Published in final edited form as:

Mol Microbiol. 2010 November ; 78(3): 636–650. doi:10.1111/j.1365-2958.2010.07347.x.

A backtrack-inducing sequence is an essential component of *E. coli* σ^{70} -dependent promoter-proximal pausing

Sarah A. Perdue and Jeffrey W. Roberts

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

SUMMARY

RNA polymerase of both bacteria and eukaryotes can stall or pause within tens of base pairs of its initiation site at the promoter, a state that may reflect important regulatory events in early transcription. In the bacterial model system, the σ^{70} initiation factor stabilizes such pauses by binding a downstream repeat of a promoter segment, especially the “–10” promoter element. We first show here that the “–35” promoter element also can stabilize promoter-proximal pausing, through interaction with σ^{70} region 4. We further show that an essential element of either type of pause is a sequence just upstream of the site of pausing that stabilizes RNA polymerase backtracking. Although the pause is not intrinsically backtracked, we suggest that the same sequence element is required both to stabilize the paused state and to potentiate backtracking.

Keywords

transcription; pausing; backtracking; antitermination

INTRODUCTION

The progress of RNA synthesis by multisubunit RNA polymerases can be interrupted a few tens of nucleotides after it starts, producing “paused” or “stalled” ternary transcription complexes that may be important regulatory intermediates (Hatoum & Roberts, 2008, Saunders *et al.*, 2006). Thus, at certain promoters *E. coli* RNA polymerase (RNAP) pauses after making ~15–25 nucleotides of RNA when the σ^{70} initiation subunit binds a reiteration of a component of the original promoter sequence (Ring *et al.*, 1996). Such paused ternary complexes are the substrate modified by the bacteriophage λ family gene *Q* transcription antiterminators, which then fundamentally alter the elongation properties of the enzyme as it extends the paused complex (Roberts *et al.*, 1998, Shankar *et al.*, 2007). Paused transcription complexes also exist in *E. coli* chromosomal transcription, although their function is unknown (Hatoum & Roberts, 2008, Brodolin *et al.*, 2004, Nickels *et al.*, 2004). In eukaryotes, as characterized especially in *Drosophila*, stalled complexes of 30–50 nucleotides are prevalent (Core *et al.*, 2008, Nechaev *et al.*, 2010), and reflect an important step at which transcription elongation is regulated. For all such complexes, the origin and structure of the paused or stalled state is of considerable interest.

The simple constitution of prokaryotic transcription complexes allows the molecular basis of promoter-proximal pausing to be understood in some detail. The well-characterized element associated with previously known examples of such pausing in *E. coli* has a significant match to the “–10” consensus sequence TATAAT and, as for initial promoter recognition, this sequence acts by engaging σ^{70} region 2 (Ring *et al.*, 1996). Although σ^{70} eventually is

released from the RNAP core during elongation, it has sufficient affinity to remain bound during early transcription and thereby to induce pausing (Mooney *et al.*, 2005, Kapanidis *et al.*, 2005). In transcription from the λ late gene promoter $\lambda pR'$, σ^{70} region 2 binds a nontemplate -10 -like element displaced 12 bp downstream from the promoter -10 element and induces a pause at +16 and +17, during which λQ protein acts (Grayhack *et al.*, 1985). This is a viable transcription complex, because it is a precursor to longer transcripts *in vitro*. The displacement of the -10 -like element implies that the pause should occur at +12, but, like the initial “abortive” complexes (Kapanidis *et al.*, 2006, Revyakin *et al.*, 2006), the paused complex appears to exist predominantly “scrunched”, meaning that transcription proceeds to +16 and +17 while σ^{70} region 2 remains bound to the -10 -like element, causing the corresponding 4 or 5 bp segment of melted template DNA to be drawn into the enzyme (Marr & Roberts, 2000, Kainz & Roberts, 1992). This same behavior is found for the lambdoid phage 82 pR' σ^{70} region 2-dependent +25 pause, which naturally should pause at +20, but scrunches to +25. (Similar scrunching and relaxation movements of RNAP are illustrated in the final model Figure.)

A notable property of the *E. coli* σ^{70} -dependent paused complexes is their tendency to undergo a form of backtracking (Marr & Roberts, 2000), the reaction believed to be universal to multisubunit RNA polymerases in which the enzyme spools backwards while the 3' end of newly synthesized RNA is extruded from the secondary channel of the enzyme (Komissarova & Kashlev, 1997, Nudler *et al.*, 1997, Marr & Roberts, 2000, Reeder & Hawley, 1996). Backtracking particularly occurs at sites of a weak (e.g. U-rich in RNA) terminal sequence, but also occurs at arbitrary sequences in response to a barrier that stops RNAP (Walter *et al.*, 2003). In an example of the backtracking-like reaction of σ^{70} -dependent paused complexes, the +17 λ pause component has relaxed the scrunch, causing the RNA end to be extruded and trapping the complex in an inactive state. (The evidence for this view is that a preformed +17 λ pause is sensitive to GreB-mediated cleavage *in vitro*, and that +17 λ pause does not accumulate significantly during active synthesis if GreB is present; however, it cannot be ruled out that during active synthesis GreB also inhibits backtracking of the +17 complex and thus promotes its elongation, rather than simply destroying it after it backtracks.) As in elongation backtracking, the extruded RNA is cleaved by the active center-mediated and Gre protein-stimulated hydrolysis reaction (Borukhov *et al.*, 2005) that trims the RNA back to about +12, placing the new 3' end in the active center ready for continued elongation (Marr & Roberts, 2000). This differs from backtracking in elongation in that the upstream edge of RNAP is held fixed by σ^{70} . On the other hand the +16 λ pause component does not relax the scrunch and is not Gre sensitive. This backtracking-like reaction of paused complexes occurs *in vitro* and *in vivo*, and the activity of Gre enhances the efficiency of Q function at the paused complex, presumably because the relaxed form is not elongated efficiently (Marr & Roberts, 2000). Perhaps similarly, the stalled Pol II complexes are sensitive to the action of SII, the eukaryotic counterpart of Gre (Adelman *et al.*, 2005).

