3078–3091 *Nucleic Acids Research, 2012, Vol. 40, No. 7* *Published online 2 December 2011*
doi:10.1093/nar/gkr1158

Structural transitions in the transcription elongation complexes of bacterial RNA polymerase during σ-dependent pausing

Ekaterina Zhilina[1], Daria Esyunina[1,2], Konstantin Brodolin[3] and Andrey Kulbachinskiy[1,*]

[1]Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia, [2]Molecular Biology Department, Biological Faculty, Moscow State University, Moscow 119991, Russia and [3]Université Montpellier 1, Université Montpellier 2, CNRS UMR 5236, Centre d'études d'agents Pathogènes et Biotechnologies pour la Santé, Montpellier 34293, France

Received July 13, 2011; Revised November 8, 2011; Accepted November 10, 2011

ABSTRACT

A transcription initiation factor, the σ^{70} subunit of *Escherichia coli* RNA polymerase (RNAP) induces transcription pausing through the binding to a promoter-like pause-inducing sequence in the DNA template during transcription elongation. Here, we investigated the mechanism of σ-dependent pausing using reconstituted transcription elongation complexes which allowed highly efficient and precisely controlled pause formation. We demonstrated that, following engagement of the σ subunit to the pause site, RNAP continues RNA synthesis leading to formation of stressed elongation complexes, in which the nascent RNA remains resistant to Gre-induced cleavage while the transcription bubble and RNAP footprint on the DNA template extend in downstream direction, likely accompanied by DNA scrunching. The stressed complexes can then either break σ-mediated contacts and continue elongation or isomerize to a backtracked conformation. Suppressing of the RNAP backtracking decreases pausing and increases productive elongation. On the contrary, core RNAP mutations that impair RNAP interactions with the downstream part of the DNA template stimulate pausing, presumably by destabilizing the stressed complexes. We propose that interplay between DNA scrunching and RNAP backtracking may have an essential role in transcription pausing and its regulation in various systems.

INTRODUCTION

Transcription elongation by multisubunit RNAPs is not a monotonous process and is punctuated by pauses that differ in nature and functional output (1). Multiple factors have been described that affect the rate and efficiency of RNA synthesis by both bacterial and eukaryotic RNAPs through modulating transcription pausing and termination (2–6). In particular, bacterial transcription initiation factor, the σ^{70} subunit of *Escherichia coli* RNAP, was shown to induce transcription pausing through binding to specific promoter-like pause-inducing sequences in the transcribed DNA template (7,8).

The phenomenon of σ^{70}-dependent pausing was first described during analysis of transcription regulation at late promoters of lambdoid phages, where σ^{70}-dependent promoter-proximal pauses, observed at positions +16/+17, +25, and +18 on the $\lambda P_R'$, $82P_R'$ and $21P_R'$ promoters, respectively, were shown to be required for recruitment of a phage-encoded Q protein that acts as an antitermination factor at later steps of transcription elongation (7–10). Later, a promoter-proximal σ^{70}-dependent pause was detected during transcription from cellular *E. coli lac*-promoter (11,12). Subsequent *in vivo* experiments revealed σ^{70}-dependent pauses on a significant number (up to 10–20%) of randomly chosen *E. coli* promoters (13), and a similar fraction of predicted *E. coli* promoters were shown to contain pause-inducing signals in their initially transcribed sequences (12,14). This suggested that σ^{70}-dependent pausing is widespread in transcription of *E. coli* genes and may play an important regulatory role in gene expression. Intriguingly, promoter-proximal pausing was also shown to be a ubiquitous phenomenon in transcription of eukaryotic genes by RNAPII, where it was found to be essential for efficient

*To whom correspondence should be addressed. Tel: +7 499 196 00 15; Fax: +7 499 196 00 15; Email: akulb@img.ras.ru

gene expression; however, its molecular mechanisms are largely unknown (15,16).

The observation that σ^{70} can induce pausing during transcription elongation suggested that it does not obligatory dissociate during the initiation-to-elongation transition. Indeed, FRET and chromatin-immunoprecipitation experiments demonstrated that a significant fraction of transcription elongation complexes (TECs) may contain the σ^{70} subunit both *in vitro* and *in vivo*, and supported a stochastic release model of σ^{70} dissociation from the TEC as a result of weakening σ-core interactions following initiation (17–21). Besides, *in vitro* experiments demonstrated that σ^{70} can rebind σ-free TEC and induce promoter-distal pausing when present at sufficiently high concentrations (11,22–24). Recently, it has been shown that the presence of a promoter-proximal pause signal increases σ^{70} occupancy of TECs over distances of at least 700 nucleotides downstream thus suggesting that promoter-distal σ^{70}-dependent pausing may also occur *in vivo* (14).

Available data reveal several characteristic features of σ^{70}-dependent pausing which are summarized below (8).

(i) The pausing depends on interactions of σ^{70} region 2 with a −10-like element (TATAAT) in the non-template DNA strand (7,8,11). Accordingly, σ^{70} was proposed to bind the TEC through interactions between σ region 2 and a coiled-coil element of the β′ clamp domain (10,20,24,25), with contribution from σ region 1.2 (23). In *lac*UV5 and $82P_R'$ templates, the −10-like element is preceded by a TG-like element but the role of the latter in pausing has not been explored. Recently, region 4 of σ^{70} was shown to induce an additional pause at positions +14/+15 of the $82P_R'$ promoter through interactions with a −35-like element (10). However, such pausing can likely occur only during early steps of elongation, since RNA longer than 15 nucleotides would displace σ4 (10,26), and its functional importance is unknown.

(ii) RNA cleavage factors, GreA and GreB in *E. coli* suppress RNAP pausing, and RNA transcripts in several analyzed paused TECs were shown to be highly sensitive to the Gre-induced cleavage, indicating that the paused complexes adopt a backtracked conformation (9–12). Furthermore, a specialized 'backtrack-inducing sequence' located just upstream of the RNA 3′-end at the site of pausing in late phage promoters was shown to stimulate RNAP backtracking and pausing even in the absence of the −10-like sequence (10). However, it remains unknown whether backtracking is a prerequisite or a consequence of σ-dependent pausing and how the backtrack-inducing sequence can regulate pausing in various systems.

(iii) A stressed intermediate complex is likely to form during σ^{70}-dependent pausing, as a result of RNA elongation without breaking the upstream σ^{70}-mediated RNAP-DNA contacts. In this complex, the transcribed DNA was proposed to be 'scrunched' within RNAP, without changes in the geometry of the RNAP molecule (8,9). Indeed, several paused complexes displayed reduced sensitivity to Gre-induced RNA cleavage in comparison with backtracked TECs suggesting that they adopt a forward translocated scrunched conformation (9,10). The scrunching can likely be followed by TEC backtracking, resulting in stress relief (8). Importantly, the scrunched TECs were proposed to serve as a preferable target for regulation by Q-proteins of lambdoid phages (9). Previously, DNA scrunching was directly observed during promoter-dependent transcription initiation by *E. coli* RNAP (27,28). However, the scrunched σ^{70}-containing pausing TECs have not been characterized from either mechanistic or structural points of view.

In this work, we describe a promoter-independent system for analysis of the pausing process, employing TECs assembled on synthetic nucleic acid scaffolds, which allows for highly efficient pause formation. By using a combination of *in vitro* transcription and footprinting assays, we characterize stepwise TEC transitions that occur during σ^{70}-dependent pausing by *E. coli* RNAP. Our results suggest that the balance between DNA scrunching, TEC backtracking and read-through elongation at the pause site can be modulated by multiple factors, including NTP concentrations, RNAP mutations, regulatory proteins and non-coding nucleic acids, establishing the basis for elaborate transcription regulation.

MATERIALS AND METHODS

Protein purification

Wild-type and the R339A mutant *E. coli* core RNAPs containing hexahistidine tags at the C-terminus of the β′ subunit were purified as described previously (29). The Ω216 *E. coli* RNAP was obtained as described in (30). The ΔSI3 and ΔJaw core RNAPs were a generous gift from I. Artsimovitch and N. Akulenko. Wild-type σ^{70} subunit was expressed from pMRG8 plasmid and purified as described (31). The GreB protein was purified as described in (32).

In vitro TEC assembly and analysis of σ-dependent pausing

TEC reconstitution was performed by a modified variant of a previously published method (33). The sequences of all oligonucleotides used for TEC reconstitution are shown in Supplementary Figure S12. RNA oligonucleotide was labeled at the 5′-end with γ-^{32}P-ATP and T4 polynucleotide kinase (New England BioLabs). Template DNA oligonucleotide (2.5 μM) was incubated with labeled RNA (250 nM) in buffer containing 40 mM Tris–HCl pH 7.9 and 40 mM KCl for 3 min at 65°C and cooled to 20°C at 1.5°C/min. The samples were diluted 3-fold with the same buffer and core RNAP was added to 250 nM. The core preparation contained <0.5% of the σ^{70}-holoenzyme as estimated by comparison of RNAP

activities in an abortive promoter-dependent transcription assay performed either in the absence or in the presence of externally added σ^{70}. The samples were incubated for 20 min at 40°C, non-template DNA oligonucleotide was added to 1 μM, and incubation was continued for another 20 min. For transcription experiments, 1–2 μl of the resulting TECs was used per 10 μl reaction point. The TECs were transferred to 25°C and diluted to the desired reaction volume with the same buffer. The σ^{70} subunit was added to 1 μM and the samples were incubated for 5 min at 37°C. In most experiments, the TECs were immobilized on Ni-NTA-agarose (Qiagen); in the case of the ΔJaw and Ω216 RNAPs, the TECs were bound to streptavidin affinity beads (Sigma) via a biotin residue introduced at the 5′-end of the non-template DNA strand. Control experiments performed with wild-type *E. coli* RNAP demonstrated that the pausing was identical in the case of both types of sorbents. For TEC immobilization, 20 μl of buffer-equilibrated affinity resin was added per 100 μl of the reaction mixture at 25°C, bound TECs were washed with 1 ml of buffer containing 500 mM KCl and three times with 1 ml of buffer containing 40 mM KCl.

