributed to two distinct phenomena: scrunching and displacement of the $\sigma^{3.2}$ loop by the emerging RNA product (10-13). A further barrier that could contribute to abortive initiation is the displacement of σ^4 that occurs as the RNA reaches 14-15 nt in length (38). All of these processes would store energy as the initial transcript grows, and the failure of any to continue could lead to abortive loss of RNA. However, neither $\sigma^{3.2}$ loop displacement nor σ^4 displacement can be involved in QAC formation because the RNA is too long, leaving scrunching as the only plausible process to account for QAC formation; therefore, our demonstration that the pattern of abortive release is reconstructed in the pattern of QAC formation, specifically in the case of $\lambda pR'17_{N25A}\text{-}10Like_{Cons}$, suggests that scrunching is the dominant process that forms the pattern of abortive synthesis at this promoter

Our investigation of the QAC has provided a pathway to understanding a sequence basis of the pattern of abortive transcript release. Previous work has shown, first, that stronger promoter consensus elements (e.g. -35, -10, and discriminator elements) increase the length and yield of abortive products (14), and, second, that the ITS determines both the pattern and level of abortive release (15). It was shown previously that there is a correlation between purine-richness of the ITS and abortive tendency (15), but no specific sequence elements have been identified. We have used

a systematic scan across the $N25_{Anti}$ ITS to find elements that affect QAC formation, and then to make inferences about ITS function in abortive transcript release.

A pair of G to T mutations at positions +5 and +6 of the $N25_{Anti}$ ITS (+21 and +22 in λpR'17$N25_A$) shifts the position to which RNAP transcribes before Q^{λ}–dependent arrest from a cluster at +30 to a cluster at +24, indicating that the identity of these 2 nt is essential to the properties of the $N25_{Anti}$ ITS. The critical nature of these positions is reflected by their function in abortive initiation in the $N25_{Anti}$ promoter: Introduction of the +5, +6 G to T mutations into $N25_{Anti}$ alters abortive synthesis in a manner equivalent to the effect of the corresponding changes in λpR'17$N25_A$ on Q^{λ}-dependent arrest. The bulk of $N25_{Anti}$ abortive products are 7,13, and 14 nt in length, with minor abortive products occurring at all positions between +2 and +15. The mutant $N25_{A\text{-}TT}$ promoter displays strongly decreased abortive transcripts at +13 and +14, and instead accumulates many more abortive transcripts of 6, 7, and 8 nt in length. A reasonable conjecture is that the uridine-richness at +5, +6 near the end of the 6,7, and 8 transcripts destabilizes the DNA/RNA hybrid and promotes dissociation. The decrease in 13 and 14 nt aborted RNAs in $N25_{A\text{-}TT}$ is likely the result of increased abortive release at +6, +7, and +8 reducing the number of transcripts that proceed to +13 and +14; the ratio of transcripts at +12 to +15 to full-length transcripts, i.e the probability of abortive stalling at +12 to +15, is constant across the overlapped set of 3-nt substitutions that drastically changes the ratio of +12 to +15 to smaller abortive transcripts. The abortive pattern is sensitive to NTP concentration: low NTP concentrations favor the accumulation short abortive products, whereas high NTP concentrations facilitate the

formation of longer abortive products and partly restore the 13 and 14 nt abortive transcripts in $N25_{A\text{-}TT}$. This trend implicates the rate of transcription as a key determinant of abortive synthesis and suggests that other influences on transcription rate, e.g interaction with DNA-bound factors, influence promoter escape and abortive synthesis.

It is not surprising that the differences between these complexes would produce subtle variations in the way DNA scrunching is modulated by the ITS. This is evident in a heteroduplex analysis of the $N25_{Anti}$ and $N25_{A\text{-}TT}$ ITSs in both contexts. Transcription on λpR'17$N25_{A}$ and λpR'17$N25_{A\text{-}TT}$ heteroduplexes reveals that the effect of ITS composition on formation of the QAC is solely dependent on the template strand. However, in the context of initial transcription at $N25_{Anti}$ and $N25_{A\text{-}TT}$, neither heteroduplex produces an abortive pattern identical to either of the homoduplexes from which they are derived. Thus, the effect of ITS composition on the QAC is clearly attributable to the template strand, whereas both the template and non-template strands modulate abortive initiation. It is plausible that this difference is a reflection of the numerous structural distinctions between a σ^{70}-dependent paused complex and an initial-transcribing complex that affect the interactions between RNAP and the DNA template.

References

1. McClure, W.R., Cech, C.L. and Johnston, D.E. (1978) A steady state assay for the RNA polymerase initiation reaction. *J Biol Chem*, **253**, 8941-8948.

2. Carpousis, A.J. and Gralla, J.D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry*, **19**, 3245-3253.

3. Hsu, L.M., Vo, N.V., Kane, C.M. and Chamberlin, M.J. (2003) In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry*, **42**, 3777-3786.

4. Goldman, S.R., Ebright, R.H. and Nickels, B.E. (2009) Direct detection of abortive RNA transcripts in vivo. *Science*, **324**, 927-928.

5. Kwak, H. and Lis, J.T. (2013) Control of Transcriptional Elongation. *Annu Rev Genet*.

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

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

8. Grayhack, E.J., Yang, X.J., Lau, L.F. and Roberts, J.W. (1985) Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell*, **42**, 259-269.

9. Ring, B.Z., Yarnell, W.S. and Roberts, J.W. (1996) Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, **86**, 485-493.

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

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

12. Murakami, K.S., Masuda, S. and Darst, S.A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science*, **296**, 1280-1284.

13. Samanta, S. and Martin, C.T. (2013) Insights into the Mechanism of Initial Transcription in Escherichia coli RNA Polymerase. *J Biol Chem*, **288**, 31993-32003.

14. Vo, N.V., Hsu, L.M., Kane, C.M. and Chamberlin, M.J. (2003) In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. *Biochemistry*, **42**, 3798-3811.

15. Hsu, L.M., Cobb, I.M., Ozmore, J.R., Khoo, M., Nahm, G., Xia, L., Bao, Y. and Ahn, C. (2006) Initial transcribed sequence mutations specifically affect promoter escape properties. *Biochemistry*, **45**, 8841-8854.

16. Marr, M.T. and Roberts, J.W. (2000) Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol Cell*, **6**, 1275-1285.

17. Marr, M.T., Datwyler, S.A., Meares, C.F. and Roberts, J.W. (2001) Restructuring of an RNA polymerase holoenzyme elongation complex by lambdoid phage Q proteins. *Proc Natl Acad Sci U S A*, **98**, 8972-8978.

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

19. Perdue, S.A. and Roberts, J.W. (2010) A backtrack-inducing sequence is an essential component of Escherichia coli sigma(70)-dependent promoter-proximal pausing. *Mol Microbiol*, **78**, 636-650.

20. Bird, J. (2013), Characterization of novel DNA elements necessary for sigma-dependent promoter proximal transcriptional pausing. Ph.D. Thesis, Cornell University.

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

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

23. Deighan, P., Diez, C.M., Leibman, M., Hochschild, A. and Nickels, B.E. (2008) The bacteriophage lambda Q antiterminator protein contacts the beta-flap domain of RNA polymerase. *Proc Natl Acad Sci U S A*, **105**, 15305-15310.

24. Deighan, P. and Hochschild, A. (2007) The bacteriophage lambdaQ anti-terminator protein regulates late gene expression as a stable component of the transcription elongation complex. *Mol Microbiol*, **63**, 911-920.

25. Yang, X.J. and Roberts, J.W. (1989) Gene Q antiterminator proteins of Escherichia coli phages 82 and lambda suppress pausing by RNA polymerase at a rho-dependent terminator and at other sites. *Proc Natl Acad Sci U S A*, **86**, 5301-5305.