We emphasize that the paused scrunched complex is not itself backtracked, and backtracking is not the underlying mechanism of the pause (although backtracking can itself induce pausing at other sites (Artsimovitch & Landick, 2000)); however, we argue here that there is nonetheless an intimate connection between the processes. We report two properties of the bacterial promoter-proximal paused complexes that underlie a new description of the process. Early transcription at 82 pR' reveals two pauses, the σ^{70} region 2-dependent one at +25 described above, and an additional pair at +14/15, called “the +14/15 pause” (Goliger & Roberts, 1989). In contrast to the +25 pause, we find that the +14/15 pause is associated with a reiteration of the promoter -35 sequence that is bound by σ^{70} region 4, the natural binding element for the -35 sequence in promoters. Furthermore, we find that the position and existence of the +14/15 pause is not uniquely specified by this -35 -like element, but

instead by surrounding elements that also specify efficient backtracking at the +14/15 site. A re-examination of the σ^{70} region 2-dependent pause at +25 reveals that it, too, requires the associated backtracking-inducing sequence. The revised view is that the bacteriophage and presumably bacterial chromosomal promoter-proximal pauses are induced or stabilized by sequences that also induce backtracking, and furthermore are stabilized by interaction of σ^{70} with appropriately placed binding sequences.

RESULTS

A residue substitution in the β flap reduces +14/15 pausing

Although prominent promoter-proximal pauses of both phage λ and phage 82 late gene transcription are well characterized to depend upon the binding of σ^{70} region 2 to a -10 like sequence, a second pause of the phage 82 promoter could not be associated with such a sequence. Thus, whereas the 82 promoter-associated pause at +25 pause is strongly dependent upon a -10 like sequence at +9 to +14 (see sequence in figures below), there is no sequence containing the essential A at +2 and T at +6 of the -10 promoter consensus hexamer that could induce the pause at +14/15. A separate accidental finding was that a mutation altering the surface of the RNA polymerase β subunit flap (E898R) strongly reduces pausing at +14/+15 but slightly increases pausing at +25. For both the +14 and +15 pause components, E898R reduced the percent of RNA captured in the pause (defined as pause efficiency extrapolated to zero time) and likely also the pause half-life at early times (Fig. 1).

The +14/15 pauses are affected by a -35 -like element repeat 9 nt downstream of the promoter -35

In RNAP holoenzyme, the β flap interacts with σ^{70} region 4 (Vassilyev *et al.*, 2002, Murakami *et al.*, 2002), and this interaction is required for region 4 to be positioned properly to recognize promoter -35 elements (Kuznedelov *et al.*, 2002). The reduction in +14/15 pausing by a mutational alteration of the flap surface suggested that the +14/15 pause might be induced by σ^{70} region 4 binding a -35 -like sequence. We identified a potential -35 -like sequence, TGGATT, displaced 9 nt downstream from the promoter -35 element in *82pR'* (Fig. 2A), with significant match to the -35 consensus TTGAC(T/A). To determine if this -35 -like element contributes to +14/15 pausing, we constructed a series of templates in which the -35 -like element was mutated to be either unlike (1M: CACTTT; 2M: AATCCA; 3M: ACCTAA) or more like (1C: TTGACT; 2C: TTGACA) the consensus -35 element. We performed *in vitro* transcription time course reactions on the six templates in the presence or absence of GreB protein (as discussed further below) and plotted the efficiency of pausing at +14/15 versus time (Fig. 2 and data not shown). For both the +14 and +15 pauses, mutating the -35 -like element away from consensus decreased pausing, to the same extent for 1M, 2M, and 3M, whereas mutating it to consensus increased pausing similarly for 1C and 2C; one example of each is shown in Figure 2B,C. In the absence of GreB, these effects are apparent in pause capture, or in pause half-life, or both. The template mutations had no effect on the efficiency of pausing at the σ^{70} region 2-dependent +25 pause (Fig. 2C).

As described above, the phage λ late gene promoter pause component at +17 is trapped in a backtracked-like state, and pause capture at +17 is strongly suppressed by the GreB cleavage factor *in vitro*. To determine what fraction of the phage 82 +14 and +15 pauses are trapped in this state, we performed synthesis in the presence of GreB (Fig. 2). The +15 pause is greatly reduced by GreB (Fig. 2C), percent capture decreasing of the order of tenfold. However, the +15 pause that remains still follows the trend of $C > WT > M$ for activity of the -35 -like element in inducing pause capture. The +14 pause is also reduced, although much less, in the presence of GreB, and the remaining pause behaves like the +15 pause in

response to -35 -like variants (Fig. 2C). We conclude that both pauses are subject to trapping in a backtracked state that can be eliminated or greatly reduced by GreB-mediated cleavage, revealing an underlying pause that depends on the strength of σ^{70} interaction with the -35 -like element. We have confirmed that GreB mediates cleavage of paused $+14/15$ complexes formed in its absence, consistent with its acting, as we have suggested earlier, after backtracking is established (data not shown). Overall, the σ^{70} region 4-dependent pause behaves similarly to the previously characterized σ^{70} region 2-dependent pause, in requiring a repeat of a σ^{70} promoter binding sequence and existing in a scrunched state.

The $+14/15$ pauses are stimulated by sequence-specific interactions between σ^{70} region 4 and the -35 -like element

To confirm directly that the changes in $+14/15$ pausing we observed in the template mutants are dependent on base-specific interactions between σ^{70} region 4 and the -35 -like element, we used a σ^{70} region 4 variant, R588H, which was selected to restore function to promoters having a third-position G $\rightarrow$ A mutation in the -35 element (Gardella *et al.*, 1989). We introduced a G $\rightarrow$ A mutation in the third position of the -35 -like element (template position -24) in the WT and consensus (1C) *82pR'* templates, a change that also reduced their match to the -35 consensus. We performed *in vitro* transcription time course reactions with σ^{70} WT or R588H holoenzyme (in the absence of GreB to magnify the signal) and calculated the percent capture for the $+14$ and $+15$ pauses on the WT, WT-24A, 1C and 1C-24A templates (Fig. 2D). For both $+14$ and $+15$ pauses on both WT and 1C templates, the $-24A$ mutation reduced pause capture as expected, given that these templates were mutated away from consensus. However, for both pauses on both WT-24A and 1C-24A templates, σ^{70} R588H suppressed this loss of pause capture, whereas it decreased pausing on WT and 1C templates. This result demonstrates a sequence-specific interaction between σ^{70} region 4 and the -35 -like element.

The strength of interaction between σ^{70} region 4 and the β flap affects capture of the $+14/15$ pauses

Previously, residue substitutions in σ^{70} region 4 that decrease (L607P) or increase (T544I/D581G) the relative strength of the interaction between region 4 and the β flap were identified (Nickels *et al.*, 2005). The activity of RNAP on -35 -dependent but not -35 -independent promoters was affected by these changes in interaction strength, consistent with the role of the flap in positioning region 4 to recognize promoter -35 elements. We asked if these σ^{70} region 4 substitutions correspondingly affect capture of the $+14/+15$ pauses. We performed *in vitro* transcription time course reactions on the 1C (consensus -35 -like) template with σ^{70} WT, L607P, or T544I/D581G holoenzyme and measured pausing: for both the $+14$ and $+15$ pauses, strengthening the flap/region 4 interaction relative to WT increased the amount of paused RNA, whereas weakening the flap/region 4 interaction decreased it (Supplementary Fig. 1A,B). This result mirrors the effects of these alterations on promoter function, and thus confirms the role of this specific protein/DNA interaction in pausing at $+14/+15$. Presumably β E898R similarly reduces pausing at $+14/15$ by weakening the σ^{70} – β flap interaction.