Analysis of pausing was performed at 37°C in transcription buffer containing 40 mM Tris–HCl pH 7.9, 40 mM KCl and 10 mM $MgCl_2$. NTPs were added to desired concentrations and reactions were stopped after different time intervals by addition of stop-buffer containing 8M urea and 50 mM EDTA. In Figure 6, antisense oligonucleotides (3 μM) were added to starting 20-mer TEC 5 min prior to addition of NTPs, and transcription reaction was performed at 25°C. Stalled 21, 23 and 24-mer TECs were obtained by addition of limited substrate sets (CTP, CTP/GTP and CTP/GTP/ATP, respectively) to 20-mer TEC at 100 μM for 30 s at 37°C, followed by washing of the resin-absorbed TECs with 3 × 1 ml of the transcription buffer. The 25-mer TECs were obtained either by run-off RNA synthesis (in reactions containing σ^{70}) or by walking of the 23-mer TEC (with 100 μM ATP/UTP for 30 s at 37°C) followed by washing of the beads with the buffer. Analysis of RNA cleavage in stalled TECs was performed at 37°C in the presence of 1 μM GreB. RNA products were separated in 15 or 20% denaturing PAGE and analyzed by phosphorimaging.

ExoIII, DNaseI and $KMnO_4$ footprinting

In the ExoIII footprinting experiments, the TECs were obtained as described above, except either template or non-template DNA oligonucleotides were 5′-end labeled with γ-^{32}P-ATP, and unlabeled RNA was used for TEC assembly. About 3- to 5-fold higher amounts of reconstituted TECs were used for each reaction point (50–100 nM TEC concentration) to increase the signal intensity. Ten units of ExoIII (New England BioLabs) were added per 10 μl reaction point at 30°C in the transcription buffer. For footprinting of 23, 24 and 25-mer TECs, ExoIII was added 30 s after addition of corresponding NTPs sets. The reaction was stopped by addition of 10 μl of stop solution containing EDTA (50 mM) and calf thymus DNA (1 μg/μl). RNAP–DNA complexes were eluted with 300 mM imidazole (pH 7.9) for 5 min at 65°C, treated with chlorophorm, DNA was ethanol-precipitated, washed with 70% ethanol and dissolved in formamide-containing loading buffer. The samples were analyzed by 13% denaturing PAGE. The cleavage positions were deduced from comparison of the lengths of the digestion products with A+G cleavage markers obtained by treatment of corresponding DNA fragments with formic acid and piperidine (taking into account that this treatment removes the modified residue from DNA). The DNaseI footprinting experiments were performed in a similar way but the reactions were supplemented with $CaCl_2$ (2 mM). DNaseI (Fermentas) of 0.03 U was added per 10 μl reaction point and the reactions were stopped after 1 min by addition of 10 μl of the stop solution. In the $KMnO_4$ probing experiments, the TECs were obtained in the same way; 2 μl of 20 mM $KMnO_4$ was added per 20 μl reaction point (50 nM TEC) at 37°C, the reaction was stopped after 20 s by addition of β-mercaptoethanol (to 140 mM) and the samples were processed as described above.

RESULTS

σ^{70}-dependent pausing in TECs assembled on synthetic nucleic acid scaffolds

To analyze transcription pausing, we designed a nucleic acid scaffold based on the initially transcribed sequence of the *lac*UV5 promoter (Figure 1A). Previous work identified a σ^{70}-dependent pause at positions +17/+18 of *lac*UV5 (11,12). The pause-inducing sequence on the *lac*UV5 template contains a −10-like element placed between promoter positions +1 and +6, and a TG motif at positions −2/−3 (Figure 1A). To enhance σ^{70}-DNA interactions and facilitate analysis of pausing we replaced the wild-type pause-inducing sequence with the consensus extended −10-like element (TGcTATAAT), resulting in the 'Cons' scaffold template. We also substituted several nucleotides in the pause-site sequence to make it suitable for stepwise RNAP walking during *in vitro* transcription (Figure 1A). The 3′-end of the 20 nt RNA oligonucleotide used for TEC assembly was located 8 nt downstream of the −10-like element; this position corresponded to position +2 of the promoters so that the complex would adopt a relaxed conformation similar to that of the open promoter complex with the two initiating nucleotides bound at the RNAP active center.

We assembled TECs on the scaffold template using a σ^{70}-free preparation of the *E. coli* core RNAP and analyzed kinetics of RNA synthesis and transcription pausing in the absence or in the presence of externally added σ^{70} subunit. The experiments were performed with TECs immobilized on an affinity resin which allowed simple and convenient analysis of RNA synthesis and RNA cleavage (see 'Materials and Methods' section for details). In the presence of NTP substrates, RNAP elongated the starting 20 nt RNA to the end of the template resulting in the synthesis of a 52 nt long 'run-off' transcript (Figure 1B, the 'Cons/-σ' reaction). When the σ^{70} subunit was added, the pausing was

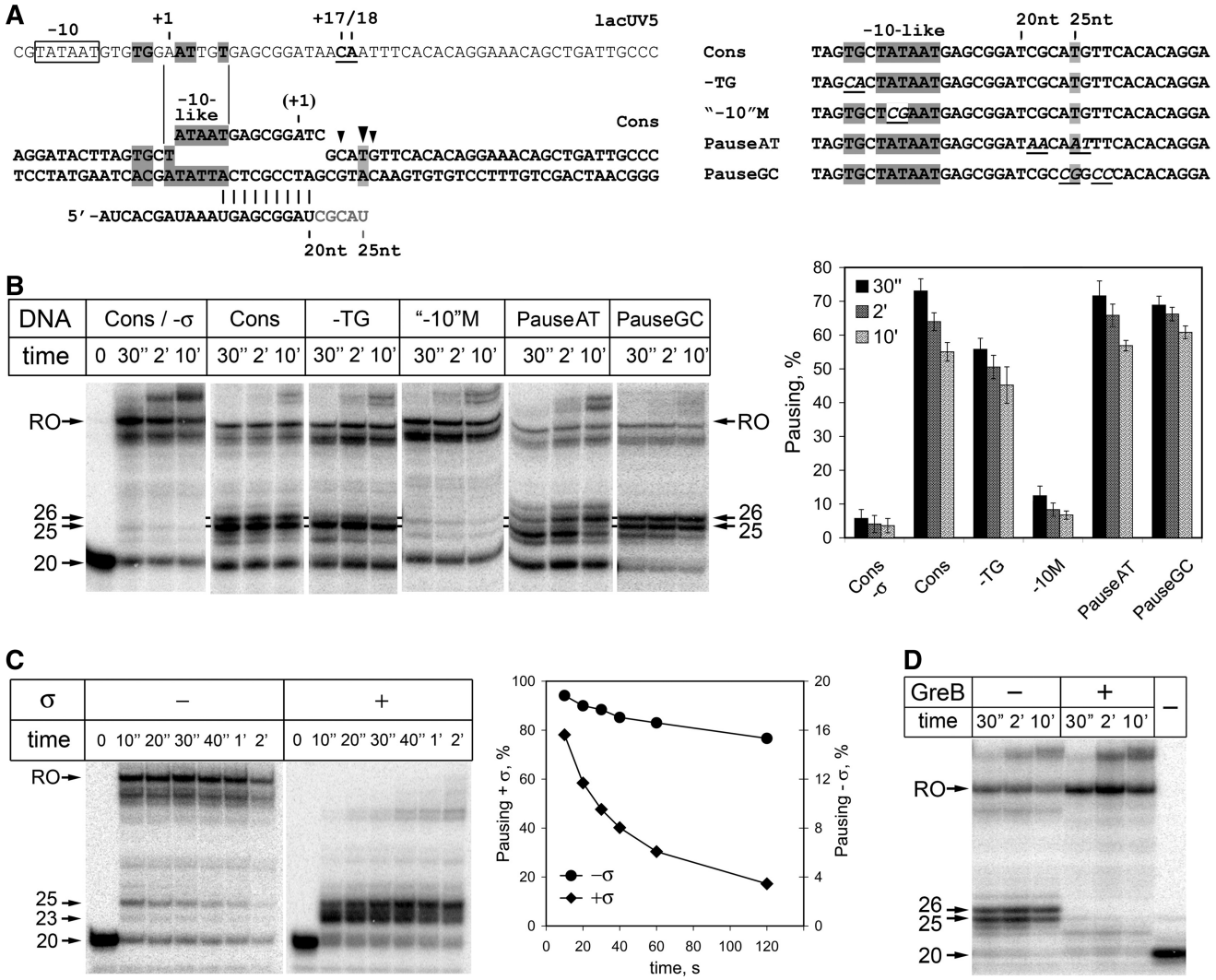

Figure 1. σ^{70}-dependent pausing in reconstituted TECs. (**A**) Synthetic nucleic acid scaffolds. Left panel, the structure of the 'Cons' scaffold used for analysis of pausing is shown in comparison with the initially transcribed region of the *lac*UV5 promoter. The −10 promoter element of *lac*UV5 is boxed, the transcription start point (+1) and the pause positions (+17/18) are shown above the sequence. The −10 and TG-like elements of the pause-inducing sequence are shadowed. The pause positions observed on the scaffold template are shown with arrowheads, the major position is shadowed. RNA nucleotides that are added to the starting 20-mer RNA during transcription at the pause site are gray. Right panel, sequences of different scaffold variants; nucleotides substituted in comparison with the 'Cons' scaffold are underlined italics. (**B**) Kinetics of σ^{70}-dependent pausing on different scaffold variants; transcription was performed at 100 μM NTPs. Positions of the starting 20-mer RNA, paused 25, 26-mer RNAs and the run-off (RO) transcript are indicated. The plot shows averages and standard deviations from three to four independent measurements. (**C**) Kinetics of RNAP pausing on the 'Cons' scaffold at 3 μM NTP concentration. (**D**) σ^{70}-dependent pausing on the 'Cons' scaffold at 1 mM NTPs analyzed either in the absence or in the presence of the GreB protein.

observed in TECs containing 25-mer and, to a smaller extent, 26-mer RNA transcripts, i.e. after addition of 5 or 6 nt to the starting RNA (Figure 1B, the 'Cons' reaction). The pause efficiency reached 70% and the pause half-life time exceeded 10 min (Figure 1B).