26. Hsu, L.M., Vo, N.V. and Chamberlin, M.J. (1995) Escherichia coli transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression in vivo and in vitro. *Proc Natl Acad Sci U S A*, **92**, 11588-11592.

27. Ko, D.C., Marr, M.T., Guo, J. and Roberts, J.W. (1998) A surface of Escherichia coli sigma 70 required for promoter function and antitermination by phage lambda Q protein. *Genes Dev*, **12**, 3276-3285.

28. Guo, J. and Roberts, J.W. (2004) DNA binding regions of Q proteins of phages lambda and phi80. *J Bacteriol*, **186**, 3599-3608.

29. Marr, M.T. and Roberts, J.W. (1997) Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, **276**, 1258-1260.

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

31. Perederina, A.A., Vassylyeva, M.N., Berezin, I.A., Svetlov, V., Artsimovitch, I. and Vassylyev, D.G. (2006) Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription elongation factors GreB from Escherichia coli and Gfh1 from Thermus thermophilus. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **62**, 44-46.

32. Schmidt, M.C. and Chamberlin, M.J. (1984) Amplification and isolation of Escherichia coli nusA protein and studies of its effects on in vitro RNA chain elongation. *Biochemistry*, **23**, 197-203.

33. Ring, B.Z. and Roberts, J.W. (1994) Function of a nontranscribed DNA strand site in transcription elongation. *Cell*, **78**, 317-324.

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

35. Kammerer, W., Deuschle, U., Gentz, R. and Bujard, H. (1986) Functional dissection of Escherichia coli promoters: information in the transcribed region is involved in late steps of the overall process. *EMBO J*, **5**, 2995-3000.

36. Turnbough, C.L., Jr. (2011) Regulation of gene expression by reiterative transcription. *Curr Opin Microbiol*, **14**, 142-147.

37. Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C. and Gourse, R.L. (2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell*, **125**, 1069-1082.

38. Nickels BE, Garrity SJ, Mekler V, Minakhin L, Severinov K, Ebright RH, and Hochschild A. (2005) The interaction between sigma70 and the beta-flap of Escherichia coli RNA polymerase inhibits extension of nascent RNA during early elongation. *Proc Natl Acad Sci U S A*, **102**, 4488-4493.

CHAPTER THREE:

FUNCTION OF A TRANSLOCATION PAUSE ELEMENT AT A REGULATORY TRANSCRIPTION PAUSE

Summary

The movement of RNA polymerase (RNAP) during transcription elongation is modulated by DNA-encoded elements that cause the elongation complex to pause. The translocation pause element (TPE) has been found to induce transcription pausing throughout the *Escherichia coli* genome and to underlie regulatory pause elements such as the *ops* and σ^{70}-dependent pause. Here, we show that, in the context of σ^{70}-dependent pausing, the TPE functions primarily by modulating the rate at which complexes attempt to disengage a σ^{70}:DNA interaction. Further, we identify a novel element within the TPE. Our findings establish the σ^{70}-dependent pause-encoding region as a compound element in which two pause-inducing components make distinct mechanistic contributions to induce and maintain a regulatory transcription pause.

Introduction

Transcription pausing has emerged as an important element of gene regulation at the level of RNA synthesis in both prokaryotes and eukaryotes. The mechanisms by which nucleic acid sequence elements and transcription factors induce transcriptional pausing are varied, but are unified in that they must accomplish two functions: a pause element must first induce capture of an elongation complex, and second, must retain

that complex at a defined position on the DNA template for a meaningful period of time. A prototypical example of regulatory transcription pausing exists in the lambdoid phage late gene operon where expression of the controlled genes is dependent on the conferral of terminator-resistance to RNA polymerase by the incorporation of the antiterminator Q at a pause site during early elongation (1,2). The terminator-resistant elongation complex can then "read through" an intrinsic terminator that precedes the coding sequence of the operon (2). The assembly of Q into complex with RNAP is mediated by a pause dependent on the σ^{70} initiation factor that positions the elongation complex such that it is a substrate for the binding and incorporation of Q (3,4).

The eponymous element of the σ^{70}-dependent pause-encoding region, the -10-like sequence, halts transcription through a binding event between σ^{70} and the non-transcribed DNA strand and establishes the major properties of σ^{70}-dependent pausing (3,4). Namely, the σ^{70}:DNA interaction confers upon the paused complex the essential qualities of an initial transcribing complex by inducing the formation of a "scrunched" DNA structure in which the upstream edge of RNAP is maintained at a defined position while the active center continues to synthesize RNA (5,6). Maintenance of the upstream edge of the elongation complex positions RNAP such that Q is able to bind to both a DNA Q binding element upstream of RNAP and to elements within RNAP (including specifically the β flap), allowing Q to assemble into the elongation complex (7,8). The binding of Q drives escape from the pause by promoting the formation of an extended scrunched state, which is proposed to provide the energy necessary for rapid

disengagement of σ^{70} from the -10-like sequence and likely Q from its DNA binding site (6).

In addition to the -10-like sequence, a transcribed sequence motif spanning from the position immediately downstream of the -10-like sequence through the pause site segment was identified as an essential component of the 82pR' σ^{70}-dependent pause (9). Although this motif was described as a backtrack-inducing sequence, a property it does display, it primarily acts as a translocation pause element (TPE), a pause-inducing sequence that is believed to function by inhibiting the progression of RNAP through the nucleotide addition cycle by favoring a "pre-translocated state" in which the 3' terminus of the nascent RNA is maintained in the active center i+1 site, rendering it unavailable for binding and incorporation of the incoming nucleoside triphosphate (10-13). A "GC-rich" segment and the pause site nucleotide of the σ^{70}-dependent pause-encoding region correspond to determinants of the TPE and play a critical role in the establishment of a σ^{70}-dependent pause (13). Importantly, the location of the TPE has been shown to define the position at which the pause occurs (13).

In this work, we have determined that the DNA segment internal to the TPE is critical to the function of the TPE in σ^{70}-dependent pausing. Furthermore, we show that a major function of the TPE in the context of σ^{70}-dependent pausing is to modulate the rate of pause escape. This insight leads us to propose a refined model for σ^{70}-dependent pausing in which two distinct pause elements act in conjunction to capture and maintain RNAP in a paused state during early elongation. In our model, the -10-like sequence is responsible for inducing the capture of RNAP and plays an important role in

determining the probability of a successful attempt at pause escape. The TPE plays an important role in determining the frequency at which a paused complex enters into a conformation in which an escape attempt can be made, namely, the postranslocated state. Thus, σ^{70}-dependent pausing uses two distinct pause elements to accomplish the capture and retention of elongation complexes.

Materials and Methods

Plasmids

pM650 (14), p'QE-30 (15), and pET-28a- σ^{70} (16) have been described. pES3 is *GreB*-6xHis in pET-28b(+). pVS10 (a gift from I. Artsimovitch) contains all four core RNAP subunits. All mutants were constructed using Quikchange site-directed mutagenesis.

Proteins

RNA polymerase (17), GreB, σ^{70} (16), and NusA (19) were purified as described.

In Vitro transcription

Reaction mixtures containing 5 nM template and 25 nM RNAP (25 nM core reconstituted with 50 nM σ^{70}) were incubated in transcription buffer (20 mM Tris-HCL pH 8.0, 0.1 mM EDTA, 1 mM dithiothreitol and 50 mM KCl), 0.1 mg/ml bovine serum albumin, a nucleotide mix of either 200 µM ATP, GTP, CTP and 50 µM UTP (containing 0.5 µCi/µl [α-^{32}P]-UTP) or 200 µM ATP, UTP, CTP and 50 µM GTP (containing 0.5 µCi/µl α-^{32}P-GTP) and 150 nM NusA for 15 min at 37°C to form open complex. Single-

round transcription reactions were initiated by addition of magnesium chloride to 5 mM and rifampicin to 10 µg/ml. Total reaction volume was 25 µl. When present, GreB was added to 200 nM before formation of open complex. Reactions were stopped by adding 125 µl of stop solution (0.6 M Tris pH 8.0, 12 mM EDTA, 0.16 mg/ml transfer RNA).