Other elements of RNAP holoenzyme that make contacts with promoter DNA are not involved in $+14/15$ pausing

σ^{70} region 1.2 acts in conjunction with σ^{70} region 2 in promoter recognition (Haugen *et al.*, 2006), extending the nontemplate DNA strand interaction downstream of the -10 consensus; correspondingly, σ^{70} region 1.2 is required along with σ^{70} region 2 for σ -dependent pausing at the $\lambda +16/17$, *lac* $+17/18$, and phage 82 $+25$ pauses (Zenkin *et al.*, 2007; J. Filter & JWR, unpublished). In addition, the RNAP α subunit C-terminal domain (α CTD) contributes to promoter recognition by binding DNA upstream of the -35 element,

particularly in sequence-specific binding to the upstream UP element (Ross *et al.*, 1993), and could make sequence-specific contacts with DNA at downstream pauses. To determine if any of these three elements affects the phage 82 +14/15 pause, we performed *in vitro* transcription time course reactions on WT, 1M and 1C 82pR' templates with variants of RNAP that affect region 2, region 1.2, and α CTD function: the σ^{70} region 2 variant L402F (Ko *et al.*, 1998); the σ^{70} region 1.2 variant M102A (Haugen *et al.*, 2008); and holoenzyme lacking the C-terminal domain of the α subunit (Δ CTD). We calculated the ratio of WT to mutant RNAP capture in the +14/15 pauses for each variant and found that they had little to no effect on the efficiency of pausing (Supplementary Fig. 1C). As expected, σ^{70} L402F and M102A severely impaired capture of the +25 pause. Thus, only region 4 of σ^{70} affects the +14/15 pause.

The –35-like element supports pausing *in vivo*

To determine if the σ^{70} region 4-dependent pause occurs *in vivo*, we used KMnO₄ footprinting to probe for paused complexes on plasmids in growing cells (Sasse-Dwight & Gralla, 1991; Kainz & Roberts, 1992; Marr & Roberts, 2000). KMnO₄ modifies single-stranded thymidine residues exposed in transcription bubbles, and these modified thymidines are detected by the inability of *Taq* polymerase to elongate a primer through them. To detect promoter-proximal pauses on transcription from 82pR', we treated MG1655 cultures harboring plasmids carrying the promoter and either WT or consensus (1C) –35-like pause-inducing elements with KMnO₄ for 1 minute periods during a time course after rifampicin addition (Fig. 3B). Because rifampicin inhibits initiation, the rifampicin time course identifies open complex as reactivity that increases over time as the promoter fills, whereas reactivity that decays over time is due to a downstream pause. We also treated cells harboring plasmids with A+10T variants of the promoters to identify the σ^{70} region 2- and –10-like element dependent +25 pause (data not shown); A+10T mutates the +25 pause-inducing –10-like element, so reactivity lost in the +10T variants was assigned to +25 pausing (Kainz & Roberts, 1992). The remaining reactivity was assigned to +14/15 pausing (Fig. 3A, B).

For the +14/15 signal, the presence of a consensus –35-like element enhances reactivity over the WT template (Fig. 3C). The reactivity at +14/15 was not detectable above background for a mutated –35-like element, template 1M (data not shown). There was no significant difference in reactivity of the +25 pause between the WT and 1C templates, confirming that mutating the –35-like element does not affect the +25 pause as seen *in vitro* (Fig. 3C). These results confirm the identification of reactivity due to the +14/15 pause and establish its existence *in vivo*.

A repeat of the promoter –35 element is not sufficient to induce a pause

The current model for σ^{70} -dependent pausing suggests that only a –10-like sequence is required to induce pausing; there is evidence, for example, that moving the –10-like sequence (and surrounding sequences, as discussed below) induces a new pause at a corresponding displacement (Ring *et al.*, 1996; Mooney & Landick, 2003). A relevant question then is whether a –35-like element in a different position could either enhance pausing at +14/15 or induce a pause at a correspondingly displaced position. Thus, we constructed a series of 82pR' template variants in which we placed a consensus –35 sequence either upstream by 2 nts (cons-2 template) or downstream by 4 or 7 nts (cons+4 and cons+7 templates) of the natural –35-like element, leaving no trace of consensus in the original position (Fig. 4A), and performed *in vitro* time course reactions in the presence or absence of GreB. We did not detect a new RNA length associated with paused complexes in either condition (Fig. 4B and data not shown); each template behaved similarly to the original mutant (1M), with +14/15 being the predominant, albeit weak, paused species (Fig.

4C; compare to 1M template in Fig. 2C). The displaced elements also did not enhance the +14/15 pause. Consistent with the previous data, mutating the sequence of 82pR' between the promoter -35 and -10 elements has no effect on the pause efficiency at +25 (Fig. 4C). Thus, the presence of a consensus -35-like element downstream of the promoter -35 element is not sufficient to induce a region 4-dependent pause; this result implies that a second sequence element is required for the -35-like element to support pausing, and that this second element also determines the position of the pause. We identify this element below, and we further show that, contrary to previous notions, this element also is required for σ^{70} region 2 – dependent pauses.

Increasing the GC-content of the 3' end of the hybrid of the σ region 4-dependent promoter-proximal pause strongly decreases pause efficiency

We described above the property of σ^{70} -dependent pauses to undergo a backtracking-like reaction, in which relaxation of the scrunch allows the paused RNA 3' end to emerge from the secondary channel. We now consider the possibility that this reaction is determined by sequences that also are the second element required for σ^{70} -dependent pausing, and that determine the position of the pause. We emphasize that the σ^{70} -dependent scrunched pause has not yet relaxed into the backtracked-like state, so that we are proposing a distinct role for these sequences in determining the pause end and in stabilizing the scrunched state, in addition to their potentiating the backtracking-like reaction. We conjecture that the backtracking-like reaction has the same sequence basis as normal elongation backtracking, which is favored by a weak terminal hybrid (Nudler et al., 1997) and by sequences upstream that stabilize the hybrid after backtracking (Lewis *et al.*, 2008). At +14/15, a 9 nt DNA:RNA hybrid would contain 7 dA/U or T/rA basepairs. An AU-rich segment in the hybrid, particularly at the 3' end of the paused RNA, is common to all previously identified region 2-dependent pauses; furthermore, all of these AU-rich hybrid RNA sequences are preceded by a GC-rich sequence (Fig. 5 and Supplemental Fig. 2). When no σ^{70} -binding element is present to anchor the scrunched complex, then the remaining pause would be a normal elongation backtrack to a more stable position, like other “class II” backtracked pauses (Artsimovitch & Landick, 2000).