To analyze individual contributions of the pause-inducing and pause-site sequences to σ^{70}-dependent RNAP pausing, we generated several scaffold variants that lacked the TG element (the '−TG' scaffold), contained substitutions in the −10 element (the '−10M' scaffold), and contained an A/T-rich sequence (taken from *lac*UV5, the 'PauseAT' scaffold) or G/C-rich sequence ('PauseGC') at the pause site (Figure 1A, right). Substitutions in the consensus −10 element strongly decreased pausing, while substitutions of the TG motif resulted in ~10–15% decrease in the pausing efficiency (Figure 1B and Supplementary Figure S1). Neither of these substitutions affected the pattern of pausing. In contrast, substitutions in the pause-site sequence did not significantly affect the pausing efficiency but changed the pattern of pausing (Figure 1B). In the case of the 'PauseAT' template, the major pausing position was shifted upstream and corresponded to 24-mer complexes; these complexes were elongated to 26-mers at longer reaction times. In the case of the 'PauseGC' template, increased pausing was observed in 26-mer TEC indicating that this substitution facilitated downstream RNA elongation at the pause site in

comparison with the 'Cons' template (Figure 1B). Similar pattern was observed on a related template with another GC-rich sequence at the pause site ('PauseGC2', Supplementary Figure S2).

Conditions for σ^{70}-dependent and σ^{70}-independent pausing

In the experiment shown on Figure 1B, a fraction of complexes (~30%) escaped pausing even in the case of templates containing the consensus pause-inducing sequence. The read-through fraction might correspond to: (i) TECs that did not bind the σ^{70} subunit and did not form the σ^{70} contacts with the −10-like element, or (ii) TECs that broke the −10-element–σ^{70} interactions during transcription through the pause site. The experiments described above were performed in the presence of 100 μM NTPs and 1 μM σ^{70}. Analysis of pausing at different σ^{70} concentrations revealed that the apparent K_d for the σ^{70}-TEC interaction was ~200 nM but the pausing efficiency did not reach 100% even at saturating σ^{70} concentrations (≥1 μM, Figure 1 and Supplementary Figure S3). However, when the experiments were repeated at lower NTP concentrations (3 μM) almost 100% of TECs paused (Figure 1C, right). Thus, all TECs in our system can bind the σ^{70} subunit, and the lack of pausing in a fraction of complexes at high NTP concentrations is likely explained by disruption of σ^{70}-DNA contacts in these complexes during transcription through the pause site.

Further analysis of σ^{70}-dependent pausing at NTP concentrations ranging from 10 μM to 1 mM revealed that the pausing efficiency gradually decreases with increasing NTP concentrations (Supplementary Figure S1). Importantly, changes in the NTP concentrations also affected the pausing positions that were predominantly observed in 23/25-mer complexes at low NTP concentrations and in 25/26-mer complexes at high NTP concentrations (Figure 1C and Supplementary Figure S1). This suggests that changes in cellular NTP concentrations may regulate the σ-dependent pausing *in vivo* (see 'Discussion' section).

Noticeably, a small but detectable fraction of TECs assembled on the 'Cons' template paused even in the absence of σ^{70} (Figure 1B, the 'Cons/-σ' reaction). This fraction was significantly increased (to >10%) in the transcription reactions performed at low NTP concentrations (Figure 1C, left). The σ-independent pausing occurred in 25-mer TECs, at the position that coincided with the major position of σ-dependent pausing. However, the σ-less pausing complexes escaped the pause site with a much faster kinetics than the corresponding σ^{70}-containing complexes (plot in Figure 1C).

Previously, it was shown that an A/T-rich backtrack-inducing sequence located at the pause site in late promoters of lambdoid phages can stimulate RNAP pausing even when the −10-like pause-inducing element is mutated (see Introduction) (10). We therefore proposed that the σ-independent pausing observed on the 'Cons' template can also be explained by transient RNAP backtracking, stimulated by the presence of A/T pairs at the pause site. To test this proposal, we analyzed RNAP pausing on two templates with G/C-only pause-site sequences ('PauseGC' and 'PauseGC2'). The σ-independent pausing was abolished on both templates (Supplementary Figure S2). Furthermore, combination of mutations in the −10-like element with the G/C-rich pause-site sequence in the same template (the '−10 M_PauseGC' template) completely eliminated pausing both in the absence and in the presence of σ^{70} (Supplementary Figure S2), suggesting that the low-efficiency pausing observed on the '−10'M template (Figure 1B) was also dependent on the presence of the A/T base pairs at the 3′-end of the pause site.

RNAP backtracking and GreB-induced RNA cleavage during σ^{70}-dependent pausing

Previously, it was shown that the RNA cleavage factors GreA and GreB suppress σ^{70}-dependent promoter-proximal pausing of *E. coli* RNAP both *in vitro* and *in vivo* (see Introduction). In agreement with published data, GreB dramatically reduced σ^{70}-dependent pausing and stimulated run-off transcription on the 'Cons' template (Figure 1D), suggesting that TEC backtracking plays an essential role in the pausing. A catalytically dead GreB variant with a D41A substitution was unable to alleviate pausing indicating that the cleavage-stimulating activity of GreB was required for its antipausing action (AK, unpublished observation).

To understand the role of the backtracking during transcription pausing in more detail, we performed analysis of Gre-induced RNA cleavage in 20, 21, 23 and 24-mer TECs, stalled at different positions near the pause site, and in 25-mer paused TEC. No significant RNA cleavage occurred during the course of the experiment in the absence of Gre-factors (Supplementary Figure S4). In the presence of GreB, RNA in the starting 20-mer complex was also resistant to cleavage. However, the efficiency of cleavage increased with increasing the RNA length and reached ~100% in the case of the paused 25-mer complex, in which the reaction was essentially complete within 20 s (Figure 2). In all cases, the major reaction product was 20-mer RNA, resulting from cleavage 8 nt downstream of the pause-inducing sequence. Thus, the distance between the −10-like element and the RNAP active center in the paused relaxed TEC after backtracking corresponds to that in the open promoter complex.

Remarkably, a fraction of 21, 23 and 24-mer complexes (up to 40%) was much less sensitive to GreB action and significant amounts of RNAs in these complexes remained uncleaved even after 10-min incubation with GreB (Figure 2). These complexes may therefore correspond to stressed scrunched TECs that form during σ^{70}-dependent pausing. To check whether the cleavage-resistant TECs can be converted to cleavage-sensitive complexes, we obtained the σ^{70}-containing 23-mer TEC as described above, washed NTPs out and walked the complex two nucleotides downstream. The resulting 25-mer TEC was highly susceptible to GreB-induced RNA cleavage and was not readily elongated to full-length RNA upon addition of all four NTP substrates, indicating that it adopted a fully backtracked paused conformation

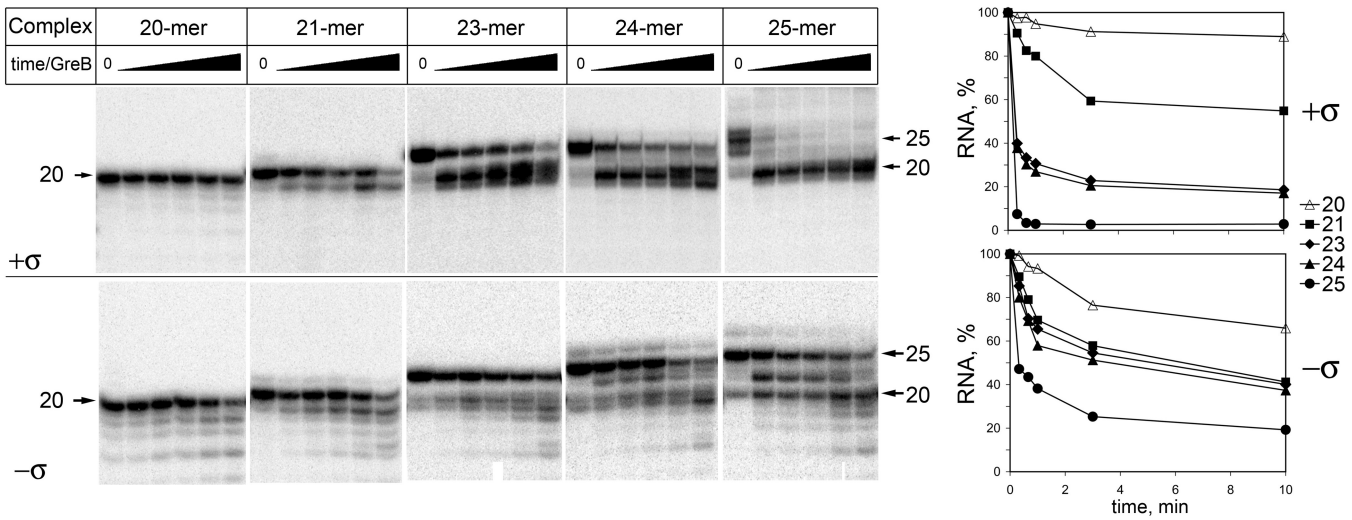

Figure 2. RNA cleavage in TECs stalled at different positions at the pause site in the presence (top) or in absence (bottom) of the σ^{70} subunit. The 25-mer TECs were obtained either by run-off RNA synthesis in the presence of σ^{70}, or by walking of the 23-mer complex in the absence σ^{70}.

(Supplementary Figure S5). Thus, the σ^{70}-containing TECs backtrack after reaching a critical RNA length (25 or 26 nt in our system) that corresponds to the preferable pausing position.

RNA cleavage in TECs obtained in the absence of σ^{70} was also increased when RNAP approached the pause site (Figure 2). However, the cleavage efficiency in σ-less 23, 24 and 25-mer TECs was significantly lower than in corresponding σ-containing TECs indicating that the presence of σ^{70} stimulated backtracking. At the same time, the extent of backtracking (i.e. the number of base pairs that TEC backtracks) was higher in σ-less complexes since RNA cleavage products shorter than 20 nt were detected in reactions with TECs lacking σ^{70} but not with σ^{70}-containing complexes (Figure 2).