Heteroduplex templates

Heteroduplex templates were constructed as described previously and sequenced to confirm purity (3). Transcription was performed as described above.

Purification, fractionation, and analysis of transcription reactions

Stopped transcription reactions were phenol extracted by addition of 150 µl of phenol/chloroform/isoamyl alcohol (25:24:1), vortexing, centrifugation, and collection of the aqueous phase. Ethanol precipitation of RNA was performed by adding 450 µl of 100% ethanol to each reaction, followed by storage at -20 °C overnight. Precipitated RNA was resuspended in transcription loading dye (1x T buffer, 80% formamide, 0.05% bromophenol blue and xylene cyanol). Reactions were fractionated by electrophoresis using 12 or 15% denaturing polyacrylamide gels containing 6M urea. Reactive bases were detected by an Amersham Biosciences Typhoon 9400 Variable Mode Imager. Quantitation was performed using ImageQuant.

Results

The λpR' +10 to +15 region modulates the rate of σ^{70}-dependent pause escape

To assess the role of the TPE internal segment in σ^{70}-dependent pausing, we replaced the λpR' native sequence of 5'-TTAATT-3' with the sequence 5'-CCGGCC-3' to create the mutant "IE[CG]" (Figure 3.1A). This mutation was designed following previous indications that the base composition, and particularly A/T-richness, of this region is important for pausing (9). *In vitro* transcription reactions were carried out in the presence of the cleavage factor GreB to allow complexes that backtrack upon failing to escape to undergo a cleavage event, thus permitting paused complexes to make multiple attempts at escape (5). In the presence of GreB, complexes transcribing IE[CG] displayed a ~3.5-fold reduction in the level σ^{70}-dependent pausing 15 seconds after initiating the reaction, demonstrating that the base composition of the λpR' +10 to +15 region is important to the formation of σ^{70}-dependent paused complexes (Figure 3.1B, C). Whether the reduction of σ^{70}-dependent pausing at IE[CG] is a product of a defect in pause capture, the result of an enhanced rate of escape, or both is not clear from this result due to the substantial difference in the level of σ^{70}-dependent pausing observed at the earliest time point. To overcome this limitation, we created -10-like sequence consensus mutations in both the λpR' and IE[CG] templates (ConsλpR' and ConsIE[CG], respectively), which we anticipated would reduce the rate of pause escape and allow us to observe the decay of σ^{70}-dependent paused complexes from comparable levels at the initial time point (Figure 3.1A). As predicted, in the presence of GreB, at 15 seconds ConsλpR' and ConsIE[CG] had sigma-dependent pause capture

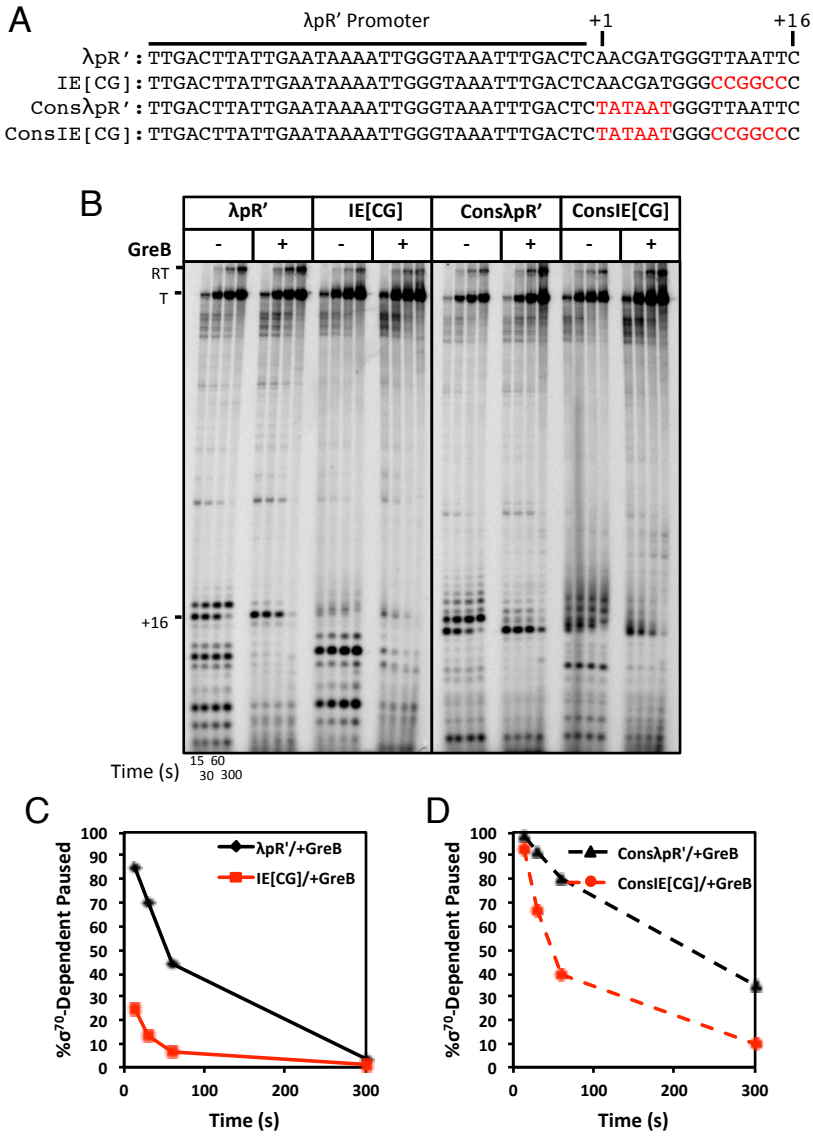
Figure 3.1 The TPE internal element modulates the rate of σ^{70}-dependent pause escape

(A) Sequence comparison of λpR', IE[CG] ConsλpR', and ConsIE[CG].

(B) *In vitro* transcription from λpR', IE[CG] ConsλpR', and ConsIE[CG] in the presence and absence of GreB.

(C) Quantitation of λpR' and IE[CG] in the presence of GreB.

(D) Quantitation of ConsλpR' and ConsIE[CG) in the presence of GreB.


A
λpR' Promoter
+1
+16
λpR': TTGACTTATTGAATAAAATTGGGTAAATTTGACTCAACGATGGGTTAATTC
IE[CG]: TTGACTTATTGAATAAAATTGGGTAAATTTGACTCAACGATGGGCCGGCCC
ConsλpR': TTGACTTATTGAATAAAATTGGGTAAATTTGACTCTATAATGGGTTAATTC
ConsIE[CG]: TTGACTTATTGAATAAAATTGGGTAAATTTGACTCTATAATGGGCCGGCCC
B
λpR'
IE[CG]
ConsλpR'
ConsIE[CG]
GreB
-
+
RT
T
+16
Time (s)
15 30 60 300
C
%σ70-Dependent Paused
λpR'/+GreB
IE[CG]/+GreB
Time (s)
D
%σ70-Dependent Paused
ConsλpR'/+GreB
ConsIE[CG]/+GreB
Time (s)

levels of nearly 100% (Figure 3.1B, D). Given that the wt -10-like sequence exhibits 85% pause capture at 15 seconds and that the IE[CG] mutant is not predicted to interfere with the σ^{70}:DNA interaction, it is plausible that the wt -10-like sequence also engages 100% of elongation complexes and that the increase in σ^{70}-dependent paused complexes observed at 15 seconds on consensus templates is entirely due the a decreased rate of escape. In spite of nearly 100% σ^{70}-dependent pause capture observed at the first time point, complexes paused at ConsIE[CG] undergo pause escape at a rate ~3X faster than those paused at ConsλpR' (Figure 3.1D). These data show that the sequence composition of the λpR' +10 to +15 segment modulates the rate of σ^{70}-dependent pause escape.