To determine if these sequences induce the pause that remains in the absence of σ^{70} binding (i.e. behave like a class II pause), we constructed 82pR' template variants with mutant -35-like elements and a more GC-rich region at the 3' end of the +14/15 hybrid (Fig. 6A); this change should suppress backtracking because the new terminal +14/15 hybrid is more stable than WT and is no longer preceded by a more stable hybrid. We first showed that such a template artificially stalled at +15 does show less tendency to backtrack, as evidenced by decreased sensitivity to GreB-induced cleavage (Supplementary Fig. 4). We then performed *in vitro* transcription time course reactions on the -35-like 1M templates with either WT or GC-rich hybrids at +14/15 (Fig. 6B). Increasing hybrid stability of complexes at +14 and +15 further reduced to a nearly undetectable level the pause that remains after mutation of the -35-like sequence, indicating that the -35-like element and an AU-rich RNA are the two components of pausing at these positions (Fig. 6C). (Note that the +25 pause is affected for a different reason, as described below.)

Most importantly, we also saw a strong reduction in pause efficiency due to the GC-rich hybrid modification when a consensus -35-like element was present, i.e. the condition that supports the σ^{70} -dependent scrunched pause (Supplementary Fig. 3B,C). However, the pauses are not eliminated completely. If, as we propose, the backtrack-inducing sequence is essential for pausing, this can be explained plausibly as a slight remaining backtracking-like activity of the GC-substituted terminal sequences that supports some σ^{70} -dependent scrunched pause. Note that the GC-rich template transcript stopped at +15 is still sensitive to GreB, although much less so than the wild type template (Supplementary Fig. 4).

A region 2-dependent pause is similarly affected by hybrid stability

Like the +14/15 pause that depends on σ^{70} region 4, the well-characterized σ^{70} region 2-dependent pauses have an AU-rich terminal segment preceded by a GC-rich segment that could be essential for the scrunched pause (Fig. 5); as described, the +17 component of the λ σ^{70} region 2-dependent pause is strongly backtracked, whereas the scrunched +16 component is not. We asked if this backtrack-inducing sequence is essential to a σ^{70} region 2-dependent pause as well. First, in order to destroy the element completely, we changed 5 bp of the -10-like element that is required for the phage 82pR' +25 pause. (Because the sites coincide by chance, this is the same template used above to change the +14/15 hybrid, and the result is thus shown in Fig. 6C, +25 plot.) The region 2-induced +25 pause was greatly reduced by destroying the -10-like element, but in fact it was not completely eliminated. This result suggests that, like the +14/15 pauses, the +25 pause must be induced by more than one component.

To determine if a backtrack-inducing sequence also contributes to the +25 pause, we constructed templates with a more GC-rich sequence in the terminal portion of the +25 hybrid, a change predicted to disable the backtrack-inducing sequence at +25, (Fig. 6D; Supplementary Fig. 5A), and performed *in vitro* transcription time course reactions on variants with either a WT or mutated -10-like element. When the wild type -10-like element is present, the template with a GC-rich hybrid at +25 itself has greatly reduced +25 pause efficiency compared to the WT template (Supplementary Fig. 5B,C). This result indicates that the presence of a -10-like element is not sufficient to induce fully this promoter-proximal pause. As expected this template is defective in supporting 82Q-mediated antitermination (Fig. 7). When both pause-inducing elements are mutated, no detectable +25 pause remains; this result indicates again that both σ^{70} -binding and backtrack-inducing sequences are required for the pause (Fig. 6D,E). Furthermore, since there is a pause at +25 when the -10-like sequence is destroyed completely, the backtrack-inducing sequence must determine the position of this pause.

DISCUSSION

We have presented a novel view of the phenomenon of promoter-proximal pausing in the bacterial model system provided by bacteriophage λ and its relatives. The initial description showed pausing dependent on σ^{70} region 2, which recognizes a downstream repeat of the promoter -10 element, or a close variant of it (Ring et al., 1996). In the natural pathway to the pause, σ^{70} releases the promoter contacts but remains bound to core, through several separate protein-protein contacts, and is available to bind the downstream -10 element repeat (Marr et al., 2001). Likely this binding is weakened relative to that in free holoenzyme or in open complex, particularly when σ^{70} region 4 is displaced from its binding site on the RNAP β flap as the emerging transcript reaches a length of about 16 nucleotides (Nickels et al., 2006); thus, σ^{70} does dissociate eventually from core. Even after σ^{70} is released from RNAP core, a segment including a -10 consensus and surrounding sequences placed well downstream of the promoter induces a pause *in vitro* if sufficient σ^{70} is present in solution to maintain its association with core (Mooney & Landick, 2003). Here we show that a σ^{70} binding element is not sufficient to induce a pause, but instead that an associated backtrack-inducing sequence is required.

Our first finding that a pause depends upon σ^{70} region 4 instead of σ^{70} region 2 is novel but not surprising. The fact that the -35-like element we identified would be encountered when the transcript is 9 nucleotides long, and that σ^{70} region 4 binding to this element strengthens pauses at lengths of 14 and 15, is consistent with the persistence of the σ^{70} region 4 interaction with the RNAP β flap for RNAs of these lengths (Nickels et al., 2005). Similarly, it would not be expected that this interaction could anchor a pause with longer RNA.

Although there is no known regulatory role for the σ^{70} region 4-dependent pause at this promoter, since it appears not to be involved in 82Q function (Yang *et al.*, 1989), we showed that this pause exists in vivo and thus represents a potential alternate mode of promoter-proximal pausing in *E. coli*.

More significant is the realization that the σ^{70} -dependent pauses have an essential sequence that by itself can induce backtracking of elongation complexes, but which here functions to define the end of the scrunched, paused complex, and is in fact required for the establishment of the stable scrunched pause. Presumably there is a property of this sequence that both stabilizes the scrunch and potentiates backtracking; likely this is an initial disruption of the elongating complex at the site of backtracking that prevents the catalytic function of elongation (Toulokhonov *et al.*, 2007, Bar-Nahum *et al.*, 2005) and thus enables both stable scrunching and backtracking. The notion that these sequences are essential makes sense of the conserved AT-rich terminal sequence and the GC-rich element just downstream of the -10 -like element in the σ^{70} region 2-dependent pauses. It is noteworthy that Pol II promoter-proximal stalling also is associated with sequences that favor backtracking (Nechaev *et al.*, 2010).

