6732–6740 *Nucleic Acids Research, 2016, Vol. 44, No. 14* *Published online 20 April 2016*
doi: 10.1093/nar/gkw285

A universal transcription pause sequence is an element of initiation factor σ70-dependent pausing

Jeremy G. Bird, Eric J. Strobel and Jeffrey W. Roberts*

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

Received November 25, 2015; Revised April 5, 2016; Accepted April 7, 2016

ABSTRACT

The *Escherichia coli* σ70 initiation factor is required for a post-initiation, promoter-proximal pause essential for regulation of lambdoid phage late gene expression; potentially, σ70 acts at other sites during transcription elongation as well. The pause is induced by σ70 binding to a repeat of the promoter −10 sequence. After σ70 binding, further RNA synthesis occurs as DNA is drawn (or 'scrunched') into the enzyme complex, presumably exactly as occurs during initial synthesis from the promoter; this synthesis then pauses at a defined site several nucleotides downstream from the active center position when σ70 first engages the −10 sequence repeat. We show that the actual pause site in the stabilized, scrunched complex is the 'elemental pause sequence' recognized from its frequent occurrence in the *E. coli* genome. σ70 binding and the elemental pause sequence together, but neither alone, produce a substantial transcription pause.

INTRODUCTION

Transcription pausing at specific DNA regions is an important step in gene regulation in all organisms (1,2). A particularly well-understood example is the initiation factor σ70-dependent pause that enables modification of *Escherichia coli* RNA polymerase (RNAP) by the bacteriophage lambda gene *Q* antiterminator protein (3,4). Within tens of nucleotides of transcription initiation, but well into the elongation phase, σ70 in the RNAP transcription complex re-binds DNA at a near duplication of the original '−10' promoter binding site, called the '−10-like sequence' (5). This binding anchors a transcription pause that lasts on the order of tens of seconds, and provides a substrate for the Q protein to bind both RNAP and a Q binding element in the upstream DNA (6,7). σ70-dependent promoter-proximal pauses exist also in bacterial transcription (8–10), and σ70 can be shown to act during extended elongation *in vivo* and *in vitro* as well (11–13).

The function of σ70 in this pause is well characterized. Mutations in the −10-like sequence corresponding to mutations that impair function of the promoter −10 sequence also impair pausing (5). A mutationally altered σ70 (L402F) selected to inhibit lambda late gene regulation impairs pausing during transcription *in vitro* and *in vivo* (14), although it is competent for initiation and does not prevent cell growth; a suppressor of *L402F* also enhances pausing (15). In addition, the paused complex has been characterized physically so that its structure is well understood (5,7,16,17).

Although σ70 is required for the pause, the actual site of pausing is not determined directly by the interaction of σ70 and DNA. Binding of σ70 regions 2 and 1.2 to DNA anchors the upstream edge of RNAP in a configuration like that of an open initiation complex; however, the catalytic center remains active, allowing polymerization of RNA to continue for 3–4 nucleotides after σ70 has engaged (16). Analogous to the scrunching that has been demonstrated to occur during initial RNA synthesis from a promoter, downstream DNA at the pause is 'scrunched' into the enzyme as it is unwound (Figure 1A) (16,17).

These findings raise the question of why scrunching in the paused complex stops where it does; what determines the actual RNA end in the paused complex? We show here that a G/C-rich segment at the upstream edge of the templating RNA/DNA hybrid in the transcribing enzyme largely determines the position of the pause in polymerization. The G/C-rich segment is part of a previously recognized universal pause-inducing sequence: the 'elemental pause site' (EPS), which traps the transcription complex from the pre-translocated position (18–20). Consistently, we find that other features characteristic of the EPS, especially the terminal RNA nucleotide, also strongly affect the strength of the σ70-dependent pause.

*To whom correspondence should be addressed. Tel: +1 607 255 2430; Fax: +1 607 255 6249; Email: jwr7@cornell.edu
Present addresses:
Jeremy G. Bird, Departments of Genetics and Chemistry and Chemical Biology, and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA.
Eric J. Strobel, School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853, USA.

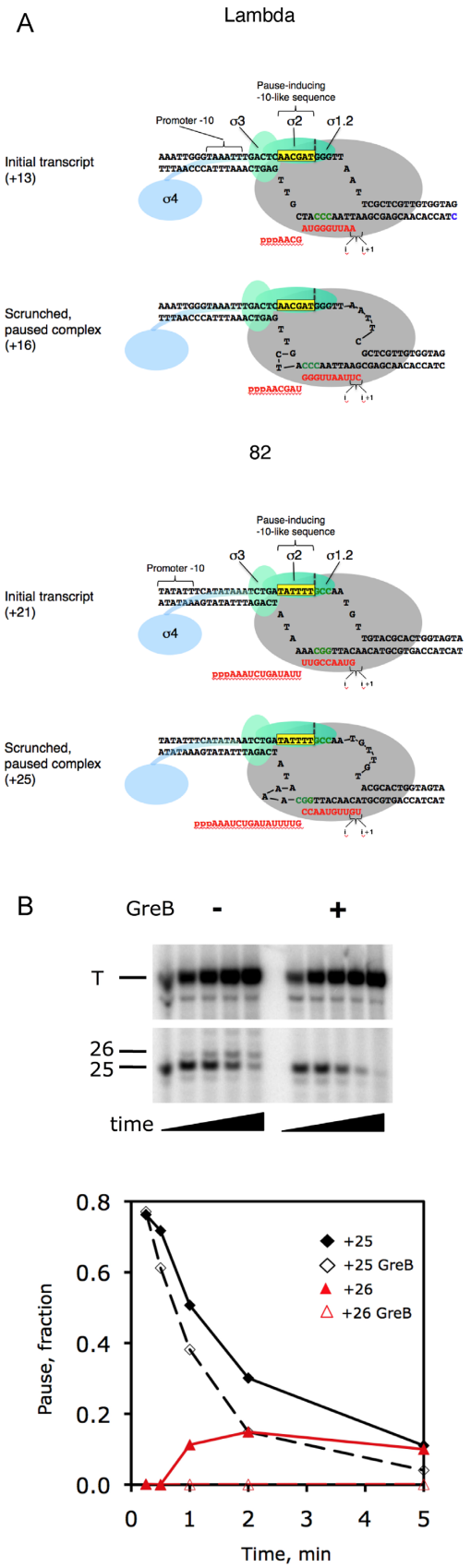

Figure 1. The phage λ and 82 *pR'* σ70-dependent pauses. (**A**) Model of λ and 82 *pR'* σ70-dependent paused states. RNAP escapes the promoter after initiation and begins actively transcribing the gene. The σ70 initiation subunit, still associated with RNAP, then recognizes the −10-like sequence at a transcript length of 13 (λ *pR'*) or 21(82*pR'*) nucleotides. Once the σ70/DNA interaction is made RNAP continues to transcribe several nucleotides downstream, to +16 for λ *pR'* or to +25 for 82*pR'*, 'scrunching' DNA into the complex before actually pausing. (**B**) Single round transcription time course on 82 *pR'* template. Terminated and paused RNAs are shown, and +25 (black) and +26 (red) RNAs are plotted as fraction of total RNA in each lane. Solid lines are +GreB and dashed lines are -GreB.