To test whether RNAP backtracking at the pause site is stimulated by the presence of A/T base pairs at the pause site, we analyzed GreB-induced cleavage in stalled σ-less 24-mer TECs obtained on the 'PauseGC2' template. In contrast to 24-mer TECs obtained on the 'Cons' template, such complexes were essentially resistant to cleavage indicating that the G/C-rich pause-site sequence stabilizes the TECs in the active forward-translocated conformation (Supplementary Figure S6A). At the same time, the paused 24– 26-mer TECs formed on this template in the presence of σ^{70} were highly susceptible to cleavage (Supplementary Figure S6B) demonstrating that σ can stimulate TEC backtracking even in the absence of a backtrack-inducing sequence.

Previously, it was hypothesized that TECs may bind σ^{70} and pause not only during initial transcription through the pause-inducing sequence but also after backtracking that can occur several nucleotides downstream of the pause site and bring the TEC back to the pause-inducing sequence to allow σ^{70} binding (10). Differential susceptibility of stalled σ-less TECs obtained on the 'Cons' and 'PauseGC2' templates to backtracking allowed us to directly test this proposal. When σ^{70} was added to 21, 23, 24 and 25-mer TECs obtained on the 'Cons' template, followed by addition of nucleotide substrates, it induced pausing at the 25/26-mer positions (Supplementary Figure S7A and C). The efficiency of pausing reached ~70% in the case of 21 and 23-mer TECs and ~50% in the case of 24 and 25-mer TECs. In contrast, much less efficient pausing was observed when the experiment was repeated on the 'PauseGC2' template that prevents TEC backtracking at the pause site (Supplementary Figure S7B and C). Thus, the σ^{70} subunit can indeed bind backtracked TECs but cannot interact with forward translocated TECs that have transcribed through the pause site.

Analysis of DNA melting in TECs during σ^{70}-dependent pausing by $KMnO_4$ probing

Our further experiments were aimed at structural characterization of the TECs in the process of pausing. RNA elongation at the pause site should be accompanied by downstream extension of the transcription bubble and loading of an extra segment of downstream DNA into the RNAP molecule, without changes in the position of the upstream TEC border. To detect formation of such stressed TECs, we analyzed changes in the length of the melted DNA region and the RNAP footprint on DNA in TECs stalled at defined positions near the pause site.

To detect changes in the structure of the transcription bubble during pausing, we performed footprinting with $KMnO_4$, a reagent that specifically modifies single-stranded thymines. The pattern of thymine modifications in the starting 20-mer σ-less TEC corresponded to the expected size of the transcription bubble in the elongation complex (~13–14 bp). The leftmost modified thymine in the melted region was located at position −11 on the template strand and the rightmost thymine was located at non-template position +2 (with position +1 placed 7 bp downstream of the −10-like element) (Figure 3A,

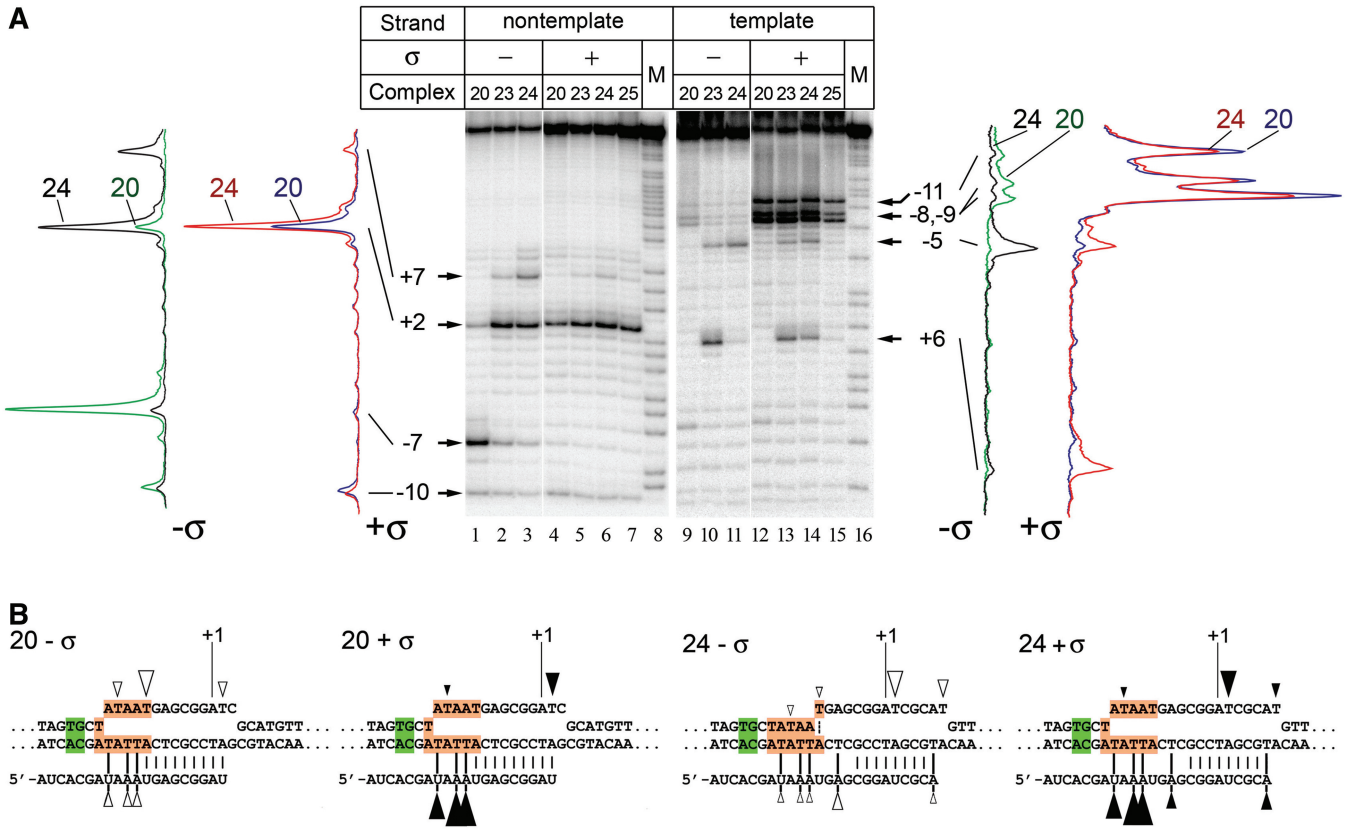

Figure 3. Analysis of DNA melting in TECs during σ^{70}-dependent pausing by $KMnO_4$ probing. (**A**) Analysis of the cleavage products by gel-electrophoresis. The scanned cleavage profiles in the 20-mer and 24-mer TECs obtained either in the absence or in the presence of the σ^{70} subunit are shown on the left (for the non-template strand) and on the right (for the template strand) of the gel. For each strand, the relative signal amplitudes are on the same scale. (**B**) Schematic representation of the footprinting results in 20-mer and 24-mer TECs. The central part of the 'Cons' scaffold template is shown. Positions of modified thymines in σ-less and σ^{70}-containing TECs are shown with open and closed arrowheads, respectively; the sizes of the arrowheads correspond to the observed modification efficiencies.

lanes 1 and 9, and Figure 3B). Similar pattern was observed in σ^{70}-containing 20-mer TEC, with some changes in thymine reactivities. In particular, σ^{70} protected non-template thymine −7 from modification, indicative of close σ contacts with the −10-like element (lane 4), and increased template thymine reactivities at positions −11, −9 and −8 in the same region (lane 12).

RNA extension in 23 and 24-mer σ-less TECs was accompanied by the melting of downstream DNA, as revealed by modification of thymines at positions +6 and +7 in the template and non-template strands, respectively (Figure 3A, lanes 2, 3 and 10, 11). In the 24-mer TEC, the template +6T became partially protected due to its pairing with the RNA 3′-end. Concurrently, the reactivity of upstream thymine residues decreased, indicative of gradual TEC translocation. At the same time, a low level of modification was still observed with upstream thymines indicating that a small fraction of complexes adopted a backtracked conformation.

Analysis of σ^{70}-containing 23 and 24-mer TECs also revealed additional downstream DNA melting in comparison with the 20-mer TEC, at positions +6 and +7 of the template and non-template strands (lanes 13, 14). Remarkably, in contrast to σ-less complexes, RNA extension in σ^{70}-containing TECs was not accompanied by a decrease in the reactivities of upstream thymine residues (compare lanes 9–11 and 12–14 and cleavage profiles for the template DNA strand). Thus, the observed downstream extension of the transcription bubble can be attributed to formation of scrunched TECs during transcription at the pause site. It should be noted that the actual efficiency of downstream DNA melting in scrunched TECs is likely higher, since the reactions also contained a fraction of backtracked TECs that did not have the extended transcription bubble.

ExoIII footprinting reveals TECs with extended DNA footprint during σ^{70}-dependent pausing

To detect changes in RNAP-DNA contacts during formation of stressed TECs, we performed ExoIII footprinting experiments. The ExoIII exonuclease progressively degrades one of the two DNA strands in the 3′→5′ direction which allows to determine downstream and upstream borders in TECs containing 5′-labels in the non-template and template DNA strands, respectively. To precisely locate the TEC borders we used a longer version of the 'Cons' scaffold that was extended by 10 nt in its upstream part (Figure 4C). Control experiments demonstrated that