A detailed mutant scan of the λpR' +10 to +15 region

To identify specific elements within the GC-rich segment of IE[CG] that interfere with σ^{70}-dependent pausing, we performed a mutant scan that consisted of overlapping three, two, and one bp mutations in which the IE[CG] sequence was restored to wt. The nomenclature of these templates is *n*x*p*, where *n* refers to the number of bases restored and *p* refers to the position of the first restored base. For example, 3x10 refers to a template in which positions 10, 11, and 12 have been restored from the IE[CG] sequence to the wt sequence. We chose to perform this mutant scan using templates that contain the wt λpR' -10-like sequence to initially identify IE[CG] derivatives that restore the observed level of σ^{70}-dependent pausing to near-wt levels. Notable mutants identified using this method could then be examined for effects on pause half-life using

a -10-like sequence consensus variant. Of the fifteen scan mutants, we identified four that were noteworthy for their ability to restore the percentage of σ^{70}-dependent paused elongation complexes observed at 15 seconds following the start of the reaction to levels approaching that of λpR' (Figure 3.2A). Strikingly, the scan mutants that restore the level of σ^{70}-dependent pausing most effectively contain mutations that are clustered within a three bp stretch of the +10 to +15 segment, from +12 to +14.

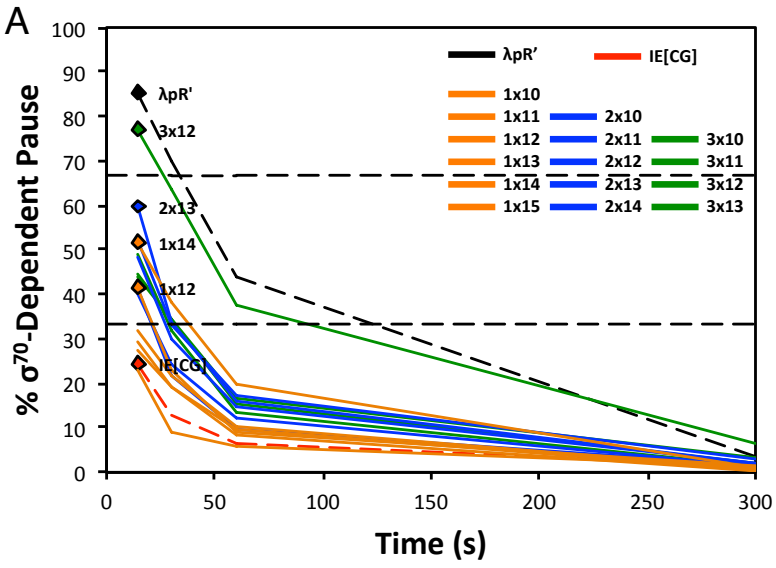
Our analysis of the restoring scan mutants focuses on the level of σ^{70}-dependent pausing at 15 seconds as an indicator of the rate of pause escape. The values described below refer to the fraction of complexes that are engaged in a σ^{70}-dependent pause at 15 seconds. Of the fifteen mutants tested, 3x12 was most effective at restoring σ^{70}-dependent pausing to near wt levels (λpR': 85% vs. 3x12: 77%) (Figure 3.2A). In contrast, none of the other three bp mutants restore pausing to a level above 49%. Of the two bp mutants, 2x13 restores pausing to 60%, whereas all other two bp mutants clustered between 40% and 53% (Figure 3.2A). Notably, 3x12 and 2x13 both restore positions 13 and 14 to wt. Finally, four of the six one bp behaved similarly to the parental IE[CG] mutant. The two exceptions, 1x12 and 1x14, displayed 42% and 52% pausing respectively (Figure 3.2A). Once again, these two mutations are included in the 3x12 mutation.

Our mutant scan successfully identified three positions within the IE[CG] mutant that, when restored to the wt sequence, restore the level of σ^{70}-dependent pausing observed at 15 seconds to wt levels. To determine if this subset of the IE[CG] mutation is sufficient to produce the effect of mutating the entire +10 to +15 segment, we created

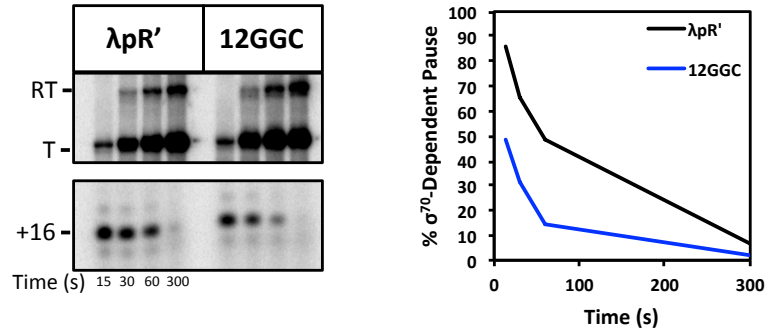
Figure 3.2 A detailed mutant scan of IE[CG]

(A) Percent σ^{70}-dependent paused complexes of IE[CG] restoring scan mutants. Scan mutants with notable pause restoring effects are labeled and shown beneath the graph.

(B) *In vitro* transcription from λpR' and 12GGC in the presence of GreB.



λpR':AACGATGGGTTAATTC (85%)
3x12:AACGATGGGCCAATCC (77%)
2x13:AACGATGGGCCGATCC (60%)
1x14:AACGATGGGCCGGTCC (52%)
1x12:AACGATGGGCCAGCCC (42%)
CG:AACGATGGGCCGGCCC (24%)



a λpR' derivative in which IE[CG] mutant sequence is substituted in place of wt sequence from +12 to +14. The resulting mutant, termed '12GGC', reduced the level of pausing observed at 15 seconds from 86% in the case of wt λpR', to 49% (Figure 3.2B). Thus, while substituting these positions in the IE[CG] mutant with wt sequence is sufficient to restore wt levels of pausing, this subset of the IE[CG] mutation is insufficient to reproduce the full effect of its parental mutation. This result is consistent with the observation that two and three bp substitutions outside the +12 to +14 segment can also significantly affect the function of IE[CG].

σ^{70}-dependent pause restoring mutations function by slowing the rate of pause escape

A subset of the IE[CG] derivatives described above restores the level of σ^{70}-dependent pausing observed at 15 seconds close to that of wt. As demonstrated above, by using a consensus -10-like sequence to raise the energetic barrier to escape so that the initial pause of fast escape mutants is observable, we were able to determine that the IE[CG] mutant reduces the observed level of σ^{70}-dependent pausing by increasing the rate at which paused elongation complexes escape the pause (Figure 3.1B, D). Because the effect of the IE[CG] mutant is to increase the rate of pause escape, we predicted that IE[CG] derivatives that restored σ^{70}-dependent pausing do so by slowing the rate of pause escape. To test this we created consensus -10-like sequence versions of the restoring scan mutants 3x12, 2x13, 1x14, and 1x12 (designated by the