MATERIALS AND METHODS

Plasmids and DNA templates

Linear DNA templates for transcription were amplified from plasmids by polymerase chain reaction using primers indicated in supplemental materials and purified by gel extraction (Qiagen QIAquick gel extraction kit). All of the DNA templates were created by Agilent Quickchange mutagenesis from a p82*pR'* +2G 1M −35-like mutant plasmid (equivalent to wild-type (WT) for these experiments) or the λ*pR'* pM650 plasmid (21,22). The 82*pR'* +2G mutant was used throughout to prevent initiation stuttering from the WT initial AAA sequence; it is designated simply 82*pR'* in the text. Heteroduplex templates were constructed as described by Ring *et al.* (5).

Proteins

RNAP protein was the gift of M. Suh (Cornell University, Ithaca, NY, USA). The following proteins were purified as previously described: 6×His-σ70 and its L402F and M102A variants (23); GreB (24).

In vitro transcription

Reaction mixtures contained 2 nM DNA template, 10 nM RNAP holoenzyme (10 nM RNAP core reconstituted with 50 nM WT or mutant σ70), 200 mM each of adenosine triphosphate, guanosine triphosphate and cytidine triphosphate, and 50 mM UTP (supplemented with 0.5 mCi/ml [a-^{32}P]-UTP) in Transcription buffer (20 mM Tris–HCl (pH 8.0), 50 mM KCl, 0.1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol, 0.1 mg/ml bovine serum albumin and 5% glycerol). Experiments shown in Figure 1B also contained 100 nM GreB in the reaction mixture. Reaction mixtures were first incubated at 37°C for 10 min to form open complexes. A single round of transcription was then initiated by the addition of $MgCl_2$ to 5mM and rifampicin to 10 μg/ml and incubation continued at 37°C. Aliquots were taken at indicated times after the start of synthesis and mixed on ice with five volumes Stop solution (600 mM Tris–HCl (pH 8.0), 12 mM EDTA and 0.16 mg/ml tRNA). RNA was extracted by mixing with a equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). A total of 2.5 volumes of 100% ethanol was added to the aqueous layer to precipitate samples overnight at −20°C. Samples were run on 12% polyacrylamide gels containing 6M Urea (National Diagnostics Urea gel). RNA species were detected using a Molecular Dynamics Typhoon 9400 and images were analyzed using GE Imagequant software. To determine pause efficiency, transcript quantities were first normalized based

on the number of radiolabeled residues, to enable comparison of numbers of molecules. For all 82pR' pause measurements, the normalized values of the pause bands were then divided by the total of all molecules present in a lane (paused RNAs/(paused RNAs + terminated RNAs)). For the λ*pR'* pause measurement (Supplementary Figure S4), total RNA at 300 s was used for total molecules at all times because too many unquantifiable intermediate pauses were present at earlier times.

RESULTS

Promoter-proximal pausing occurs similarly from λ and 82*pR'* promoters

Comparison of σ70-dependent pausing from the phage λ and phage 82 late gene *pR'* late gene promoters emphasizes its essential and common features (6,25). The −10-like pause inducing sequences are slightly different: the 82 version matches the −10 promoter consensus TATAAT more closely than that of λ, and also includes an extended −10 sequence. The locations of the pause-inducing sequences and the sites of pausing (Figure 1A) differ by 9 bp. Transcription of the λ*pR'* template pauses primarily at +16 and +17 (16), whereas the major pauses from 82*pR'* are +25 and +26, so that the pauses are similarly displaced from the −10-like element (Figure 1A).

When *in vitro* transcription from λ*pR'* is performed in the presence of excess transcription cleavage factor GreB, the pause at +17 disappears whereas the pause at +16 increases somewhat (16). Since GreB mediates cleavage the 3′ overhanging RNA of a backtracked elongation complex (EC) (26,27), the +17 paused complex is backtracked whereas the +16 complex is not (26,27). We envision the +16 complex to be in a stable paused, scrunched state, and that synthesizing one more nucleotide to +17 leads either to pause escape or to destabilization of the scrunched state and backtracking due to the continued interaction of σ70 region 2 with the −10-like element. A λpR' paused complex presumably must synthesize past +16 in order to disrupt the σ70 region 2/−10-like element interaction and escape the pause.

The pausing behavior of transcripts from 82*pR'* is similar: paused species occur at both +25 and +26 in the absence of GreB (Figure 1B). Excess GreB eliminates +26 whereas +25 is only modestly sensitive to GreB, similar to the properties of λ*pR'* +16 and +17 (Figure 1B). The +26 pause of 82*pR'* forms more slowly than the +17 pause of λ, presumably because complexes at +25 of 82*pR'* are somewhat more stable than those at +16 of λ. But the overall similar behavior supports the model that the first site of pausing, 16 or 25, is special in representing a stable, forward scrunched complex, and that addition of one more nucleotide destabilizes it, causing backtracking to the original position of σ70 binding (+13 or +21) (Figure 1A).