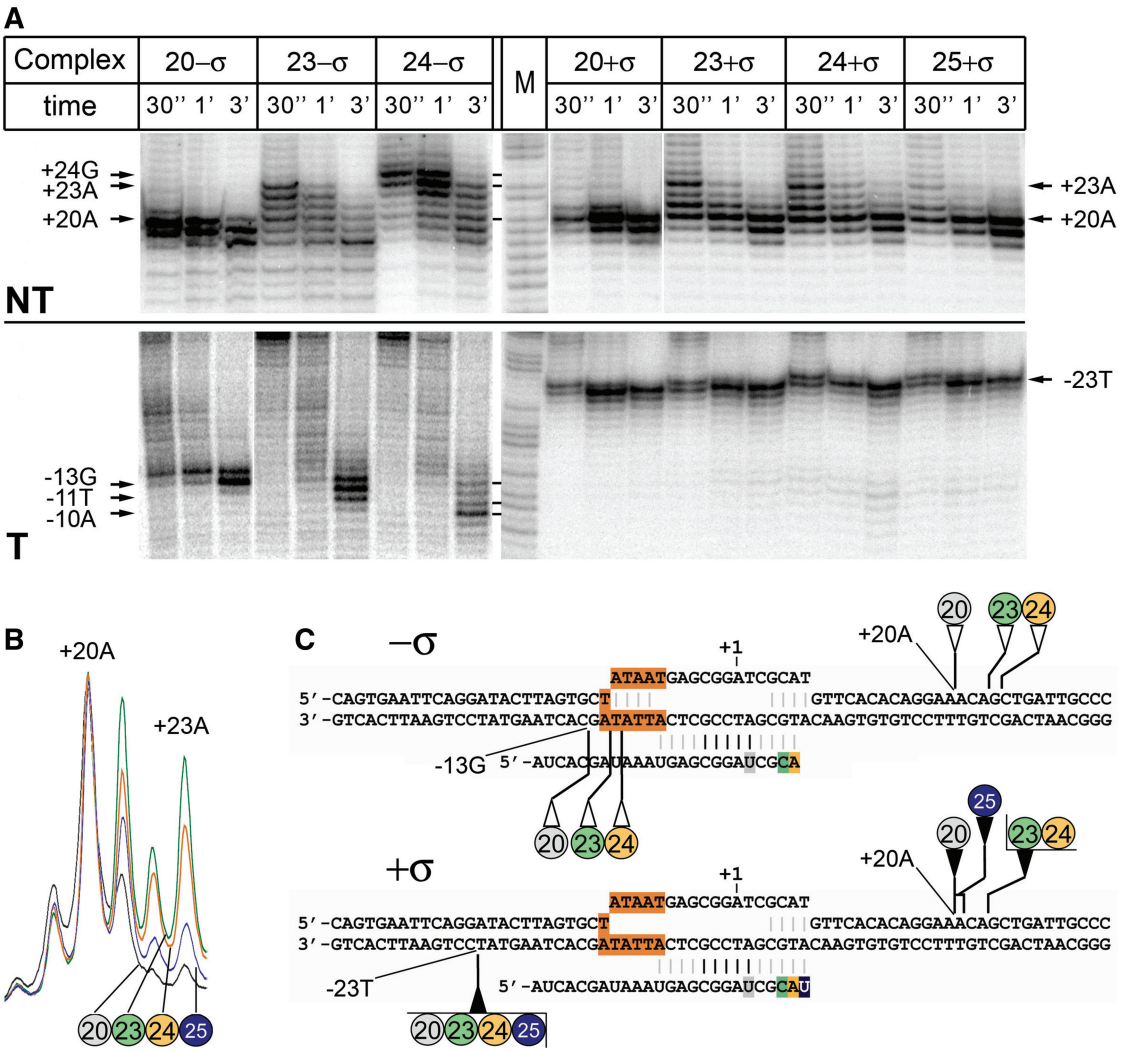

Figure 4. Analysis of the TEC conformation during σ^{70}-dependent pausing by ExoIII footprinting. (**A**) ExoIII footprinting on the non-template (up) and template (bottom) DNA strands in TECs obtained either in the absence or in the presence of the σ^{70} subunit. 'M' is an A+G cleavage marker. (**B**) ExoIII footprinting profiles of the front TEC border on the non-template DNA strand in σ^{70}-containing complexes [the data for the 30'' point from (A)]. The profiles are normalized by the signal intensities at position +20 A. (**C**) The summary of the footprinting results for σ-less (up) and σ^{70}-containing (bottom) complexes. Complementary interactions that change during transcription at the pause site are shown with gray lines. The downstream-most positions of the ExoIII stops in different TECs on the template and non-template DNA strands are indicated below and above the scaffolds, respectively.

the σ^{70}-dependent pausing on this scaffold was indistinguishable from the pausing on the original 'Cons' template.

We first analyzed changes in downstream TEC positions on the non-template DNA strand in complexes approaching the pause site (Figure 4A, upper panel). The front TEC borders in 20-mer complexes were identical both in the presence and in the absence of the σ^{70} subunit. In both cases, the major ExoIII stops on the non-template DNA strand were revealed before adenine residue located at position +20 (+20A).

In the 23 and 24-mer TECs obtained in the absence of σ^{70}, the front TEC border moved 3–4 nucleotides downstream, with the downstream-most ExoIII stops located at positions +23A and +24G, respectively, indicative of gradual RNAP translocation. In the 23 and 24-mer TECs containing σ^{70}, the front TEC border also moved downstream, with ExoIII stops observed down to position +23A (Figure 4A, the 30'' time point, and Figure 4B). Thus, a fraction of 23 and 24-mer complexes adopted a forward-translocated conformation, in which several additional nucleotides of the downstream DNA were placed within the RNAP molecule.

At the same time, in agreement with the data on RNA cleavage, a significant fraction of the σ^{70}-containing 23 and 24-mer TECs adopted a backtracked conformation, as revealed by ExoIII stops at position +20A. Furthermore, when the incubation of the σ^{70}-containing 23 and 24-mer TECs with ExoIII was prolonged to 3 min, the downstream stops at position +23A disappeared, with

the major stops revealed at position +20A (Figure 4A). Thus, the stressed TECs are intrinsically unstable and ExoIII can likely stimulate backward RNAP translocation in these complexes.

In agreement with its high sensitivity to Gre-induced RNA cleavage, the 25-mer paused TEC predominantly adopted the backtracked conformation, since the major ExoIII stops were revealed at positions +20A and +21A, with no significant signal corresponding to the downstream position +23A (Figure 4A–C).

We then analyzed changes in the upstream TEC positions on the template DNA strand during pausing (Figure 4A, bottom panel). In the absence of the σ^{70} subunit, the major ExoIII stops in the starting 20-mer complex were observed before positions −14C and −13G. These positions correspond to the positions of upstream ExoIII stops observed previously in elongation complexes [e.g. (34)] and in promoter-proximal transcription complexes lacking σ^{70} (35). In the 23 and 24-mer TECs, the ExoIII stops gradually moved to position −10A, corresponding to downstream TEC translocation. Several adjacent ExoIII stops were detected in stalled 23 and 24-mer σ-less TECs confirming that these complexes likely oscillate between different translocation states.

In σ^{70}-containing 20-mer TEC, the rear ExoIII barrier was shifted to position −23T, 10 nucleotides upstream in comparison with the σ-less 20-mer TEC, indicating that the σ^{70} binding increases RNAP contacts with DNA upstream of the −10-like element (Figure 4A and C). This position is identical to the positions of upstream ExoIII barriers that were observed previously in σ^{70}-containing promoter-proximal paused complexes (35,36). As opposed to the σ-less reactions, the patterns of upstream ExoIII digestions in σ^{70}-containing TECs were identical for stalled 20, 23, 24-mer and paused 25-mer complexes (Figure 4A and C). Importantly, the relative intensity of signals at positions −14C to −10A, corresponding to σ-less TECs, did not exceed 3–5% confirming that essentially all complexes contained the σ^{70} subunit (Figure 4A). Thus, the rear TEC border does not move during transcription at the pause site, resulting in formation of stressed TECs in which the DNA footprint is extended in downstream direction.

To understand the nature of the upstream ExoIII barrier in σ^{70}-containing TECs, we performed experiments with σ^{70} fragment 1–448 lacking regions 3 and 4 that might participate in interactions with the upstream DNA during pausing (10,20). In agreement with published data (23), this fragment induced efficient pausing (Supplementary Figure S8A). Positions of upstream ExoIII stops in both 20-mer and 23-mer TECs containing this fragment were similar to those observed with the wild-type σ^{70} subunit, the only difference being a 2–3 nt downstream shift in these positions at long reaction times (Supplementary Figure S8B). Thus, the upstream shift of the ExoIII barrier in σ^{70}-containing TECs cannot be explained by direct protection of the upstream DNA by the C-terminal part of σ^{70} but may result from interactions of σ region 2 with the −10-like element and/or from σ-induced changes in the TEC architecture and core RNAP-DNA interactions.

To further characterize these changes, we performed DNaseI footprinting experiments with stalled TECs obtained in the absence or in the presence of σ^{70}. Core RNAP footprint in σ-less 20-mer TEC extended from −19T to about +24G on the non-template DNA strand and from −14C to +17C on the template strand (Supplementary Figure S9). RNA extension in 23 and 24 σ-less TECs was accompanied by downstream movement of the rear footprint border (to positions −15T and −10A on the non-template and template strands, respectively). The pattern of downstream DNA protection was not changed significantly, probably due to the absence of suitably positioned DNaseI-sensitive sites in this region. The σ^{70} subunit increased DNA protection in the 20-mer TEC in the upstream direction, to positions −23A and −25C on the non-template and template strands (Supplementary Figure S9). Significantly, this σ-dependent protection did not change in 23 and 24-mer TECs (Supplementary Figure S9). Overall, the positions of the DNaseI footprint were in agreement with the data of the ExoIII footprinting. Thus, the upstream RNAP–DNA contacts in the pausing complex are extended as a result of σ^{70} binding and do not change during transcription at the pause site.

Core RNAP mutations stimulate σ^{70}-dependent pausing

Previously, a number of *E. coli* RNAP mutations were described that decreased σ^{70}-dependend pausing by changing the interactions of σ^{70} with core RNAP, including L402F, E407K, N409D substitutions in σ^{70} region 2.2 (24,25) and the R275Q substitution in the β′ coiled-coil motif in core RNAP (24). We supposed that RNAP mutations that would affect interactions of RNAP with the DNA template (in particular, those that would change the structure or conformation of the clamp domain of RNAP) may, on the contrary, stimulate σ^{70}-dependent pausing, by affecting the stability of stressed TECs that are formed during DNA scrunching. To test this conjecture, we analyzed four mutations in the β′ subunit of *E. coli* core RNAP that were previously shown to change its interactions with the DNA template (Figure 5): (1) the R339A substitution in the switch2 (SW2) region that interacts with the template DNA strand at the active site and connects the clamp domain with the main RNAP body, (2) insertion of eight amino acids (His_6GlnLeu) at position 216 in the clamphead that interacts with the downstream DNA duplex, (3) deletion in the downstream jaw domain (Δ1149–1190) that is located close to the downstream DNA duplex in the *T. thermophilus* TEC structure, and (4) deletion of the *E. coli*-specific SI3 domain (Δ943–1130) that is inserted in the trigger loop and was modeled near the downstream DNA duplex (37–39). All four mutations were previously shown to significantly decrease the stability of transcription complexes; the ΔJaw and ΔSI3 deletions also impaired recognition of *his* and *ops* pause signals during elongation (29,30,40,41).