An outstanding question about the mechanism of σ^{70} -dependent pauses is the order of events that stabilize them, and thus the precise role of σ^{70} -binding and backtrack-inducing sequences. We consider two models, not necessarily exclusive, for formation of the paused scrunched complex. The first model proposes that upon first encountering the σ^{70} binding sequence RNAP stops and further RNA synthesis is accompanied by scrunching that ends when RNA reaches the pause site (1 $\rightarrow$ 2 $\rightarrow$ 4 in Fig. 8). We emphasize that the site of the pause in RNA is different from that predicted by the position of the growing RNA chain when σ^{70} could first engage. By this measure, σ^{70} region 4 should engage the -35 -like element of the phage 82 *pR'* promoter when the RNA end is at +9, whereas the pause is at +14/15; and σ^{70} region 2 should engage the -10 -like element at +20, whereas the pause is at +25. By analogy with the structure of initial transcribing RNAP complexes, the paused complex should continue RNA synthesis from the initial engagement site to the pause as DNA is scrunched into the enzyme (2 $\rightarrow$ 4 in Fig. 8). It is striking that the pause, where scrunching stops, occurs just where backtracking also initiates in the absence of any σ^{70} involvement, i.e. on templates with a mutated σ^{70} binding sequence. This coincidence suggests strongly that the scrunch advance is stopped by the same instability of the growing complex that gives rise to backtracking, perhaps the “elemental pause” (Toulokhonov *et al.*, 2007), and implies an intimate connection between the backtracking tendency and the structure of the paused complex. The scrunched complex with σ^{70} then has a substantial lifetime, just as do initial transcribing complexes containing σ^{70} (Margeat *et al.*, 2006).

We suggest that the distinct behavior of the 82*pR'* +14 and +15 pause components reflects the differing stabilities of these scrunched states: +15 defaults more quickly to the backtracking-like condition (4 $\rightarrow$ 5 in Fig. 8), whereupon the pause is trapped until it is relieved by Gre-mediated or the slower endogenous cleavage (5 $\rightarrow$ 2 in Fig. 8). Thus, the +15 pause has greater overall sensitivity to Gre.

The involvement of a site where backtracking can occur in elongation suggests a second model for formation of the scrunched complex: the complex could continue advancing when it first encounters the σ^{70} binding element in the correct register (Fig. 8, #2), presumably because more time is required for this engagement than is available during translocation, and instead proceed to the site where backtracking starts (2 $\rightarrow$ 3 in Fig. 8). Backtracking (3 $\rightarrow$ 5 in Fig. 8) then would realign σ^{70} with its binding site, and, since the backtracked state could persist much longer than the average translocation time during RNA synthesis, σ^{70} binding could become established. Escape from the pause would likely occur through cleavage, by

either the Gre-mediated or the slower but significant endogenous pathway (Sigurdsson *et al.*, 2010). Gre facilitates pause escape and enhances Q function at the pause, but is not required for either; whether the endogenous cleavage activity might be essential is unknown. Conceivably the backtracked complex also could isomerize directly to the scrunched, paused configuration (5→4 in Fig. 8), although there might exist a difficult kinetic barrier in this direction.

An important unresolved question is the precise role of the GC-rich and AU-rich components of the backtrack-inducing sequence in σ^{70} – dependent pausing. In the second model we described, an actual elongation backtrack occurs (1→2→3→5→2→4 of Figure 8), which these sequences would induce in addition to their role in capturing the scrunch. However, in the first model (1→2→4 of Figure 8) there is no elongation backtrack to be stabilized; instead, the backtrack-inducing sequence acts only to stop elongation and capture the scrunch. It seems plausible that the AU-rich terminal portion is responsible for this, but how the GC-rich element might contribute is unclear. Mutating the GC-rich sequence strongly reduces the phage λ σ^{70} region 2 – dependent pauses, although this could reflect binding of σ^{70} region 1.2 in this segment. These questions await further investigation.

Our findings strengthen understanding of the role of sequences that induce backtracking. Backtracking appears critical to efficient function of the universal multisubunit family of RNA polymerases, in conjunction with the active center mediated hydrolysis of an internal RNA bond stimulated by either prokaryotic Gre, eukaryotic SII, or archaeal TFS proteins (Borukhov *et al.*, 2005) (Sigurdsson *et al.*, 2010). Backtracking may facilitate RNAP progress through an obstruction, as in the experimentally contrived collision of bacterial RNAP with a nucleosome (Walter *et al.*, 2003); Gre-mediated cleavage of the retreated complex could allow further collisions and favors eventual passage through the barrier. Backtracking by a base pair is thought to underlie removal of terminal mismatched ribonucleotides, which destabilize the hybrid, and thus underlies a proofreading activity (Sosunov *et al.*, 2003, Zenkin *et al.*, 2006, Thomas *et al.*, 1998, Erie *et al.*, 1993). A further notion is that a ubiquitous tendency of RNAP to backtrack and stall underlies regulation of the rate of transcription elongation: nudges from a trailing RNAP or presence of a translating ribosome (or potentially another element) that binds emerging RNA and inhibits its retraction into the enzyme during backtracking would promote forward progression of RNAP (Proshkin *et al.*, 2010). We describe here a variant function of backtrack-inducing sequences to act along with the σ^{70} initiation factor to stabilize a paused transcription complex that is the target of an antiterminator in *E. coli*.

EXPERIMENTAL PROCEDURES

Templates and plasmids

Templates for transcription were made from plasmids by PCR amplification and gel purified. Plasmid p82a+2G contains the wild-type phage 82 late gene promoter followed by the intrinsic terminator t82 at position +82, modified by a A→G substitution at +2. The +2G mutation interrupts the +1→+3 AAA initiating sequence that causes heterogeneous initiation on the wild-type promoter (Guo & Roberts, 1990). Mutations were introduced into p82a+2G via Quikchange site-directed mutagenesis (Stratagene).

Protein purification

The following proteins were purified as described: RNAP core (Hager *et al.*, 1990), 6xHis- σ^{70} wild-type and R588H (Marr & Roberts, 1997), 6xHis- β wild-type and E898R RNAP core (Santangelo *et al.*, 2003). 6xHis- σ^{70} variants L607P and T544I/D581G were a gift of A. Hochschild and B. Nickels (Harvard Medical School, Boston, MA). 6xHis- σ^{70} M102A was

a gift of J. Filter (Cornell University, Ithaca, NY). GreB protein was a gift of I. Artsimovitch (Ohio State University, Columbus, OH).