Sequence at the site of the discriminator element is necessary for promoter proximal pause formation on 82*pR'*, but, surprisingly, acts primarily through the template strand

The DNA segment of ∼7 bp between the *E. coli* promoter −10 element and the start site of transcription is the 'discriminator', the composition of which is important for the behavior of major classes of promoters (28–32). Within the discriminator, the three nucleotides adjacent to the −10 element constitute the 'discriminator element', which is the binding site for σ70 region 1.2 in the non-template strand of open promoter DNA. Closer matches of the discriminator element to a consensus (approximately GGG) (29,30) yield stronger binding, stabilizing open complex—although the discriminator element of a 'good' discriminator matches the consensus poorly, because its role is to destabilize the open complex (28). Similarly, the discriminator element (when sufficiently close to the consensus) should stabilize the σ70 interaction that anchors the pause; according to match to consensus, the discriminator element associated with the pause -10-like element of λ*pR'* (+7–9 GGG) should bind σ70 strongly and that of 82*pR'* (+15–17 GCC) should bind σ70 less well (Figure 2A). We tested the role of the 82*pR'* +15–17 GCC sequence in σ70-dependent pausing by measuring pausing on phage 82*pR'* templates containing individual and combined mutations of each G and C to A, chosen because of the lower preference for A than G or C in the promoter discriminator element consensus (29). Mutating each base individually significantly reduced the half-life of the pause (∼2-fold), and mutating all three reduced the half-life ∼5-fold. In effect, this sequence is necessary for pause formation (Figure 2B and C).

Next, we determined if the 82*pR'* discriminator element acts in the same manner as a promoter discriminator element sequence, i.e. through a non-template strand interaction with σ70. The +7–9 GGG sequence of the λ*pR'* pause does support pausing primarily through the non-template strand (33). However, we show here that the 82*pR'* pause behaves differently. We transcribed *in vitro* heteroduplex DNAs containing a mutation in either strand of the +15–17 GCC sequence of 82*pR'* DNA and compared the result to homoduplex WT and mutant templates (Figure 3 and Supplementary Figure S1).

Surprisingly, for both the triple mutant (Figure 3) and the +15 and +16 single mutants (Supplementary Figure S1), the non-template strand mutant heteroduplex DNA supported an identical level of pausing as the WT template (Figure 3), whereas in all three cases, pausing supported by the template strand mutant DNA most closely resembled that of the homoduplex mutant. The 82*pR'* +17 C:G to A:T heteroduplex mutants act differently from +15 and +16 mutants, as both strands contribute to the reduction in pause (Supplementary Figure S1). This result is opposite to expectation if this sequence functions through an interaction with σ70 region 1.2, and is different from the result with the λ pause, for which discriminator element mutants act mostly through the non-template strand (33). Consistently, a mutation inactivating σ70 region 1.2 prevents non-template strand function in supporting the λ pause (Filter, J. and Roberts, J., unpublished data).

The σ70 region 1.2/non-template strand interaction on 82*pR'* DNA is still necessary for pause formation

To examine further the relation of the 82*pR'* +15–17 sequence to the discriminator element, we changed it to the λ sequence GGG, the consensus sequence expected to maximize binding of σ70 region 1.2 to the non-template strand

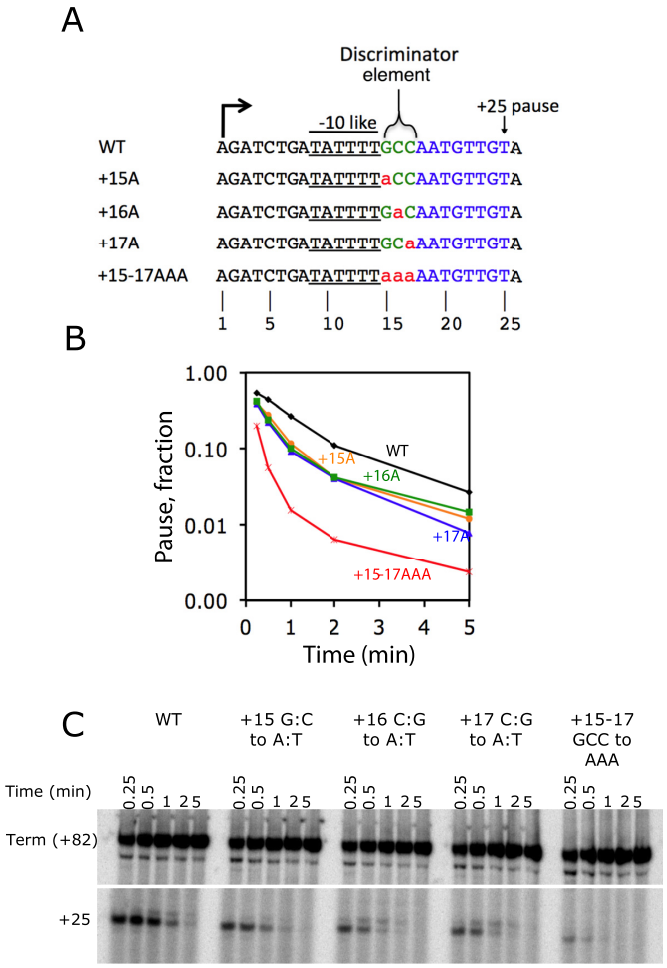

Figure 2. The discriminator element is necessary for promoter proximal pause formation. (**A**) Transcription template sequences from position +1 to +26 for WT and mutant 82 *pR'* templates. The discriminator element sequence is marked in green, the AT-rich region preceding the pause is marked in blue and mutations are in lowercase red. (**B**) The +25/26 paused and terminated RNAs from (**C**) were quantified and paused RNAs plotted as percent total RNA. WT is black; +15–17AAA triple mutant is red; and +15A, +16A and +17A mutants are gold, green and blue respectively. Approximate pause half-lives calculated from the initial parts of the curves are 45 s. (WT), 25 s. (+15A), 22 s (+16A), 22 s (+17A) and 8 s (+15–17AAA). (C) Single round *invitro* transcription time courses of templates from (A).