We found that the mutations affected the σ^{70}-dependent pausing in two ways. First, the R339A and, to a lesser extent, Ω216 and ΔSI3 mutations increased the efficiency

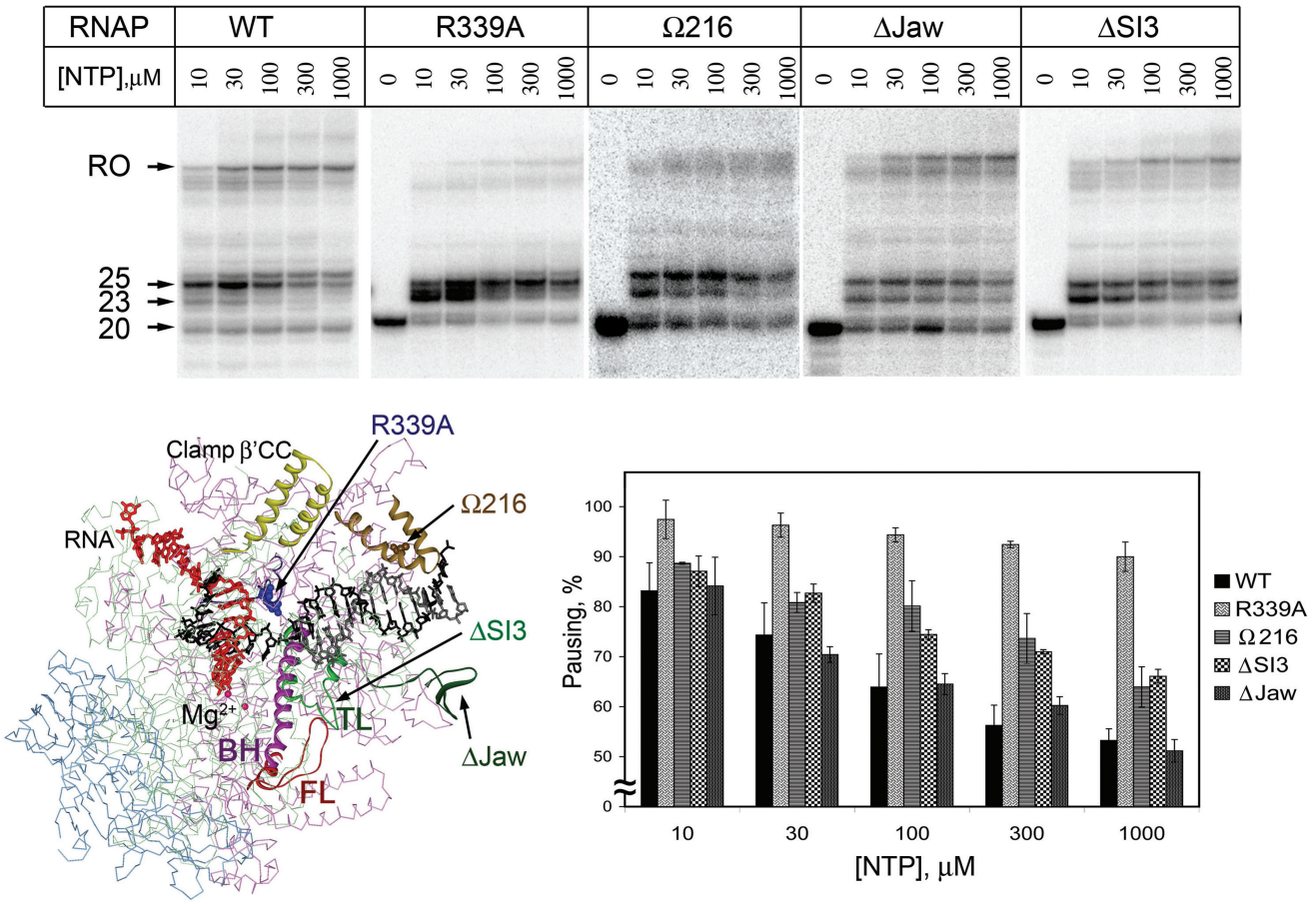

Figure 5. Effects of mutations in *E. coli* core RNAP on transcription pausing. The reaction was performed on the 'Cons' scaffold template at different NTP concentrations for 1′ at 37°C. The plot shows averages and standard deviations from two to three independent experiments. Location of the mutations on the *T. thermophilus* TEC structure [2PPB, (48)] is shown in the left bottom part of the figure. The template and non-template DNA strands are black and gray, respectively; RNA is red. Trigger loop (TL), bridge helix (BH) and F-loop (FL) in the RNAP active center are shown in green, magenta and red, respectively. The β′ coiled-coil (β′CC) region of the clamp domain is shown in yellow, SW2 is blue (with R339 residue shown as a CPK model), the ΔJaw deletion is dark green, the region of Ω216 insertion is brown; the site of the ΔSI3 deletion in the TL is indicated with an arrow.

of pausing (Figure 5). This effect was especially pronounced at high NTP concentrations; for example, at 1 mM NTPs the pausing efficiency of the R339A, Ω216 and ΔSI3 RNAP reached 90, 64 and 66% as compared to 55% for wild-type RNAP. At the same time, the ΔJaw RNAP did not reveal changes in the pausing efficiency. Second, all four mutations affected the pattern of pausing. Even at high NTP concentrations (≥100 μM), the pausing by the mutant RNAPs was observed in 23 and 25-mer TECs while the wild-type *E. coli* RNAP paused in 25 and 26-mer TECs (Figure 5). This indicated that the mutations might stimulate premature RNAP backtracking during transcription at the pause site.

To understand the effects of the mutations on pausing in more detail we chose two RNAPs, R339A and ΔSI3, for further characterization. Control experiments confirmed that the increased pausing by these mutant RNAPs was dependent on the σ^{70} subunit (Supplementary Figure S10A). The paused complexes formed by the R339A RNAP were highly susceptible to GreB-induced RNA cleavage, indicative of the backtracked TEC conformation (Supplementary Figure S10B). The ΔSI3 RNAP was previously shown to be resistant to stimulation of cleavage by Gre-factors (42). However, we observed that GreB increased the rate of RNA cleavage by the ΔSI3 RNAP in the paused complexes about 10- to 20-fold, suggesting that they were also backtracked (Supplementary Figure S10B). Finally, we observed that the mutant RNAPs were characterized by lower efficiency of promoter escape and synthesized higher amounts of short abortive RNAs than wild-type *E. coli* RNAP, indicative of problems with DNA scrunching during transcription initiation (Supplementary Text and Figure S10C).

Oligonucleotides complementary to the RNA 5′-end suppress pausing

Since the σ^{70}-dependent pausing is associated with TEC backtracking, we analyzed whether factors that prevent backtracking can suppress the pausing. Previously, it was shown that oligonucleotides complementary to the

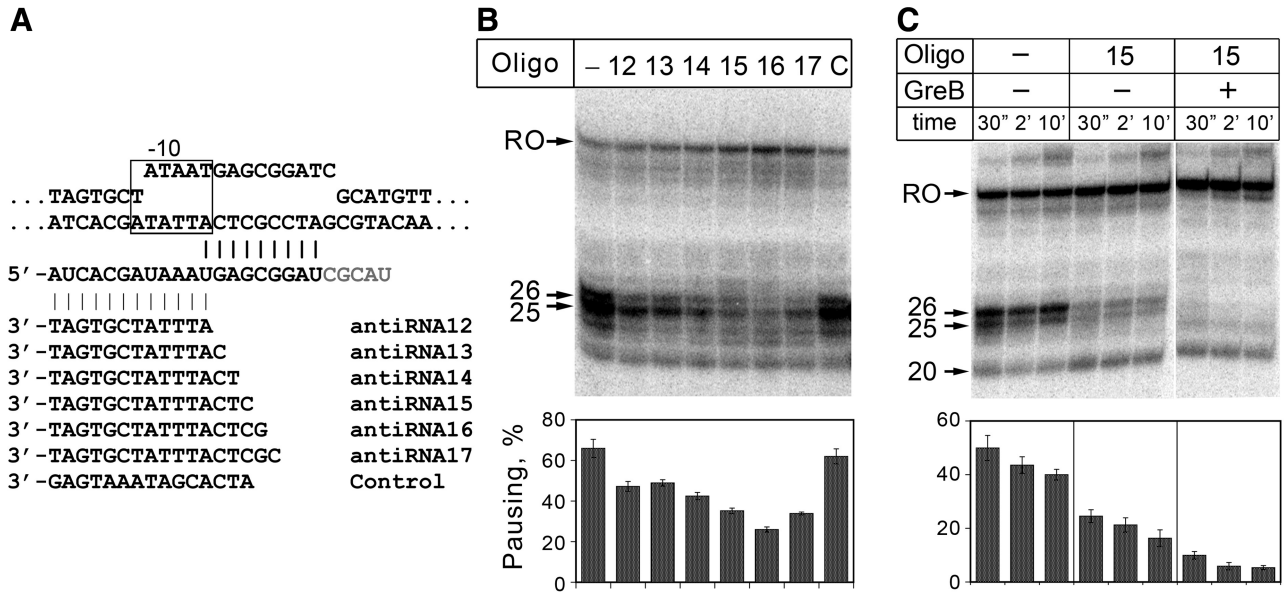

Figure 6. The effect of antisense oligonucleotides on σ^{70}-dependent pausing. (**A**) DNA oligonucleotides used in the study are shown below the 'Cons' scaffold. The −10-like element is boxed; RNA nucleotides 21 through 25 are shown in gray. (**B**) Transcription pausing by wild-type *E. coli* RNAP in the presence of oligonucleotides. The reaction was performed in the presence of 100 μM NTPs for 1′ at 25°C. (**C**) Kinetics of the pausing in the presence of the oligonucleotide antiRNA15 and GreB measured at 1 mM NTPs. The plots show the efficiencies of pausing at positions 25 and 26.