***In vitro* transcription reactions**

Reactions containing 2nM linear templates and 20nM RNAP (20nM core reconstituted with 100nM σ^{70}) were incubated in TB (20mM Tris-HCl pH 8.0, 0.1mM EDTA, 1mM DTT and 50mM KCl), 0.1 mg/ml BSA and 10% glycerol, plus 200 μ M ATP, GTP and CTP and 50 μ MUTP (supplemented with 0.5 μ Ci/ μ l α^{32} P-UTP) at 37°C for 5–10 minutes. For the reactions in Figure 1, 6xHis- β wild-type or E898R RNAP core was used with 6xHis- σ^{70} wild-type. For the reactions in Figures 2, 4 and 5, wild-type RNAP core was used with 6xHis- σ^{70} WT or mutants. When present, GreB was used at 100nM. Transcription was initiated by addition of MgCl₂ to 5mM and rifampicin to 10 μ g/ml; this scheme allows a single round of synthesis because initiation occurs before rifampicin can bind to inhibit further rounds. Aliquots were taken at indicated times after initiation and added to 5 volumes cold 1.2x STOP buffer (600mM Tris-HCl, pH 8.0, 12mM EDTA, 0.16 mg/ml tRNA) on ice. Samples were extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), 2.5 volumes 100% EtOH was added, and samples were precipitated overnight at –20°C. RNAs were resolved on a 15% polyacrylamide denaturing gel, detected by a Molecular Dynamics Typhoon PhosphorImager, and quantified by Imagequant software. RNAs were normalized to the number of radiolabeled residues in the transcript, and pause efficiency was calculated. Pause efficiency is defined as the percent of an RNA that remains of the total RNAs that were synthesized to and past a specific pause. For example, +15 pause efficiency = 100 * (15 nt normalized RNA intensity)/(15 nt normalized RNA intensity + 25 nt normalized RNA intensity + terminated normalized RNA intensity).

82Q *in vitro* antitermination assays were performed and analyzed as described (Santangelo et al., 2003). The reactions contained 100nM GreB and 150nM NusA (NusA has no effect on pause efficiency at +14, +15 or +25 (data not shown)).

***In vivo* KMnO₄ footprinting on plasmid DNA**

Footprinting was performed essentially as described (Marr & Roberts, 2000). MG1655 cells harboring plasmids encoding 82pR' templates were grown in M9 glucose minimal media + 0.5% casamino acids to mid-log phase and split into 5 × 10ml flasks. For the steady-state sample, 1ml of 110mM KMnO₄ was added and the culture was shaken for 1 min at 37°C, then the 11mls were transferred to a centrifuge tube containing 1ml of 12x STE (300mM Tris-Cl; 600mM NaCl; 6mM EDTA; 900mM fresh β -mercaptoethanol) on ice to stop the reaction. Rifampicin was added to 200 μ g/ml to the remaining flasks, incubation continued for the indicated times, and samples were treated with KMnO₄ as above. For the no-treatment sample, 1ml H₂O was added to 10mls of cells and the 11mls transferred to 1ml 12x STE on ice. The sequencing ladder was generated from wild type p82A +2G plasmid.

After centrifugation (8 min. at 7000 × g), pellets were resuspended by vortexing in 2 × 200 μ l STET (8% sucrose, 5% Triton X-100, 50 mM Tris-HCl (pH 8.0), and 50 mM EDTA). RNaseA and lysozyme were added to 10 μ g/ml each and samples were boiled 5 min. After centrifugation (5 min. at 16000 × g), supernatants were transferred to clean tubes and 5 volumes Qiagen buffer PB was added. The mixture was transferred to Qiagen mini-prep columns and plasmid purification continued as per the manufacturer's instructions.

To detect modified thymidine, equal amounts of modified template DNA (~1 μ g) were used in 10 rounds of linear PCR; labeled oligonucleotide primer was in at least 5-fold molar excess throughout the reaction. After electrophoretic resolution, reactive bases were detected by a Molecular Dynamics Typhoon PhosphorImager and quantified by Imagequant

software. Reactivity was normalized to a non-specific band to account for differences in plasmid concentration and efficiency of the extension reaction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Ann Hochschild, Bryce Nickels, and members of the laboratory for reading the manuscript and the NIH (Grant GM 21941) for support.

References

- Adelman K, Marr MT, Werner J, Saunders A, Ni Z, Andrulis ED, Lis JT. Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFHS. *Mol Cell*. 2005; 17:103–112. [PubMed: 15629721]
- Artsimovitch I, Landick R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc Natl Acad Sci U S A*. 2000; 97:7090–7095. [PubMed: 10860976]
- Bar-Nahum G, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A, Nudler E. A ratchet mechanism of transcription elongation and its control. *Cell*. 2005; 120:183–193. [PubMed: 15680325]
- Borukhov S, Lee J, Laptenko O. Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol Microbiol*. 2005; 55(1):315–1324.
- Brodolin K, Zenkin N, Mustaev A, Mamaeva D, Heumann H. The sigma 70 subunit of RNA polymerase induces lacUV5 promoter-proximal pausing of transcription. *Nat Struct Mol Biol*. 2004; 11:551–557. [PubMed: 15122346]
- Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*. 2008; 322:1845–1848. [PubMed: 19056941]
- Erie DA, Hajiseyedi O, Young MC, von Hippel PH. Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science*. 1993; 262:867–873. [PubMed: 8235608]
- Gardella T, Moyle H, Susskind MM. A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J Mol Biol*. 1989; 206:579–590. [PubMed: 2661827]
- Goliger JA, Roberts JW. Sequences required for antitermination by phage 82 Q protein. *J Mol Biol*. 1989; 210:461–471. [PubMed: 2559207]
- Grayhack EJ, Yang XJ, Lau LF, Roberts JW. Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell*. 1985; 42:259–269. [PubMed: 2990726]
- Guo HC, Roberts JW. Heterogeneous initiation due to slippage at the bacteriophage 82 late gene promoter in vitro. *Biochemistry*. 1990; 29:10702–10709. [PubMed: 2271677]
- Hager DA, Jin DJ, Burgess RR. Use of Mono Q high-resolution ion-exchange chromatography to obtain highly pure and active *Escherichia coli* RNA polymerase. *Biochemistry*. 1990; 29:7890–7894. [PubMed: 2261443]
- Hatoum A, Roberts J. Prevalence of RNA polymerase stalling at *Escherichia coli* promoters after open complex formation. *Mol Microbiol*. 2008; 68:17–28. [PubMed: 18333883]
- Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, Gourse RL. rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell*. 2006; 125:1069–1082. [PubMed: 16777598]
- Haugen SP, Ross W, Manrique M, Gourse RL. Fine structure of the promoter-sigma region 1.2 interaction. *Proc Natl Acad Sci U S A*. 2008; 105:3292–3297. [PubMed: 18287032]
- Kainz M, Roberts J. Structure of transcription elongation complexes in vivo. *Science*. 1992; 255:838–841. [PubMed: 1536008]
- Kapanidis AN, Margeat E, Ho SO, Kortkhonjia E, Weiss S, Ebright RH. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*. 2006; 314:1144–1147. [PubMed: 17110578]