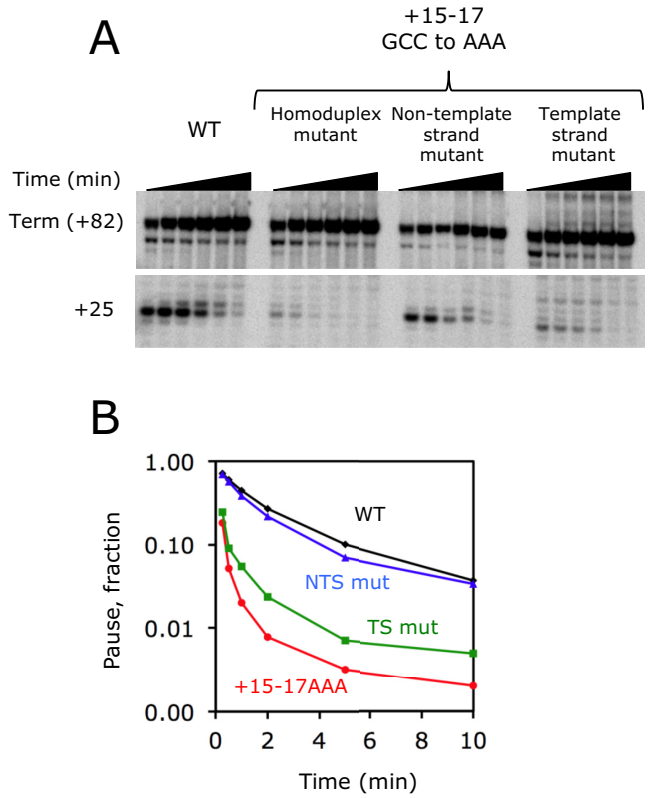

Figure 3. Heteroduplex template analysis reveals that the 82 *pR'* GC-rich discriminator element acts primarily through the template strand. (**A**) Single round *invitro* transcription time courses of WT, +15–17AAA mutant and heteroduplex mutant templates. (**B**) Quantification of the +25/26 paused RNAs as percent of total RNA over time. WT is black, homoduplex mutant is red, non-template strand mutant heteroduplex is blue and template strand mutant heteroduplex is green.

(29). Mutating the 82pR' pause discriminator element to consensus increases the pause lifetime about 4-fold, from 87 sec to 360 s (Figure 4B). Transcription of the heteroduplex DNAs generated from the WT and GGG templates showed that when the discriminator element is changed to consensus, the non-template strand contributes much more significantly to the pause effect, similar to the behavior of λ*pR'*, although both strands still contribute to pausing. This result implies that the σ70 region 1.2 interacts with the 82*pR'* discriminator element non-template strand as expected (see also below), in addition to the clear template strand effect of the sequence. In fact the template strand effect of the GGG mutation is greater than that of WT, consistent with the role of the −10 G in pausing (see 'Discussion' section).

To confirm that σ70 region 1.2 acts in the σ70-dependent pause of 82*pR'*, we used the σ70 M102A mutant, which prevents σ70 region 1.2 interaction with the non-template strand discriminator element (31). We performed *in vitro* transcription of both WT and +15–17 GGG 82*pR'* mutant templates using either WT σ70 or σ70 M102A (Figure 4C). The M102A alteration of σ70 region 1.2 reduces the pause lifetime drastically on both templates, at least for the initial ∼90% of pause decay. (The decay curves throughout are complex, and there appears to be a component (<10%) in each case that is much more stable). Thus although the WT discriminator element GCC is not optimal for σ70 region 1.2 binding, this interaction is nonetheless critical for σ70 function at the pause site.

The discriminator element sequence of the template strand determines the site of the pause

We have shown that the GC-rich discriminator element of the 82*pR'* promoter-associated pause acts substantially through the template strand, in addition to its expected function in the non-template strand. To probe further the function of this sequence in pause formation, we determined the effect of displacing the GC-rich sequence (Figure 5A) away from the −10-like sequence. According to previous understanding that the discriminator element acts only as a spatially constrained component of the σ70 re-

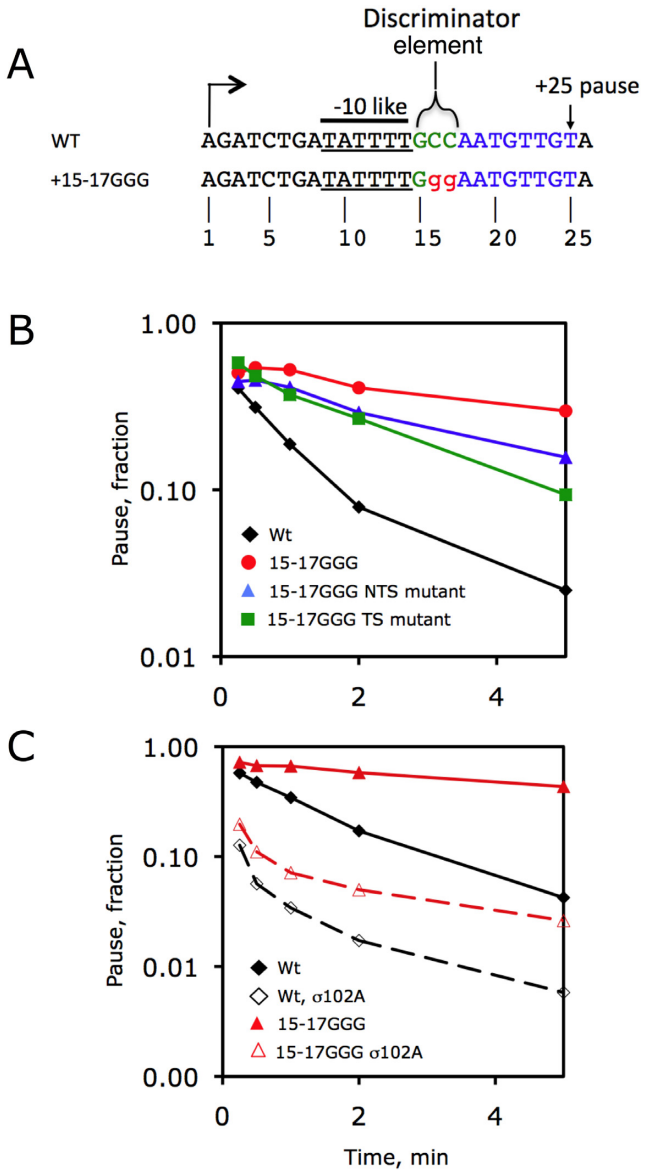

Figure 4. A consensus discriminator element increases pausing and reveals that the σ70 region 1.2/non-template DNA interaction is still necessary for pause formation. (**A**) Sequences of WT and consensus discriminator element mutant. The discriminator element is marked in green; the AT-rich sequence is in blue; and mutations are marked in red. (**B**) +25/26 paused RNAs and terminated RNA from single round *invitro* transcription time courses of WT, +15 ± 17 GGG mutant and the two heteroduplex mutant templates were quantified and paused RNA was plotted as percent total RNA over time. WT is black, homoduplex mutant is red, non-template strand mutant heteroduplex is blue triangles and template strand mutant heteroduplex is green squares. (**C**) Effect of σ102A mutation on pausing from WT and +15 ± 17 GGG mutant templates; symbols are as in (B), with dashed lines designating reactions with the σ102A mutation.