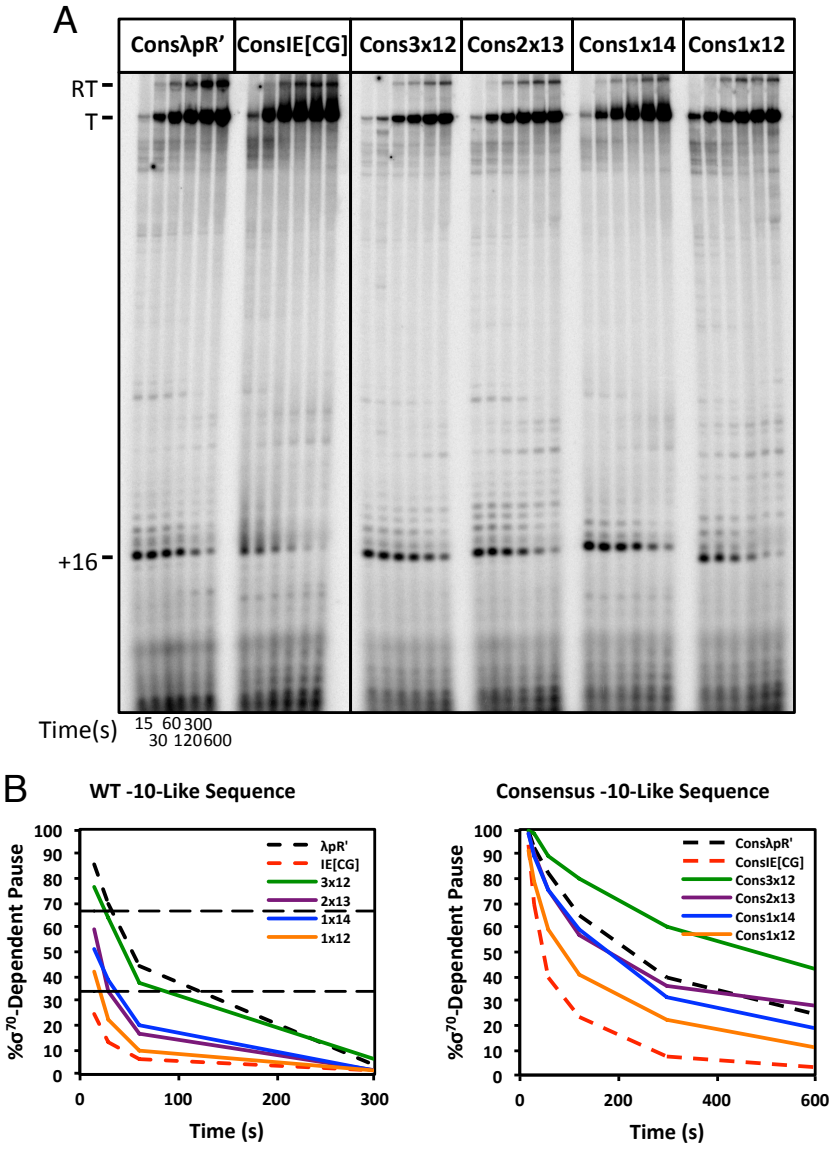
antecedent “Cons”) and performed *in vitro* transcription in the presence of GreB (Figure 3.3A).

As predicted, all of the scan mutant consensus derivatives exhibit both complete capture of the complexes at the pause and a decreased rate of pause escape relative to the parental ConsIE[CG] mutant. Furthermore, the pause half-lives of complexes paused at these mutants reflect the level of σ^{70}-dependent pausing observed at 15 seconds on their non-consensus counterparts: the higher the level of σ^{70}-dependent pausing observed at 15 seconds using non-consensus templates, the slower the rate of escape on their consensus counterparts (Figure 3.3B, C). The only anomaly is the pause escape rate at λpR’ itself, in which the effect of the consensus -10-like sequence on the pause half-life is less relative to its effect on the restoring scan mutants. This observation may be attributable to the complexity of the contributions made by individual sequence elements to the unified σ^{70}-dependent pause element; namely, the cumulative effect of individual pause elements is not necessarily additive. This notion is supported by the observation that ConsIE[CG] and its derivatives, which vary only at positions +12 to +14, perfectly follow the predicted trend that pause level at a non-consensus pause site inversely correlates with escape rate at a consensus pause site. In λpR’, the added variability at positions +10, +11, and +15 complicates its comparison with ConsIE[CG] and the consensus scan mutants. Regardless, the relationship between the restoring scan non-consensus and consensus mutants confirms our observation that the λpR’ +10 to +15 region primarily functions by modulating the rate at which RNAP escapes from the σ^{70}-dependent pause.

Figure 3.3 The level of pausing at non-consensus IE[CG] derivatives reflects the pause half-life

(A) *In vitro* transcription from consensus derivatives of the noteworthy restoring mutants described in Figure 3.2A.

(B) Quantitation from Figure 3.2A and Figure 3.3B. The level of σ^{70}-dependent paused complexes in the presence of a wt -10-like sequence reflects the rate of escape in the presence of a consensus -10-like sequence.


A
ConsλpR'
ConsIE[CG]
Cons3x12
Cons2x13
Cons1x14
Cons1x12
RT
T
+16
Time(s)
15 60 300
30 120 600
B
WT -10-Like Sequence
Consensus -10-Like Sequence
%σ70-Dependent Pause
Time (s)
λpR'
IE[CG]
3x12
2x13
1x14
1x12
ConsλpR'
ConsIE[CG]
Cons3x12
Cons2x13
Cons1x14
Cons1x12

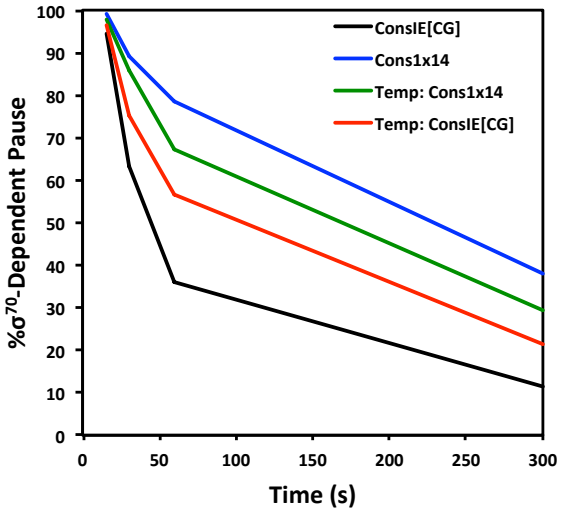
The identification of the IE[CG] derivative 1x14, which dramatically altered the σ^{70}-dependent pause half-life as the result of a point mutation, presented us with an opportunity to examine the non-transcribed and transcribed strand contributions of the IE[CG] mutant using a heteroduplex analysis, because IE[CG] and 1x14 have only a single nucleotide mismatch, which will not significantly disrupt the DNA template structure. *In vitro* transcription of IE[CG] and 1x14 heteroduplexes revealed that the lambda +10 to +15 region contributes to σ^{70}-dependent pausing through both strands of the DNA duplex (Figure 3.4). Complexes paused on both heteroduplex templates display escape rates intermediate to the parental homoduplexes, although the contribution of the template strand is greater than the contribution of the non-template strand. Given the location of the +10 to +15 segment within an elongation complex, it is plausible that the non-template strand contribution of this region to σ^{70}-dependent pausing reflects contacts betwwn RNAP core and scrunched DNA that affect the stability of the pause complex.

The identity of the σ^{70}-dependent pause site nucleotide modulates escape rate

It was previously demonstrated that the identity of the nucleotide at the σ^{70}-dependent pause site is essential to pause formation at 82pR' (13). Here, we use a series of λpR' non-consensus and consensus -10-like sequence mutants that contain point mutations at the pause site (from the wt 'C' to 'G', 'T', and 'A') to reproduce the result observed with 82pR' and to demonstrate that a major role of pause site nucleotide

Figure 3.4 Heteroduplex analysis of ConsIE[CG] and Cons1x14

In vitro transcription from ConsIE[CG], Cons1x14, and the corresponding heteroduplexes.