- Kapanidis AN, Margeat E, Laurence TA, Doose S, Ho SO, Mukhopadhyay J, Kortkhonja E, Mekler V, Ebright RH, Weiss S. Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis. *Mol Cell*. 2005; 20:347–356. [PubMed: 16285917]
- Ko DC, Marr MT, Guo J, Roberts JW. A surface of *Escherichia coli* sigma 70 required for promoter function and antitermination by phage lambda Q protein. *Genes Dev*. 1998; 12:3276–3285. [PubMed: 9784501]
- Komissarova N, Kashlev M. Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc Natl Acad Sci U S A*. 1997; 94:1755–1760. [PubMed: 9050851]
- Kuznedelov K, Minakhin L, Niedziela-Majka A, Dove SL, Rogulja D, Nickels BE, Hochschild A, Heyduk T, Severinov K. A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science*. 2002; 295:855–857. [PubMed: 11823642]
- Lewis DE, Komissarova N, Le P, Kashlev M, Adhya S. DNA sequences in gal operon override transcription elongation blocks. *J Mol Biol*. 2008; 382:843–858. [PubMed: 18691599]
- Margeat E, Kapanidis AN, Tinnefeld P, Wang Y, Mukhopadhyay J, Ebright RH, Weiss S. Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophys J*. 2006; 90:1419–1431. [PubMed: 16299085]
- Marr MT, Datwyler SA, Meares CF, Roberts JW. Restructuring of an RNA polymerase holoenzyme elongation complex by lambdaoid phage Q proteins. *Proc Natl Acad Sci U S A*. 2001; 98:8972–8978. [PubMed: 11481468]
- Marr MT, Roberts JW. Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*. 1997; 276:1258–1260. [PubMed: 9157885]
- Marr MT, Roberts JW. Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol Cell*. 2000; 6:1275–1285. [PubMed: 11163202]
- Mooney RA, Darst SA, Landick R. Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell*. 2005; 20:335–345. [PubMed: 16285916]
- Mooney RA, Landick R. Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. *Genes Dev*. 2003; 17:2839–2851. [PubMed: 14630944]
- Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*. 2002; 296:1285–1290. [PubMed: 12016307]
- Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science*. 2010; 327:335–338. [PubMed: 20007866]
- Nickels BE, Garrity SJ, Mekler V, Minakhin L, Severinov K, Ebright RH, Hochschild A. 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*. 2005; 102:4488–4493. [PubMed: 15761057]
- Nickels BE, Mukhopadhyay J, Garrity SJ, Ebright RH, Hochschild A. The sigma 70 subunit of RNA polymerase mediates a promoter-proximal pause at the lac promoter. *Nat Struct Mol Biol*. 2004; 11:544–550. [PubMed: 15122345]
- Nickels BE, Roberts CW, Roberts JW, Hochschild A. RNA-mediated destabilization of the sigma(70) region 4/beta flap interaction facilitates engagement of RNA polymerase by the Q antiterminator. *Mol Cell*. 2006; 24:457–468. [PubMed: 17081994]
- Nudler E, Mustaev A, Lukhtanov E, Goldfarb A. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell*. 1997; 89:33–41. [PubMed: 9094712]
- Proshkin S, Rahmouni AR, Mironov A, Nudler E. Cooperation between translating ribosomes and RNA polymerase intran transcription elongation. *Science*. 2010; 328:504–508. [PubMed: 20413502]
- Reeder TC, Hawley DK. Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. *Cell*. 1996; 87:767–777. [PubMed: 8929544]
- Revyakin A, Liu C, Ebright RH, Strick TR. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*. 2006; 314:1139–1143. [PubMed: 17110577]

- Ring BZ, Yarnell WS, Roberts JW. Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*. 1996; 86:485–493. [PubMed: 8756730]
- Roberts JW, Yarnell W, Bartlett E, Guo J, Marr M, Ko DC, Sun H, Roberts CW. Antitermination by bacteriophage lambda Q protein. *Cold Spring Harb Symp Quant Biol*. 1998; 63:319–325. [PubMed: 10384296]
- Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K, Gourse RL. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*. 1993; 262:1407–1413. [PubMed: 8248780]
- Santangelo TJ, Mooney RA, Landick R, Roberts JW. RNA polymerase mutations that impair conversion to a termination-resistant complex by Q antiterminator proteins. *Genes Dev*. 2003; 17:1281–1292. [PubMed: 12756229]
- Sasse-Dwight S, Gralla JD. Footprinting protein-DNA complexes in vivo. *Methods Enzymol*. 1991; 208:146–168. [PubMed: 1779834]
- Saunders A, Core LJ, Lis JT. Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol*. 2006; 7:557–567. [PubMed: 16936696]
- Shankar S, Hatoum A, Roberts JW. A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. *Mol Cell*. 2007; 27:914–927. [PubMed: 17889665]
- Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol Cell*. 2010; 38:202–210. [PubMed: 20417599]
- Sosunov V, Sosunova E, Mustaev A, Bass I, Nikiforov V, Goldfarb A. Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J*. 2003; 22:2234–2244. [PubMed: 12727889]
- Thomas MJ, Platas AA, Hawley DK. Transcriptional fidelity and proofreading by RNA polymerase II. *Cell*. 1998; 93:627–637. [PubMed: 9604937]
- Touloukhonov I, Zhang J, Palangat M, Landick R. A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell*. 2007; 27:406–419. [PubMed: 17679091]
- Vassilyev DG, Sekine S, Laptenko O, Lee J, Vassilyeva MN, Borukhov S, Yokoyama S. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*. 2002; 417:712–719. [PubMed: 12000971]
- Walter W, Kireeva ML, Studitsky VM, Kashlev M. Bacterial polymerase and yeast polymerase II use similar mechanisms for transcription through nucleosomes. *J Biol Chem*. 2003; 278:36148–36156. [PubMed: 12851391]
- Yang XJ, Goliger JA, Roberts JW. Specificity and mechanism of antitermination by Q proteins of bacteriophages lambda and 82. *J Mol Biol*. 1989; 210:453–460. [PubMed: 2559206]
- Zenkin N, Yuzenkova Y, Severinov K. Transcript-assisted transcriptional proofreading. *Science*. 2006; 313:518–520. [PubMed: 16873663]

Figure 1. +14/15 pause efficiency in the phage 82 late gene promoter is reduced by a residue substitution in the RNAP β flap

(left) Single-round *in vitro* transcription time course reactions performed on 82a+2G template with either WT or E898R RNAP; RNAs were resolved on a denaturing acrylamide gel. Relevant RNAs are marked. (right) Band intensities of +14, +15 and +25 RNAs were normalized within a reaction and plotted as % pause efficiency over time. See Methods for definitions.