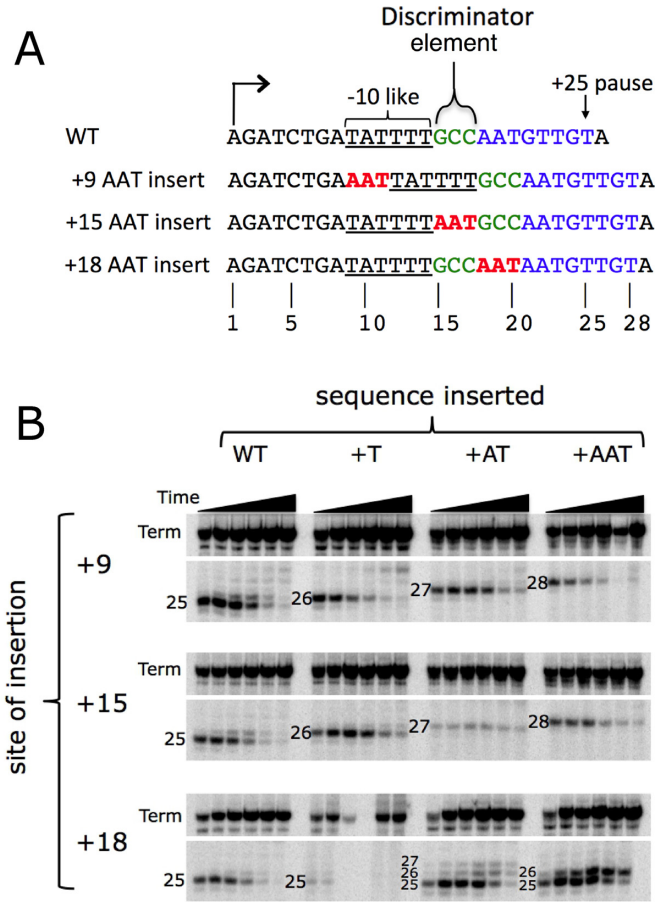

Figure 5. The GC-rich discriminator element determines the site of pausing. (**A**) Sequences of WT and insertion mutant templates. The discriminator element is marked in green; the AT-rich sequence is in blue; the −10-like element is underlined; and base insertions are marked in red. (**B**) Single round *invitro* transcription time-courses of T, AT and AAT sequence insertions at positions +9, +15 and +18 as shown in (A). Relevant pauses and terminated bands are labeled. Note that the single nucleotide insertion at +18 places a highly disfavored G at the pause end (see 'Discussion' section).

gion 1.2 non-template DNA binding, this would not have made sense; however, the essential activity of the non-consensus GCC sequence in binding σ70 region 1.2 to mediate pausing in *82pR'* transcription suggests that effective non-template strand interaction is not critically dependent upon base composition. Thus, changing the discriminator element may not seriously disrupt σ70 region 1.2 function, but would still allow determination of the effect of moving the template sequence of this region.

Two sets of insertions were controls for the effect of moving sequences surrounding the −10-like/discriminator element. In the first control set we made one, two and three nucleotide insertions at position +9 of the 82*pR'* template; these insertions move all of the sequence necessary for the pause downstream, and, unsurprisingly, gave rise to pauses at +26, +27 and +28, respectively (Figure 5B). In the second control set, nucleotides were inserted at +18, thus displacing and changing the AT-rich segment downstream of the discriminator element. These insertions do not move the site of the pause, but significantly modulate its intensity and the fine structure of paused ends; we discuss below the effect of the terminal pause nucleotide on pausing efficiency, which can account for the particularly weak pausing of the +18R insertion template. This result shows that the AT-rich sequence downstream of the discriminator element, which we knew previously affected the amount of pause formed, does not determine the site of the pause by itself (Perdue 2010).

Strikingly, inserting nucleotides at +15, thus separating the −10-like element and the discriminator element sequence, has the same effect as the +9 insertions, moving the pause downstream 1, 2 or 3 bases (Figure 5B). This result implies that the discriminator element either alone or in concert with the AT-rich sequence downstream (+18 to +25) determines the position of the pause. As we discuss below, these are components of the EPS. It is also clear that the position of the −10-like element, while necessary for inducing the pause, does not determine the specific location of the pause. If promoter-proximal pause formation is homologous to scrunching during transcription initiation, then insertions that separate the −10-like element and the discriminator element must have increased the amount of DNA that scrunches before the complex pauses, as modeled in Figure 1.

Longer insertions of 5 or 10 bases at +15 of the 82*pR'* template (insert sequences of AACAC or AACACAACAC, respectively) move the pause a corresponding number of nucleotides downstream (Supplementary Figure S2), although pausing is weak, around 10% of WT. The 5 bp insert template pauses at +26, +27, +28, +29 and +30 (Supplementary Figure S2); most of these species are partially sensitive to the RNA cleavage factor GreB, implying a tendency to backtrack, although the +30 pause is relatively unaffected, as the stable paused species should be. A GreB-resistant paused species is also seen at position +35 on the 10 base insert template, indicating that scrunching of the pause can be increased at least 10 bp.

Similar insertion mutations of the λ*pR'* template, which separate its −10-like sequence from the discriminator element by insertions after +6 of the transcript encoding region, appear to produce a displacement of the pause like that observed for 82*pR'* (Supplementary Figure S3). However, since the λ*pR'* pause is strongly dependent upon the non-template strand discriminator element, the pause becomes very weak (Supplementary Figure S3).

Pauses extended downstream by moving the discriminator element are still dependent on the σ70/−10-like site interaction

We demonstrate in two ways that pausing extended downstream by moving the discriminator element is still dependent on the original −10-like sequence in the +15 AAT construct. Conceivably, insertions at position 15 might have created a σ70-independent pausing sequence, or a new downstream −10-like sequence, especially since the inserted AAT resembles part of the −10 consensus. First, we performed *in vitro* transcription with σ70 L402F, a σ70 region 2 mutational alteration that sharply weakens pausing without preventing initiation at the promoter (14). This change nearly eliminated pausing on both WT 82*pR'* template and the +15AAT insertion (Figure 6B). Second, we disrupted the −10-like sequence of both the WT and +15AAT insert template DNAs (Figure 6A); this change also nearly eliminates the pause (Figure 6C). These results support our conclusion that inserting sequence between the −10-like element and discriminator element retains the original σ70 region 2 interaction, resulting in increased scrunching relative to the WT paused complex (Figure 7).