5′-segment of the nascent RNA inhibit TEC backtracking, likely by creating a physical barrier to backward RNAP movement (34). We therefore tested the effects of such oligonucleotides on the σ^{70}-dependent pausing on the 'Cons' scaffold template (Figure 6A). We found that antisense oligonucleotides ranging in length from 12 to 17 nucleotides decreased the efficiency of pausing, with a parallel increase in the full-length RNA synthesis (Figure 6B). The strongest effect was observed for 15 and 16 nt long oligonucleotides (antiRNA15 and antiRNA16), while a control 15 nt oligonucleotide, complementary to antiRNA15, did not affect the pausing. Analysis of the pausing kinetics in the presence of the antiRNA15 oligonucleotide revealed a significant decrease in the pausing efficiency at all time points tested (Figure 6C). The remaining paused complexes were susceptible to GreB-induced cleavage and, therefore, adopted a backtracked conformation, which likely made them resistant to the antipausing activity of the oligonucleotides (Figure 6C). Thus, the suppressing effect of oligonucleotides on σ^{70}-dependent pausing is likely due to their complementary interactions with the exiting RNA transcript and inhibition of TEC backtracking.

We then tested the effects of oligonucleotides on the pausing by the mutant variants of *E. coli* RNAP that are characterized by increased pausing. Similarly to the wild-type RNAP, the efficiencies of pausing by R339A and ΔSI3 RNAPs were decreased in the presence of the antiRNA15 oligonucleotide (Supplementary Figure S11). Thus, the defects of the mutant RNAPs in transcription through the pause site can be partially alleviated when the backtracking is suppressed.

DISCUSSION

The σ^{70}-dependent promoter-proximal pausing recently emerged as one of major types of transcription pausing in *E. coli*. Although the functional importance of such pausing remains to be established, available data suggest that it may affect TEC structure and transcription efficiency on at least a subset of genes (8,13,14,21). In this work, using an *in vitro* TEC reconstitution approach, we investigated structural rearrangements of the TEC during σ^{70}-dependent pausing and analyzed the influence of various factors on the process of pausing.

The general scheme of the σ-dependent pausing, summarizing the data from this and previously published works [reviewed in (8)], is presented on Figure 7. The recognition of the pause-inducing sequence by the σ subunit that presumably occurs in the relaxed TEC conformation is followed by RNA extension that leads to formation of stressed scrunched TECs. The scrunched complexes can then either escape the pause site and continue elongation, or adopt a paused scrunched conformation, or isomerize to paused backtracked conformation. The backtracked complexes can be reactivated through factor-mediated RNA cleavage at the RNAP active center, followed by reiterated RNA elongation; reverse isomerization to the scrunched state is probably possible at certain conditions. As was recently demonstrated, σ may remain bound to TEC following escape from the pause site thus possibly affecting TEC properties at later steps of elongation (14).

In our work, we obtained a direct experimental evidence supporting formation of stressed scrunched TECs during σ^{70}-dependent pausing. It was demonstrated that RNA elongation at the pause site is accompanied by melting of the downstream edge of the transcription bubble and

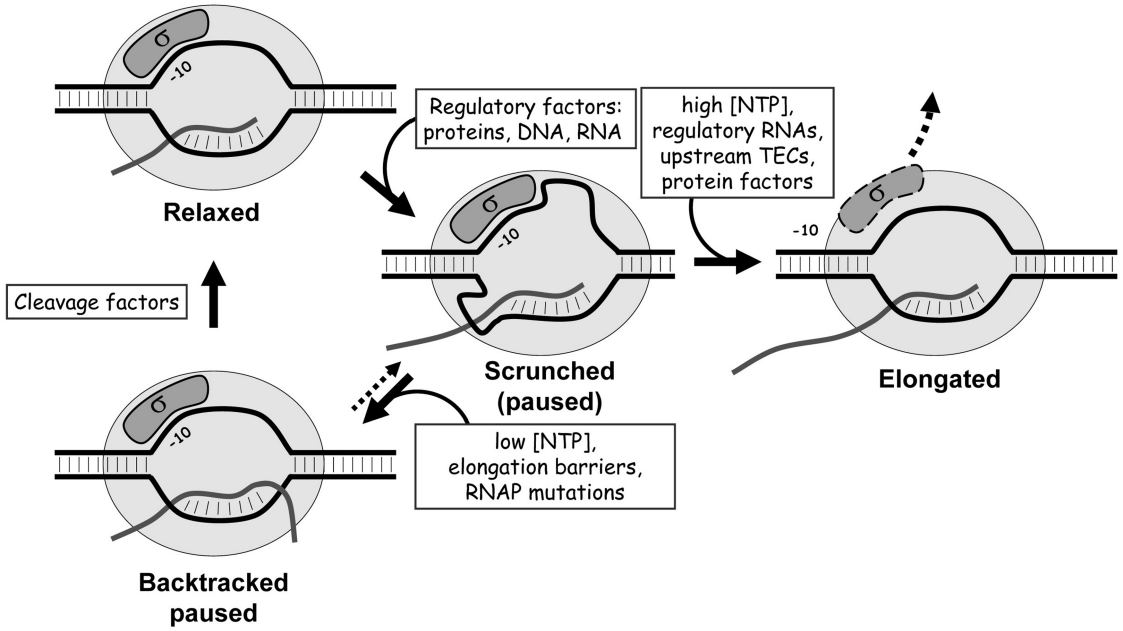

Figure 7. RNAP pausing through DNA scrunching and backtracking. See the text for details.

simultaneous movement of the front TEC border, which can be detected in a significant fraction of complexes. A similar fraction of TECs remain resistant to Gre-induced RNA cleavage confirming that they adopt a forward-translocated conformation. Importantly, the downstream extension of RNAP footprint is not accompanied by any detectable changes in the upstream RNAP–DNA contacts suggesting that excess DNA is scrunched within the RNAP molecule.

We demonstrated that RNA synthesis continues until the transcript reaches a critical length (with the last nucleotide added 13–14 nt downstream of the −10 element), after which all the complexes become highly prone to a backtracking-like transition, as manifested in their increased sensitivity to Gre-induced RNA cleavage. During such transition, the front TEC border moves upstream along the DNA template while the rear border does not change its position, as revealed in the ExoIII footprinting experiments. The extent of backtracking is strictly determined by the σ^{70} subunit, so that the RNAP active center is brought to the optimal relaxed position downstream to the −10 element. Previously, some of the paused complexes formed on the late phage promoters (e.g. +16 complex on $\lambda P_R'$) were proposed to be in the scrunched state, since they displayed decreased sensitivity to Gre-induced RNA cleavage (9,10). In our work, all paused complexes were highly susceptible to GreB-induced cleavage and, therefore, were backtracked. Similarly, the paused complexes formed on the *lac*UV5 promoter (11,12), and on the phage 82 $P_{R'}$ promoter (9) were also sensitive to Gre-induced cleavage. It remains to be established at which conditions the scrunched complexes can exist in a paused state similar to that observed on the $\lambda P_R'$ promoter template.

The transcription and footprinting analyses of TECs during σ^{70}-dependent pausing performed in this work revealed a mixed population of complexes in active, scrunched and backtracked states. This likely reflects an intrinsic stochastic nature of the pausing process during which each TEC can isomerize to various conformation species thus providing a basis for elaborate transcription regulation. Our results demonstrate that TEC transformations during the pausing may be affected by multiple factors, including template structure, nucleotide concentrations, RNAP mutations, regulatory proteins and non-coding nucleic acids (Figure 7).

In particular, recognition of the pause-inducing signal may depend on σ^{70} interactions with the −10 (7), −35-like elements (10) and, as demonstrated in this work, with the TG motif. Recently, it was shown that an additional sequence element, an A/T-rich backtrack-inducing sequence found in the lambdoid phage promoters stimulates pausing, likely by destabilizing RNA/DNA hybrid at the pause site and promoting TEC backtracking (10). We showed that the A/T-rich pause-site sequence can stimulate TEC backtracking and pausing even in the absence of the σ^{70} subunit and confirmed recent hypothesis that the σ^{70} subunit can bind backtracked TEC at the pause site (10). However, in our experimental system, efficient pausing was observed even when a G/C-only segment was present at the same place, although it did not stimulate TEC backtracking. Thus, the requirement for the backtrack-inducing sequence may be alleviated when a strong consensus σ^{70}-binding pause-inducing sequence is present.

As a first step to identification of RNAP elements that are important for formation of scrunched complexes, we showed that mutations in four *E. coli* RNAP regions (SW2, clamphead, SI3 and Jaw domains) stimulated pausing. Importantly, neither of the mutant RNAPs revealed major defects in transcription elongation under optimal conditions; moreover, the ΔJaw and ΔSI3

RNAPs displayed increased elongation rates (29,30,40,41). Most likely, the mutations destabilize stressed TECs by changing RNAP contacts with the DNA template at the active center and downstream of it, thus affecting the balance between DNA scrunching and TEC backtracking (see the model on Figure 5). The ΔSI3 mutation may also affect pausing through changes in the conformation of the trigger loop in the RNAP active center (41). Further experiments are required to establish the detailed nature of the proposed scrunching-associated defects of the mutant RNAPs and to identify other RNAP elements involved in DNA scrunching and σ^{70}-dependent pausing.

Additional factors can likely modulate pausing through changing the balance between productive elongation and TEC backtracking at the pause site. For example, the pausing is decreased at high NTP concentrations and in the presence of antisense oligonucleotides, likely as a result of suppressing the RNAP backtracking. Other factors that could suppress pausing by reducing backtracking include regulatory antisense RNAs, upstream TECs (43,44), and various protein factors. Thus, the λ Q protein was proposed to preferentially target a scrunched paused TEC and to decrease the pause half-life by an unknown mechanism (9). On the opposite, factors that stimulate backtracking should enhance σ-dependent pausing. In addition to low NTP concentrations and RNAP mutations, these factors may include elongation barriers of various natures, such as stalled transcription and replication complexes and various DNA-binding proteins.