100
90
80
70
60
50
40
30
20
10
0
%σ70-Dependent Pause
0
50
100
150
200
250
300
Time (s)
ConsIE[CG]
Cons1x14
Temp: Cons1x14
Temp: ConsIE[CG]

identity in σ^{70}-dependent pausing is to modulate the rate of pause escape (Figure 3.5A, 3.6A).

As demonstrated at the 82pR' σ^{70}-dependent pause, if the nucleotide at the λpR' pause site is 'C' or 'T', a strong pause (~80% observed at 15 seconds) is favored (Figure 3.5B, C). If this position is mutated to 'G' or 'A', the level of σ^{70}-dependent pausing observed at 15 seconds is reduced to 48% and 23% respectively (Figure 3.5B, C). The rate of σ^{70}-dependent pause escape observed on consensus versions of the pause site mutants (designated by the antecedent "Cons") reflects the level of pausing observed on the non-consensus mutants: ConsλpR' and Cons16T undergo pause escape at comparable rates whereas Cons16G and Cons16A escape so rapidly that in spite of containing a consensus -10-like sequence they display only 77% and 68% σ^{70}-dependent pause at 15 seconds respectively; nonetheless, these high numbers make it likely that capture approximates 100% in each case (Figure 3.6B, C).These results demonstrate definitively that the identity of the nucleotide at the σ^{70}-dependent pause site plays a critical role in determining the rate at which elongation complexes escape the interaction between the -10-like sequence and σ^{70}.

Discussion

In this work we have identified an element within the σ^{70}-dependent pause-encoding region that, while uninvolved in the σ^{70}:DNA interaction, is instrumental in determining the rate of pause escape. Further, we have shown that the identity of the pause site nucleotide is also a major contributor to pause half-life. These elements are

Figure 3.5 Contribution of the pause site nucleotide to σ^{70}-dependent pausing at λpR'

(A) Comparison of pause site nucleotide mutants.

(B) *In vitro* transcription from λpR', 16G, 16T, and 16A in the presence of GreB.

(C) Quantitation of (B).

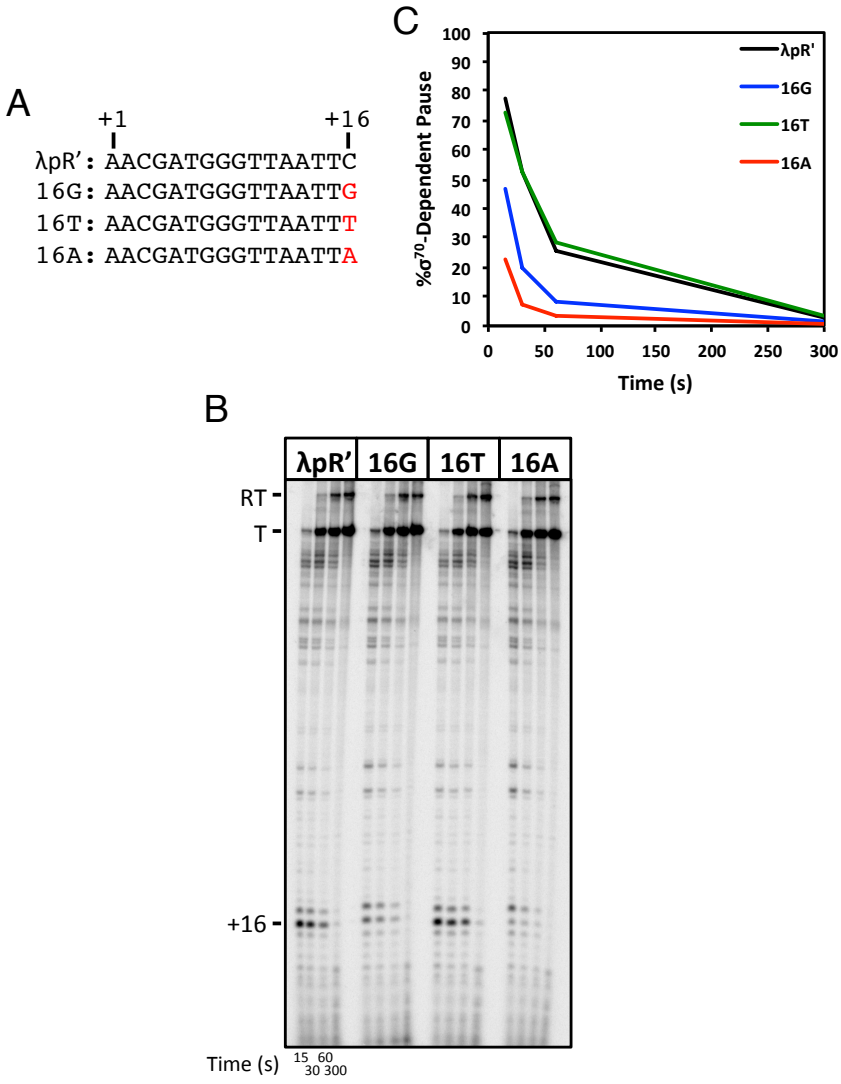
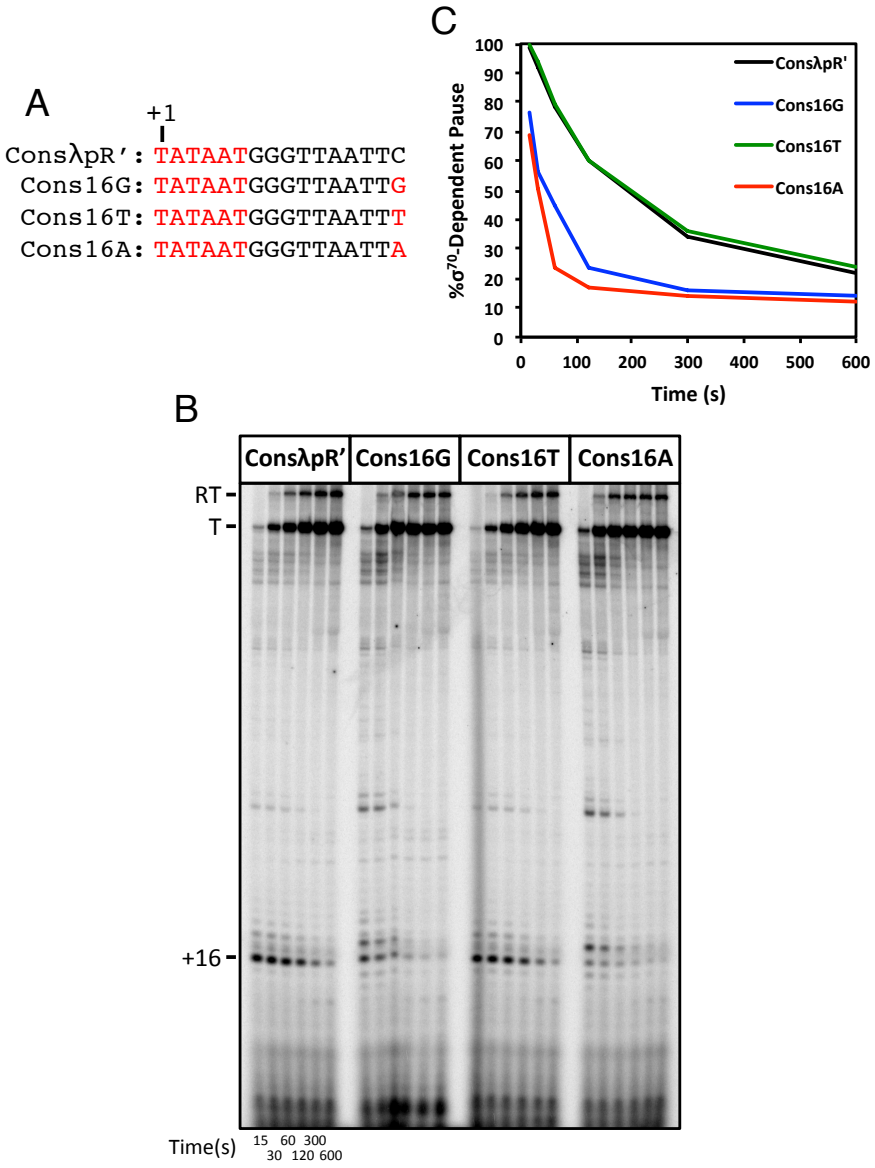

A
+1
+16
λpR': AACGATGGGTTAATTC
16G: AACGATGGGTTAATTG
16T: AACGATGGGTTAATTT
16A: AACGATGGGTTAATTA
C
%σ70-Dependent Pause
Time (s)
λpR'
16G
16T
16A
B
λpR'
16G
16T
16A
RT
T
+16
Time (s) 15 30 60 300

Figure 3.6 The pause site nucleotide modulates the rate of σ^{70}-dependent pause escape

(A) Comparison of consensus pause site nucleotide mutants.