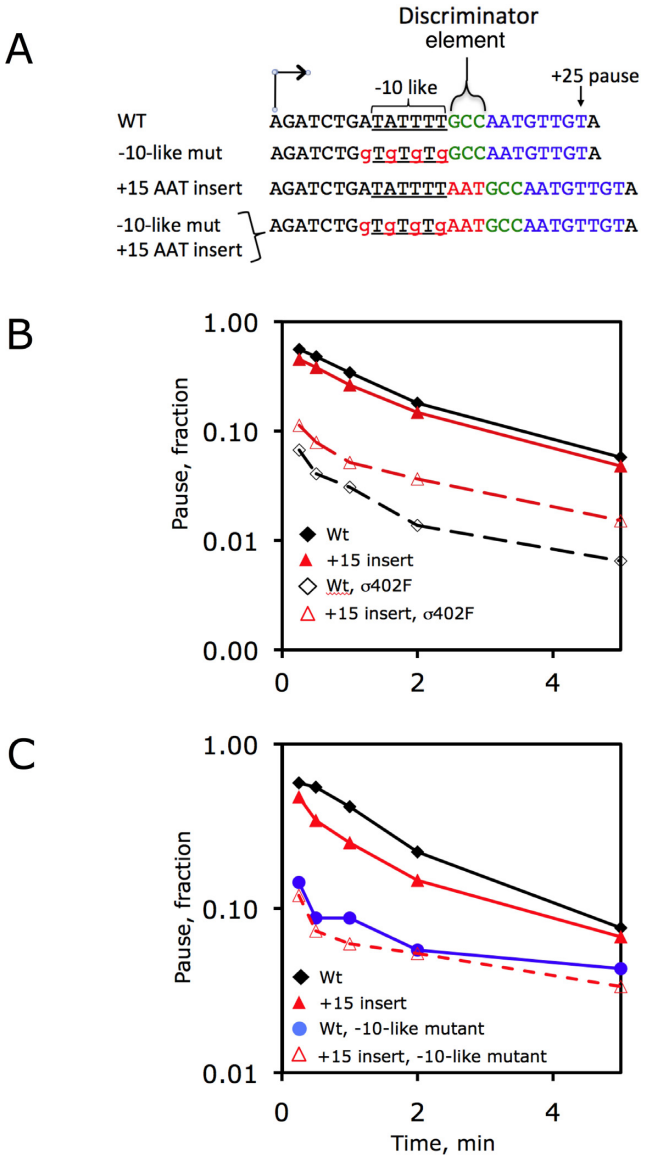

Figure 6. Pauses extended downstream are still dependent on the σ70/−10-like site interaction. (**A**) Sequences of WT and mutant templates. The discriminator element is marked in green; the AT-rich sequence is in blue; and mutations are marked in red. (**B**) Effect of the σ70 402F mutational alteration on pausing from WT and +15AAT insert DNA. Dashed lines designate pausing with σ70 402F. (**C**) Effect of −10-like sequence mutations on pausing from WT and +15AAT insert DNA. Dashed lines designate pausing with −10-like mutant templates.

The base at the 3′ terminal end of the RNA modulates σ70-dependent pause formation

Replacing the WT U at the terminal position (+25) of the paused RNA by A, G or C all reduced the +25/26 pause (Figure 8A). The initial amount of pause (pause capture) is about the same for U, C and A, and likely for G as well, but the pause lifetime clearly decreases in the order U > C > A > G for the terminal nucleotide of the pause. We show explicitly using heteroduplex templates that the terminal nucleotide effect is expressed from the template and not from the non-template DNA strand (Figure 8B); we

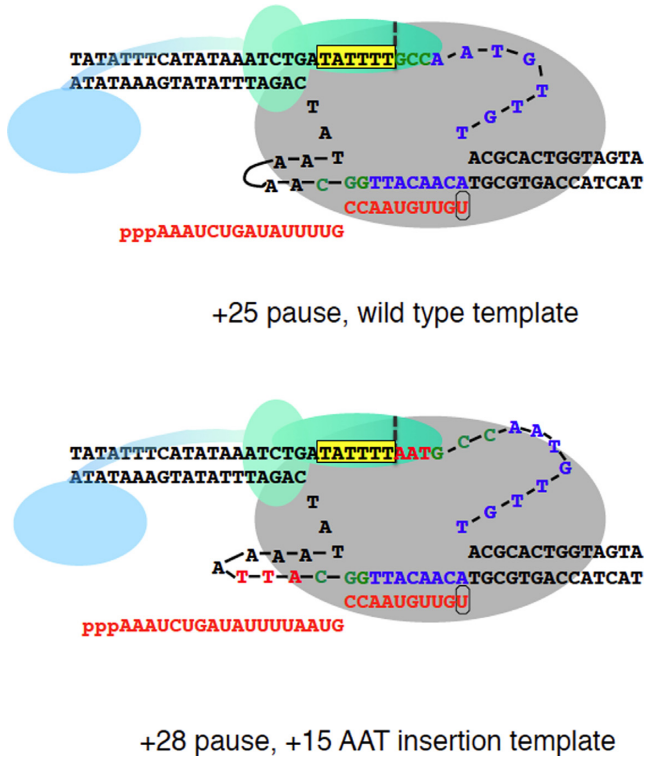

Figure 7. Models of the paused states of WT and insert template promoter proximal pauses. The GC-rich discriminator element is labeled in green, the AT-rich sequence is blue; −10-like sequence is yellow; RNAP is gray; and σ70 is light green. (**A**) Model of a transcription complex at the +25 promoter-proximal pause on WT phage 82 *pR'* DNA. (**B**) Model of a transcription complex at the +25 promoter-proximal pause on +15AAT insert DNA. The insert sequence moves the universal transcriptional pause sequence downstream, thus moving the site of the pause an equivalent distance. σ70 remains bound to the -10-like site, requiring the complex to create a larger scrunched bubble to reach the pause site.