In addition to the σ-subunit, other factors may also induce formation of the scrunched TECs and stimulate transcription pausing via a similar mechanism. In particular, it has been shown that scrunched complexes are likely formed during transcription through the *ops* pause site, with the participation of the transcription elongation factor RfaH (45). It remains to be tested whether other DNA- and RNA-encoded signals and protein factors, including alternative σ subunits, can also induce scrunching.

Recent work demonstrated that transcription of many eukaryotic genes, including highly regulated *Drosophila* and mammalian genes, is also regulated by promoter-proximal pausing of RNAP II following transcription initiation (15,16). While the detailed mechanism of this pausing remains to be uncovered, analysis of the pause-site sequences revealed a G/C-rich segment at the site of pausing, followed by A/T-rich stretches. It was therefore hypothesized that such sequences can probably stimulate transient RNAPII backtracking, which allows for binding of pause-specific transcription factors NELF and DSIF (Spt4/5) (15). In support of this view, the paused RNAPII complexes were shown to be susceptible to TFIIS-dependent RNA cleavage (46). Alternatively, the backtrack-inducing sequence in the initially transcribed region of eukaryotic genes may stimulate RNAP backtracking after the binding of NELF and DSIF has occurred. Thus, the molecular mechanism of eukaryotic promoter-proximal pausing may be strikingly similar to the σ-dependent pausing in bacteria. Comparative analysis of structural transitions of the TEC during promoter-proximal pausing in various species is an exciting area for future research.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Text, Supplementary Figures 1–12 and Supplementary References [29 and 47].

ACKNOWLEDGEMENTS

The authors thank I. Artsimovitch for critical reading of the manuscript, helpful discussions, plasmids and proteins.

FUNDING

Russian Academy of Sciences Presidium Program in Molecular and Cellular Biology (to A.K.); the grant of the President of Russian Federation (MD-618.2011.4); the Russian Foundation for Basic Research (grant number 10-04-00925); Federal Targeted Program 'Scientific and scientific-pedagogical personnel of innovative Russia 2009–2013' (state contract 02.740.11.0771); INSERM and CNRS (to K.B.). Funding for open access charge: RFBR 10-04-00925.

Conflict of interest statement. None declared.

REFERENCES

1. Landick,R. (2006) The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.*, **34**, 1062–1066.
2. Roberts,J.W., Shankar,S. and Filter,J.J. (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.*, **62**, 211–233.
3. Selth,L.A., Sigurdsson,S. and Svejstrup,J.Q. (2010) Transcript elongation by RNA polymerase II. *Annu. Rev. Biochem.*, **79**, 271–293.
4. Artsimovitch,I. and Landick,R. (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl Acad. Sci. USA*, **97**, 7090–7095.
5. Santangelo,T.J. and Artsimovitch,I. (2011) Termination and antitermination: RNA polymerase runs a stop sign. *Nat. Rev. Microbiol.*, **9**, 319–329.
6. Peters,J.M., Vangeloff,A.D. and Landick,R. (2011) Bacterial transcription terminators: the RNA 3′-end chronicles. *J. Mol. Biol.*, **412**, 793–813.
7. 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.
8. Perdue,S.A. and Roberts,J.W. (2011) sigma(70)-dependent transcription pausing in *Escherichia coli*. *J. Mol. Biol*, **412**, 782–792.
9. 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.
10. 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.
11. Brodolin,K., Zenkin,N., Mustaev,A., Mamaeva,D. and Heumann,H. (2004) The sigma 70 subunit of RNA polymerase induces lacUV5 promoter-proximal pausing of transcription. *Nat. Struct. Mol. Biol.*, **11**, 551–557.

12. 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.
13. Hatoum,A. and Roberts,J. (2008) Prevalence of RNA polymerase stalling at *Escherichia coli* promoters after open complex formation. *Mol. Microbiol.*, **68**, 17–28.
14. 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.
15. Nechaev,S. and Adelman,K. (2011) Pol II waiting in the starting gates: regulating the transition from transcription initiation into productive elongation. *Biochim. Biophys. Acta*, **1809**, 34–45.
16. Levine,M. (2011) Paused RNA polymerase II as a developmental checkpoint. *Cell*, **145**, 502–511.
17. Mukhopadhyay,J., Kapanidis,A.N., Mekler,V., Kortkhonjia,E., Ebright,Y.W. and Ebright,R.H. (2001) Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell*, **106**, 453–463.
18. Raffaelle,M., Kanin,E.I., Vogt,J., Burgess,R.R. and Ansari,A.Z. (2005) Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo. *Mol. Cell*, **20**, 357–366.
19. Kapanidis,A.N., Margeat,E., Laurence,T.A., Doose,S., Ho,S.O., Mukhopadhyay,J., Kortkhonjia,E., Mekler,V., Ebright,R.H. and Weiss,S. (2005) Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis. *Mol. Cell*, **20**, 347–356.
20. Mooney,R.A., Darst,S.A. and Landick,R. (2005) Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell*, **20**, 335–345.
21. Mooney,R.A., Davis,S.E., Peters,J.M., Rowland,J.L., Ansari,A.Z. and Landick,R. (2009) Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell*, **33**, 97–108.
22. 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.
23. Zenkin,N., Kulbachinskiy,A., Yuzenkova,Y., Mustaev,A., Bass,I., Severinov,K. and Brodolin,K. (2007) Region 1.2 of the RNA polymerase sigma subunit controls recognition of the -10 promoter element. *EMBO J.*, **26**, 955–964.
24. Sevostyanova,A., Svetlov,V., Vassylyev,D.G. and Artsimovitch,I. (2008) The elongation factor RfaH and the initiation factor sigma bind to the same site on the transcription elongation complex. *Proc. Natl Acad. Sci. USA*, **105**, 865–870.
25. 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.
26. Nickels,B.E., Garrity,S.J., Mekler,V., Minakhin,L., Severinov,K., Ebright,R.H. and Hochschild,A. (2005) The interaction between sigma70 and the beta-flap of *Escherichia coli* RNA polymerase inhibits extension of nascent RNA during early elongation. *Proc. Natl Acad. Sci. USA*, **102**, 4488–4493.
27. Kapanidis,A.N., Margeat,E., Ho,S.O., Kortkhonjia,E., Weiss,S. and Ebright,R.H. (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, **314**, 1144–1147.
28. Revyakin,A., Liu,C., Ebright,R.H. and Strick,T.R. (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, **314**, 1139–1143.
29. Pupov,D., Miropolskaya,N., Sevostyanova,A., Bass,I., Artsimovitch,I. and Kulbachinskiy,A. (2010) Multiple roles of the RNA polymerase {beta}' SW2 region in transcription initiation, promoter escape, and RNA elongation. *Nucleic Acids Res.*, **38**, 5784–5796.
30. Kulbachinskiy,A.V., Ershova,G.V., Korzheva,N.V., Brodolin,K.L. and Nikiforov,V.G. (2002) Mutations in β' subunit of the *Escherichia coli* RNA polymerase influence interaction with the downstream DNA duplex in the elongation complex. *Genetika*, **38**, 1207–1211.
31. Borukhov,S. and Goldfarb,A. (1993) Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and in vitro assembly. *Protein. Expr. Purif.*, **4**, 503–511.
32. Sosunova,E., Sosunov,V., Kozlov,M., Nikiforov,V., Goldfarb,A. and Mustaev,A. (2003) Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proc. Natl Acad. Sci. USA*, **100**, 15469–15474.
33. Sidorenkov,I., Komissarova,N. and Kashlev,M. (1998) Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol. Cell*, **2**, 55–64.
34. Komissarova,N. and Kashlev,M. (1997) Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3′ end of the RNA intact and extruded. *Proc. Natl Acad. Sci. USA*, **94**, 1755–1760.
35. Devi,P.G., Campbell,E.A., Darst,S.A. and Nickels,B.E. (2010) Utilization of variably spaced promoter-like elements by the bacterial RNA polymerase holoenzyme during early elongation. *Mol. Microbiol.*, **75**, 607–622.
36. 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.
37. Hudson,B.P., Quispe,J., Lara-Gonzalez,S., Kim,Y., Berman,H.M., Arnold,E., Ebright,R.H. and Lawson,C.L. (2009) Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proc. Natl. Acad. Sci. USA*, **106**, 19830–19835.
38. Chlenov,M., Masuda,S., Murakami,K.S., Nikiforov,V., Darst,S.A. and Mustaev,A. (2005) Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase beta' subunit. *J. Mol. Biol.*, **353**, 138–154.
39. Opalka,N., Brown,J., Lane,W.J., Twist,K.A., Landick,R., Asturias,F.J. and Darst,S.A. (2010) Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol.*, **8**, pii: e1000483.
40. Ederth,J., Artsimovitch,I., Isaksson,L.A. and Landick,R. (2002) The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J. Biol. Chem.*, **277**, 37456–37463.
41. Artsimovitch,I., Svetlov,V., Murakami,K.S. and Landick,R. (2003) Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.*, **278**, 12344–12355.
42. Zhang,J., Palangat,M. and Landick,R. (2010) Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat. Struct. Mol. Biol.*, **17**, 99–104.
43. Epshtein,V. and Nudler,E. (2003) Cooperation between RNA polymerase molecules in transcription elongation. *Science*, **300**, 801–805.
44. Epshtein,V., Toulme,F., Rahmouni,A.R., Borukhov,S. and Nudler,E. (2003) Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J.*, **22**, 4719–4727.
45. Belogurov,G.A., Sevostyanova,A., Svetlov,V. and Artsimovitch,I. (2010) Functional regions of the N-terminal domain of the antiterminator RfaH. *Mol. Microbiol.*, **76**, 286–301.
46. Nechaev,S., Fargo,D.C., dos Santos,G., Liu,L., Gao,Y. and Adelman,K. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. *Science*, **327**, 335–338.
47. Kulbachinskiy,A. and Mustaev,A. (2006) Region 3.2 of the sigma subunit contributes to the binding of the 3′-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *J. Biol. Chem.*, **281**, 18273–18276.
48. Vassylyev,D.G., Vassylyeva,M.N., Zhang,J., Palangat,M., Artsimovitch,I. and Landick,R. (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature*, **448**, 163–168.