(B) *In vitro* transcription from ConsλpR', Cons16G, Cons16T, and Cons16A in the presence of GreB.

(C) Quantitation of (B).


A
+1
ConsλpR': TATAATGGGTTAATTC
Cons16G: TATAATGGGTTAATTG
Cons16T: TATAATGGGTTAATTT
Cons16A: TATAATGGGTTAATTA
C
%σ70-Dependent Pause
Time (s)
ConsλpR'
Cons16G
Cons16T
Cons16A
B
ConsλpR'
Cons16G
Cons16T
Cons16A
RT
T
+16
Time(s)
15 60 300
30 120 600

components of a TPE, and thus, our results suggest that an important function of the TPE in σ^{70}-dependent pausing is modulation of the rate of pause escape determining the stability at the pause.

A translocation pause element modulates σ^{70}-dependent half-life

At λpR', substitution of the natively AT-rich segment from +10 to +15 with the GC-rich sequence 5'-CCGGCC-3' reduces the level of σ^{70}-dependent pausing observed during *in vitro* transcription. The use of a consensus -10-like sequence to slow the rate of pause escape revealed that this decrease in pausing results not from a defect in pause capture, but rather from decrease in pause half-life. Furthermore, mutation of the pause site nucleotide also dramatically reduces the pause half-life when analyzed in the presence of a consensus -10-like sequence. Specifically, while mutating the wt C at +16 to T does not affect pause escape, substitution of G or A substantially reduces pause half-life.

The AT-rich segment and the pause site nucleotide exist within a TPE that was previously identified as a component of the σ^{70}-dependent pause-encoding region (9,13). The pause site nucleotide in particular is known to be a critical component of the TPE, with pyrimidine favoring and purine disfavoring the pause. The location of these elements in the TPE and the observation that mutation of these sequences changes the pause half-life implicates the TPE in the modulation of the pause escape rate.

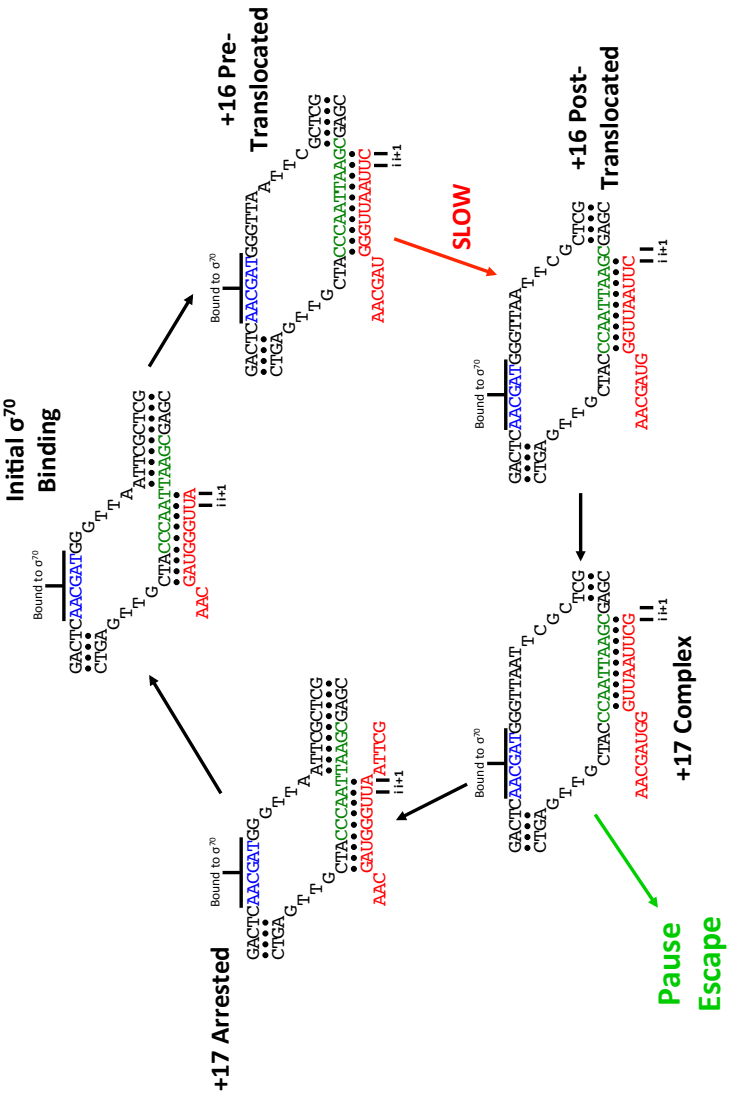
Mechanism of translocation pause element function σ^{70}-dependent pausing

How might the TPE modulate the σ^{70}-dependent pause half-life? To answer this question, we must first consider that failure to escape the pause results in backtracking and arrest (5). Arrested complexes can resume transcription following the catalysis of RNA cleavage by GreA or GreB; however, the preservation of contacts between σ^{70} and the -10-like sequence implies that upon release from the arrested state, the paused complex will resynthesize RNA up to +16 and undergo a translocation pause for the second time (5). If the probability that σ^{70} will disengage the -10-like sequence is low, it is possible that a paused complex can undergo numerous iterations of the σ^{70}-dependent pause cycle prior to escape. Thus, if the strength of the TPE determines the rate at which an escape attempt occurs, the translocation pause would exert a large influence over the σ^{70}-dependent pause half-life (Figure 3.7).

The contribution of the TPE to the pause half-life is particularly exposed when a DNA template containing a consensus -10-like sequence also contains a defect in the TPE. The probability that an elongation complex will escape a σ^{70}-dependent pause element should be modulated by any factor that favors or disfavors the disengagement of σ^{70} from the -10-like sequence. The most immediately apparent factor is the strength of the interaction between σ^{70} and the -10-like sequence: as the strength of the contacts between σ^{70} and the DNA increase, the energy necessary to disrupt this interaction increases. Thus, strengthening the σ^{70}:DNA interaction will decrease the probability of σ^{70} release during a given escape attempt because of an increased energetic requirement for the disengagement of σ^{70} from the -10-like sequence. In the case of ConsIE[CG] and Cons16A, which contain defects in the TPE, an initially high level of

Figure 3.7 The σ^{70}-Dependent Pause Cycle

σ^{70}-dependent pausing is induced when RNAP binds to a -10-like sequence, which anchors the elongation complex to the DNA template. RNAP continues transcribing until it encounters a translocation pause element, which disfavors nucleotide addition and slows the rate at which RNAP can enter into a “post-translocated state” and attempt to escape the pause. Upon entering the post-translocated state, RNAP proceeds to +17 and either escapes the pause by disengaging σ^{70} from the -10-like sequence, or fails to escape and enters into a backtracked state. Cleavage of the backtracked RNA by the transcription factor GreB allows the paused complex to repeat this cycle and attempt to escape the pause once again.