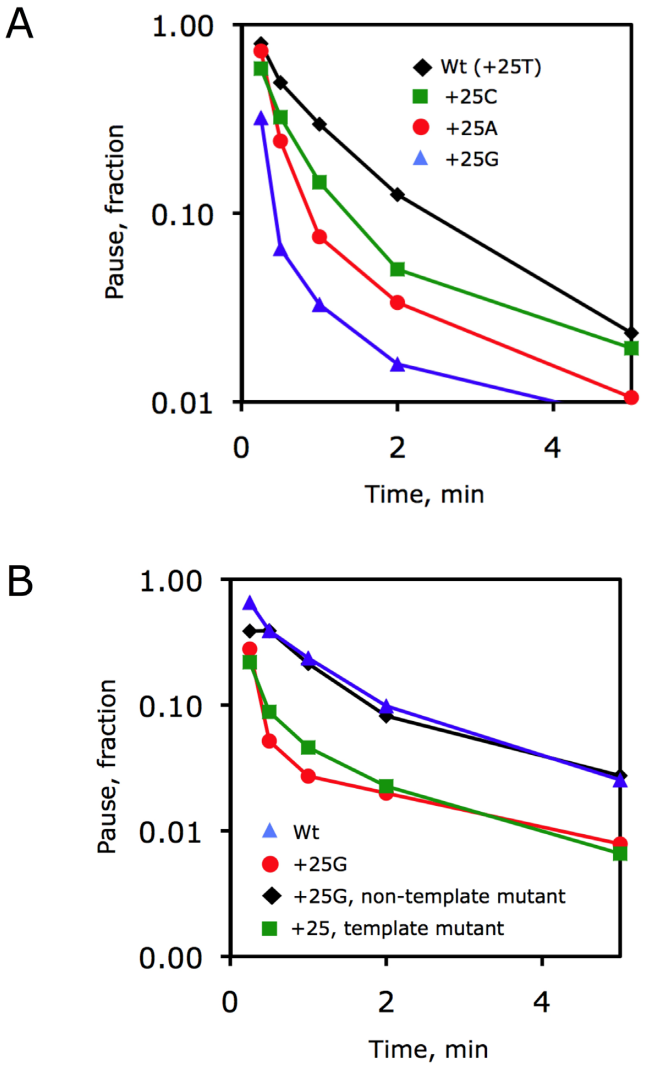

Figure 8. The nucleotide at the 3′ terminal end of the RNA modulates σ70-dependent pause formation. (**A**) Effects of single base substitution templates at position +25 of the 82 pR' template. (**B**) Heteroduplex template analysis of the +25G mutation.

also show that a similar nucleotide preference is found for the λ*pR'*-associated pause (Supplementary Figure S4). It is noteworthy that Hein *et al.* (34) found the same base preference (U > C > A > G) for the terminal nucleotide that favors persistence of the pre-translocated state of the catalytic center, as measured by susceptibility to pyrophosphorolysis. This coincidence strongly suggests that the phage 82*pR'* and λ*pR'*σ70-dependent pauses either are pre-translocated, or enter the paused state from the pre-translocated position; the result is consistent with the identification of the EPS as an essential element of σ70-dependent pausing.

DISCUSSION

We have identified and characterized essential elements of the promoter-proximal σ70-dependent transcription pause that underlies the modification of *E. coli* RNAP to an antiterminating state by lambdoid phage *Q* proteins. The pause mechanism involves universal features of the multisubunit RNAP activity, including the transition between translocation states and the ability of the active center to stabilize a scrunched DNA state. The capacity of σ70 to restrain RNAP by binding upstream elements of the enzyme while the catalytic center still polymerizes RNA is a model for processes that might induce transcription pausing by RNAPs of all three kingdoms (1).

The main question addressed here is the relation between the site of σ70 binding that initiates development of the pause and the actual, well-defined site where RNA polymerization stops at the pause. Because unwinding and scrunching of downstream DNA, accompanied by RNA synthesis on newly exposed template, continue after σ70 has bound, some other process must determine the RNA end. We show that a G/C rich segment in the template strand of DNA has this role. Notably, this G/C rich segment is at the upstream edge of the templating RNA/DNA hybrid in the stable paused complex, where RNA/DNA base pairs must be broken if RNA synthesis is to continue. This immediately suggests a role of this segment, consistent with previous proposals (18–20,35): the difficulty of unwinding the strong G/C (and perhaps C/G) base pairs at the upstream

end of the hybrid obstructs forward translocation that must accompany continued elongation.

Understanding of the function of the template strand G/C-rich segment in σ70-dependent pausing was obscured in previous work by its clear role as the non-template strand discriminator element that binds σ70 region 2 and stabilizes the open promoter complex, and, in fact, similarly stabilizes the σ70-dependent pause (29,31). This sequence in the non-template strand is particularly important as the site of the discriminator element, which when it is suboptimal destabilizes the open promoter complex through its inefficient interaction with σ70 region 1.2, a property that underlies regulation of promoters for synthesis of stable RNA species in *E. coli* (28). A previous study of the role of this sequence in σ70-dependent pausing from the λ*pR'* promoter noted primarily the non-template strand effect (33), which was dominant, although in fact the data showed substantial template strand effects at least for the first two positions downstream of the −10-like sequence. The difference is now clear: the λ*pR'* segment is much more dependent than that of 82*pR'* upon its match to a consensus for binding σ70 region 1.2, so that changing the original sequence mostly affected non-template strand function. We show here that increasing the match of the 82*pR'* sequence GCC to the consensus GGG reveals a clear non-template strand effect. This change also increases the template strand contribution to pausing (Figure 4B), consistent with the preference for G at −10 of the EPS sequence.

A critical element in efficient pausing at the site specified by the G/C-rich segment is the identity of the base at the RNA end: for the 82*pR'* pause, the natural U is most efficient, followed by C, and the purines support much weaker pausing. Note that insertion of a single nucleotide at +18 places the disfavored G at the pause site and greatly weakens the pause (Figure 5B). The λ*pR'* pause is affected similarly, with pyrimidines supporting strong pausing and purines only weak pausing. It is noteworthy that the order of pause efficiency for 82*pR'* (and nearly that of λ) is U > C > A > G; this is also the order determined by Hein *et al.* (34) for activity of the terminal base to inhibit movement of the active center from the pre-translocated to the post-translocated state—an essential step in the pathway to elongation.