Figure 2. A -35-like repeat/ σ^{70} region 4 interaction underlies the +14/15 pause
(a) Sequences of WT and -35-like element mutant templates 1M (no match to consensus -35) and 1C (consensus -35). The +14/15 pause-stimulating -35-like element is boxed, and the promoter -35 and -10 and the +25 pause-stimulating -10-like element are indicated. **(b)** Single-round *in vitro* transcription time course reactions performed on the three templates with WT RNAP and, where noted, in the presence of 100nM GreB; RNAs were resolved on a denaturing acrylamide gel. **(c)** Plots of +14, +15 and +25 pause efficiency over time. (top left) +14 pause efficiency over time for all templates in the presence or absence of GreB; (top middle) Scaled version of the +14 pause graph showing curves for WT and 1M +14 pause efficiency over time; (bottom left) +15 pause efficiency over time for all templates in the presence or absence of GreB; (bottom middle) Scaled version of the +15 pause graph showing curves for all 3 templates in the presence of GreB only; (right) +25 pause efficiency is shown in the presence or absence of GreB; for simplicity, curves are shown only for the WT template, which is representative of +25 pausing on all 3 templates. **(d)** Percent capture of the +14 and +15 pauses on the WT and 1C or -24A template variants with either σ^{70} WT or R588H holoenzyme.

Figure 3. The σ^{70} region 4-stimulated +14/15 pause occurs *in vivo*

(a) Sequences of the *82pR'* WT and 1C late gene promoters, with reactive thymidine residues due to open complex, +14/15 or +25 indicated (see text for explanations of how reactivities were assigned). (b) Cultures harboring *82pR'* +2G plasmids containing either WT or 1C promoters were untreated (NT) or treated with 10mM (final concentration) KMnO_4 for 1 minute periods after the start of incubation with 200 $\mu\text{g}/\text{ml}$ (final concentration) rifampicin for the indicated time; the 0 min culture (steady-state sample) was not treated with rifampicin. Plasmids were harvested and used as templates in primer extension reactions. Reactivity attributed to the different species is marked and a sequence ladder is shown. (c) Reactivities assigned to the +14/15 or +25 pauses were normalized to a non-specific band and plotted over time after rifampicin addition. The averages of three experiments are shown; error bars are $\pm 1\text{SEM}$.

Figure 4. A repeat of the promoter –35 consensus does not determine the position of a promoter-proximal pause

(a) Sequence of template variants with a –35 consensus sequence placed in different locations in the promoter. (b) Single-round *in vitro* transcription time course reactions performed with WT holoenzyme in the absence of GreB; RNAs were resolved on a denaturing acrylamide gel. (c) Plots of +14, +15 and +25 pause efficiencies over time on the five templates.

Promoter (pause)		hybrid pause	AU content of RNA in the 9nt hybrid at indicated pause
<i>82pR'</i> (+25)	AGAUCUGAUUUUUGCCAAUGUUGUACG		6
<i>82pR'</i> (+14/15)	AGAUCUGAUUUUUGCCAA		7/7
<i>λpR'</i> (+16/17)	AACGAUGGUUAAUUCGCUC		6/6
<i>21pR'</i> (+18)	AUAAGCUGCGUUGUGACU		5
<i>80pR'</i> (+17)	AACGCUGGGAUACGUUCAG		5
<i>lacZ</i> (+17/18)	AAUUGUGAGCGGAUAAACAUUU		5

Figure 5. Promoters with σ -dependent pauses contain weak RNA:DNA hybrids at the position of the pause

(left) Promoters with σ -dependent pauses; pause position is indicated in parentheses. Note that all pause positions were determined experimentally except that of *80pR'* which is assigned by analogy to *λpR'*. (middle) RNAs transcribed from these promoters are aligned with respect to the position of the pause; the site(s) of the pause for each promoter is underlined in the sequence, and the 9 nt hybrid at the pause is flanked by red dashed lines. The sequences begin with the +1 nucleotide; rA and U residues in the RNA:DNA hybrid are in red; rG and rC residues just upstream of the paused hybrid are in blue. (right) The number of rA and U residues out of the 9 nts in the hybrid at the position of the pause is indicated.

Figure 6. An AU-rich RNA in the RNA:DNA hybrid at the position of σ^{70} -dependent pauses is required for efficient pausing

(a) Sequences of the 1M and 1M/GC-rich hybrid at +15 templates; sites of pausing are underlined in red. (b) Single-round *in vitro* transcription reactions were performed on the templates in (a) in the presence or absence of GreB and resolved on a denaturing acrylamide gel. (c) Plots of +14, +15 and +25 pause efficiency over time on the templates described in (a). Note that +25 pause efficiency was expected to decrease on the GC-rich hybrid at +15 template because the GC-rich substitution coincidentally mutates the +25 pause-stimulating -10-like sequence (see part D, below). (d) Sequences of the -10-like mutant and -10-like mutant/GC-rich hybrid at +25 templates. The +25 pause-stimulating -10-like element is mutated in both templates; the +25 pause site is underlined in red. The -10-like mutant template change is identical to the GC-rich hybrid at +15 change shown in (a), although the -35-like element is different. This change in the -35-like element has no effect on the +25 pause efficiency. (e) Single-round *in vitro* transcription time course reactions were performed on the templates described in (d) in the presence or absence of GreB. No plots of pause efficiency are shown because the +25 pause was not detectable above background on the -10-like mutant/GC-rich hybrid at +25 template.

Figure 7. Both a -10-like element and a backtrack-inducing sequence are required for 82Q antitermination activity

(top) *In vitro* 82Q antitermination assays were performed over a range of 82Q concentrations on the templates described in Supplementary Figure 5a. (bottom) Moles of terminated (T) and readthrough (RT) RNA were quantified and the percent readthrough was plotted as a function of 82Q concentration.

Figure 8. A model for σ -stimulated promoter-proximal pausing

A model is shown for the establishment of promoter-proximal pausing by *E. coli* RNA polymerase, as discussed in the text. The case of σ^{70} region 4 is shown, but the same model applies to σ^{70} region 2-dependent pausing.