Initial σ70 Binding
+16 Pre-Translocated
SLOW
+16 Post-Translocated
+17 Arrested
+17 Complex
Pause
Escape

σ^{70}-dependent paused complexes escape at a rate of three to four times that of complexes paused at ConsλpR'. It is likely that a substantial percentage of elongation complexes that transcribe the non-consensus counterparts of these mutants, IE[CG] and 16A, escape fast enough that they could be considered transiently bound to the -10-like sequence and do not undergo any apparent pausing event.

A role for the translocation pause element in paused complex stability

Strong evidence suggests that when an elongation complex that has bound to the -10-like sequence reaches the pause site, it exists in a "scrunched" conformation in which downstream DNA has been compressed within RNAP to allow transcription beyond the position at which σ^{70} first engages the -10-like sequence (5,6). An implication of this model is that in order for a scrunched pause to form, the interaction between σ^{70} and the -10-like sequence must not immediately disfavor the continuation of the nucleotide addition cycle. Thus, RNA synthesis will continue until the elongation complex encounters a sequence element that disfavors nucleotide addition (e.g. the TPE), at which point it enters into a paused state. Presumably, the absence of such an element permits RNA synthesis to continue until the scrunched complex can no longer stably proceed forward and either pause escape occurs through rewinding of the upstream edge of the transcription bubble and consequent freeing of σ^{70} from its contacts with the -10-like sequence, or, conversely, arrest occurs through rewinding the downstream edge of the transcription bubble and forcing the elongation complex into a backtracked conformation in which the contacts between σ^{70} and the -10-like sequence

are maintained. The function of the TPE then, would be to stabilize the scrunched complex by preventing RNAP from transcribing beyond the pause site nucleotide. If a complex does transcribe beyond the pause site, it would lose the stability conferred by the TPE and reach a "tipping point" at which it either escapes pause or backtracks and arrests.

The model described above is manifest in the behavior of elongation complexes at the σ^{70}-dependent pause element in the absence of GreB. Such complexes can be divided into two categories: those that are paused and those that are arrested. Paused complexes (e.g. +16 complexes at λpR'; +25 complexes at 82pR') exist in a stable scrunched state at the pause site, which is defined by the TPE. Arrested complexes (e.g. +17 complexes at λpR'; +26 complexes at 82pR') represent a population of elongation complexes that have progressed beyond the pause site to the tipping point, but have failed to complete the escape process and backtracked.

What defines the tipping point? An experiment performed by Bird and Roberts provides some insight into the answer to this question. Insertions between the -10-like sequence and the TPE have been shown to move the location of the pause site downstream so that it always coincides with the position defined by the TPE (13). A subtle implication of this result is that the extent of scrunching alone does not define the tipping point. If the accumulation of scrunching energy beyond a limit of stability were the defining element of the tipping point, insertions that increase the degree of scrunching necessary for RNAP to reach the pause site should disfavor progression to the pause site. This is not the case. Instead, the fact that σ^{70}-dependent pausing will

occur at the pause site defined by the TPE even when an extended scrunch is required suggests that the translocation pause may favor the stabilization of the scrunched complex, and that the defining characteristic of the tipping point is the loss of this stability when RNAP proceeds beyond the pause site.

Function of Q in pause escape

The role of the TPE in σ^{70}-dependent pausing raises an intriguing notion regarding the antipausing activity of the antiterminator Q. Upon incorporation of Q, the paused elongation complex rapidly escapes the σ^{70}-dependent pause in a process that involves the formation of a high energy scrunched complex (2,6). It is plausible that the antipausing activity of Q functions by favoring nucleotide addition and the formation of a post-translocated elongation complex to overcome the barrier imposed by the TPE. Two lines of existing evidence support this model. First, the TPE induces a pre-translocated, implying that the wt σ^{70}-dependent pause is pre-translocated. The 16A mutant is particularly defective in σ^{70}-dependent pausing and pauses primarily at +17 rather than at the native site of +16. This effect may be attributable to a defect in the TPE altering the pause so that the translocation step of the nucleotide addition cycle is no longer disfavored at +16 and RNAP is allowed to proceed to +17. Second, we have previously shown that the interaction between σ^{70} and the -10-like sequence is maintained during the escape of Q-modified elongation complexes and that the disengagement of σ^{70} from DNA is the barrier that these complexes must overcome to undergo pause escape (6). Thus, it follows that Q must drive pause escape by a mechanism that does not involve

directly interfering with the σ^{70}:DNA interaction. Taken together, these data suggest that a possible mechanism for the antipausing activity of Q involves overcoming the TPE by favoring the formation of a post-translocated elongation complex.

References

1. Grayhack, E.J. and Roberts, J.W. (1982) The phage lambda Q gene product: activity of a transcription antiterminator in vitro. *Cell*, **30**, 637-648.

2. Grayhack, E.J., Yang, X.J., Lau, L.F. and Roberts, J.W. (1985) Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell*, **42**, 259-269.

3. Ring, B.Z. and Roberts, J.W. (1994) Function of a nontranscribed DNA strand site in transcription elongation. *Cell*, **78**, 317-324.

4. Ring, B.Z., Yarnell, W.S. and Roberts, J.W. (1996) Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, **86**, 485-493.

5. Marr, M.T. and Roberts, J.W. (2000) Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol Cell*, **6**, 1275-1285.

6. Strobel, E.J. and Roberts, J.W. (2014) Regulation of promoter-proximal transcription elongation: enhanced DNA scrunching drives lambdaQ antiterminator-dependent escape from a sigma70-dependent pause. *Nucleic acids research*, **42**, 5097-5108.

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

8. Deighan, P., Diez, C.M., Leibman, M., Hochschild, A. and Nickels, B.E. (2008) The bacteriophage lambda Q antiterminator protein contacts the beta-flap domain of RNA polymerase. *Proc Natl Acad Sci U S A*, **105**, 15305-15310.

9. Perdue, S.A. and Roberts, J.W. (2010) A backtrack-inducing sequence is an essential component of Escherichia coli sigma(70)-dependent promoter-proximal pausing. *Mol Microbiol*, **78**, 636-650.

10. Vvedenskaya, I.O., Vahedian-Movahed, H., Bird, J.G., Knoblauch, J.G., Goldman, S.R., Zhang, Y., Ebright, R.H. and Nickels, B.E. (2014) Transcription. Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science*, **344**, 1285-1289.

11. Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Nayak, D., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R. and Weissman, J.S. (2014) A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science*, **344**, 1042-1047.

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

13. Bird, J. (2013), Cornell University.

14. Ko, D.C., Marr, M.T., Guo, J. and Roberts, J.W. (1998) A surface of Escherichia coli sigma 70 required for promoter function and antitermination by phage lambda Q protein. *Genes Dev*, **12**, 3276-3285.

15. Guo, J. and Roberts, J.W. (2004) DNA binding regions of Q proteins of phages lambda and phi80. *J Bacteriol*, **186**, 3599-3608.

16. Marr, M.T. and Roberts, J.W. (1997) Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, **276**, 1258-1260.

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

18. Perederina, A.A., Vassylyeva, M.N., Berezin, I.A., Svetlov, V., Artsimovitch, I. and Vassylyev, D.G. (2006) Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription elongation factors GreB from Escherichia coli and Gfh1 from Thermus thermophilus. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **62**, 44-46.

19. Schmidt, M.C. and Chamberlin, M.J. (1984) Amplification and isolation of Escherichia coli nusA protein and studies of its effects on in vitro RNA chain elongation. *Biochemistry*, **23**, 197-203.