We conclude that these elements combined (G/C-rich upstream segment and terminal nucleotide), which, possibly along with neighboring sequences, specify the position and intensity of the σ70-dependent pause, are identical to the sequence features that induce the 'ubiquitous' or 'elemental' pause, the EPS; this sequence was discovered by cataloguing pause sites during transcription in single-molecule experiments, and also through global analysis of stopped transcription complexes by NetSeq (18–20). It was persuasively shown that an EPS consensus, consisting primarily of G at −10 (the upstream boundary of the RNA/DNA hybrid), a pyrimidine base at the RNA end and a G at the next position, traps the complex in a pretranslocated state (19). The λ*pR'* sequence matches this exactly. The 82*pR'* sequence lacks the −10G, but it does have a preferred −11G (19) and the terminal pyrimidine; furthermore, we showed that introducing a −10G enhances pausing. The lack of G downstream of the RNA 3′ end should not have a large effect (19). We conclude that the λ*pR'* and 82*pR'* σ70-dependent pauses originate from or are in a pretranslocated state, and that the EPS sets the exact position of the σ70-dependent pause. A quantitative analysis of σ70-dependent pausing and further characterization of the EPS element has been published (36). We suggest that the configuration of the σ70-dependent pause, in which an upstream constraint (here σ70 binding) forces scrunching until a barrier like the EPS sequence stops elongation, might occur generally in factor-dependent pausing by all multisubunit RNAPs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Josh Filter, Man-Hee Suh and Sarah Perdue for their continual advice and help. We also would like to thank Amber Krauchunas for discussions about experimental design.

FUNDING

National Institutes of Health [GM21941 to J.W.R]. Funding for open access charge: Departmental research funds.
Conflict of interest statement. None declared.

REFERENCES

1. Kwak,H. and Lis,J.T. (2013) Control of transcriptional elongation. *Annu. Rev. Genet.*, **47**, 483–508.
2. Santangelo,T.J. and Artsimovitch,I. (2011) Termination and antitermination: RNA polymerase runs a stop sign. *Nat. Rev. Microbiol.*, **9**, 319–329.
3. 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.
4. 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.
5. 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.
6. 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.
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. Nickels,B.E., Mukhopadhyay,J., Garrity,S.J., Ebright,R.H. and Hochschild,A. (2004) The sigma 70 subunit of RNA polymerase mediates a promoter-proximal pause at the lac promoter. *Nat. Struct. Mol. Biol.*, **11**, 544–550.
9. Hatoum,A. and Roberts,J. (2008) Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. *Mol. Microbiol.*, **68**, 17–28.
10. Perdue,S.A. and Roberts,J.W. (2011) Sigma(70)-dependent transcription pausing in Escherichia coli. *J. Mol. Biol.*, **412**, 782–792.
11. Deighan,P., Pukhrambam,C., Nickels,B.E. and Hochschild,A. (2011) Initial transcribed region sequences influence the composition and functional properties of the bacterial elongation complex. *Genes Dev.*, **25**, 77–88.
12. Goldman,S.R., Nair,N.U., Wells,C.D., Nickels,B.E. and Hochschild,A. (2015) The primary sigma factor in *Escherichia coli* can access the transcription elongation complex from solution. *Elife*, **4**, e10514.

13. Mooney,R.A. and Landick,R. (2003) Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. *Genes Dev.*, **17**, 2839–2851.
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. Leibman,M. and Hochschild,A. (2007) A sigma-core interaction of the RNA polymerase holoenzyme that enhances promoter escape. *EMBO J.*, **26**, 1579–1590.
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. Strobel,E.J. and Roberts,J.W. (2014) Regulation of promoter-proximal transcription elongation: enhanced DNA scrunching drives lambda Q antiterminator-dependent escape from a sigma70-dependent pause. *Nucleic Acids Res.*, **42**, 5097–5108.
18. 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.
19. 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.
20. VVedenskaya,I.O., Vahedian-Movahed,H., Bird,J.G., Knoblauch,J.G., Goldman,S., Zhang,Y., Ebright,R.H. and Nickels,B.E. (2014) Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science*, **344**, 1285–1289.
21. 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.
22. Guo,H.C. and Roberts,J.W. (1990) Heterogeneous initiation due to slippage at the bacteriophage 82 late gene promoter in vitro. *Biochemistry*, **29**, 10702–10709.
23. Marr,M.T. and Roberts,J.W. (1997) Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, **276**, 1258–1260.
24. Borukhov,S. and Goldfarb,A. (1996) Purification and assay of Escherichia coli transcript cleavage factors GreA and GreB. *Methods Enzymol.*, **274**, 315–326.
25. Yang,X.J., Goliger,J.A. and Roberts,J.W. (1989) Specificity and mechanism of antitermination by Q proteins of bacteriophages lambda and 82. *J. Mol. Biol.*, **210**, 453–460.
26. Borukhov,S., Sagitov,V. and Goldfarb,A. (1993) Transcript cleavage factors from E. coli. *Cell*, **72**, 459–466.
27. 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.
28. Barker,M.M., Gaal,T., Josaitis,C.A. and Gourse,R.L. (2001) Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *J. Mol. Biol.*, **305**, 673–688.
29. Feklistov,A., Barinova,N., Sevostyanova,A., Heyduk,E., Bass,I., Vvedenskaya,I., Kuznedelov,K., Merkiene,E., Stavrovskaya,E., Klimasauskas,S. *et al.* (2006) A basal promoter element recognized by free RNA polymerase sigma subunit determines promoter recognition by RNA polymerase holoenzyme. *Mol. Cell*, **23**, 97–107.
30. 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.
31. Haugen,S.P., Ross,W., Manrique,M. and Gourse,R.L. (2008) Fine structure of the promoter-sigma region 1.2 interaction. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 3292–3297.
32. Travers,A.A. (1980) Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. *J. Bacteriol.*, **141**, 973–976.
33. Ring,B.Z. and Roberts,J.W. (1994) Function of a nontranscribed DNA strand site in transcription elongation. *Cell*, **78**, 317–324.
34. Hein,P.P., Palangat,M. and Landick,R. (2011). RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. *Biochemistry* **50**, 7002–7014
35. Gilbert,W. (1976) Starting and Stopping Sequences for the RNA Polymerase. In: Chamberlin,RLaMJ (ed). *RNA Polymerase*. Cold Spring Harbor Laboratory, NY, pp. 193–205.
36. Strobel,E.J. and Roberts,J.W. (2015) Two transcription pause elements underlie a sigma70-dependent pause cycle. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E4374–E4380.