ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/5915047

An Alternate Mechanism of Abortive Release Marked by the Formation of Very Long Abortive Transcripts †

ARTICLE *in* BIOCHEMISTRY · NOVEMBER 2007

Impact Factor: 3.02 · DOI: 10.1021/bi701236f · Source: PubMed

CITATIONS
5

READS
13

5 AUTHORS, INCLUDING:

Monica Chander
Bryn Mawr College
14 PUBLICATIONS **303** CITATIONS
SEE PROFILE

Lilian M Hsu
Mount Holyoke College
24 PUBLICATIONS **636** CITATIONS
SEE PROFILE

Available from: Lilian M Hsu
Retrieved on: 29 January 2016

An Alternate Mechanism of Abortive Release Marked by the Formation of Very Long Abortive Transcripts†

Monica Chander, Karyn M. Austin,‡ Nwe-Nwe Aye-Han,‡ Piya Sircar,‡ and Lilian M. Hsu*

Program in Biochemistry, Mount Holyoke College, South Hadley, Massachusetts 01075

Received June 22, 2007; Revised Manuscript Received August 11, 2007

ABSTRACT: The Eσ^{70}-dependent N25 promoter is rate-limited at promoter escape. Here, RNA polymerase repeatedly initiates and aborts transcription, giving rise to a ladder of short RNAs 2−11 nucleotides long. Certain mutations in the initial transcribed sequence (ITS) of N25 lengthen the abortive initiation program, resulting in the release of very long abortive transcripts (VLATs) 16−19 nucleotides long. This phenomenon is completely dependent on sequences within the first 20 bases of the ITS since altering sequences downstream of +20 has no effect on their formation. VLAT formation also requires strong interactions between RNA polymerase and the promoter. Mutations that change the −35 and −10 hexamers and the intervening 17 base pair spacer away from consensus decrease the probability of aborting at positions +16 to +19. An unusual characteristic of the VLATs is their undiminished levels in the presence of GreB, which rescues abortive RNAs (≤15 nucleotides) associated with backtracked initial transcribing complexes. This suggests that VLATs are produced via a mechanism distinct from backtracking, which we propose entails polymerase molecules hyper forward translocating during the promoter escape transition. We discuss how certain features in the ITS, when combined with the N25 promoter, may lead to hyper forward translocation and abortive release at VLAT positions.

The macromolecular process of RNA transcription involves three sequential phases: initiation, elongation, and termination. Transcription initiation, the most highly regulated of the three phases, can be further broken down into multiple steps. Initiation begins when RNA polymerase binds to a promoter sequence to form a binary closed complex in which the DNA is double-stranded. In the second step, a 13−14 bp stretch of DNA extending from −11 to +2/+3 (relative to the transcription start site +1) is melted open to form an open complex (*1*, *2*). In the third step, RNA polymerase begins de novo RNA synthesis and, upon the production of a nascent transcript ∼8−10 nucleotides (nt)[1] long, undergoes promoter escape, which marks the end of initiation and the onset of elongation.

At many promoters, RNA polymerase repetitively performs abortive initiation and releases short RNAs ranging in size from 2 to 15 nt. Abortive initiation occurs due to the repetitive cycling of the enzyme back to the open complex conformation after the release of short RNAs and is particularly robust at promoters that interact strongly with RNA polymerase. Abortive initiation ceases when polymerase successfully negotiates escape and translocates away from the promoter, which requires release of promoter contacts, collapse of the initial open complex bubble, and proper placement of the nascent RNA in the RNA exit channel.

The reason for the high level of abortive initiation and weak efficiency of escape at strong promoters (*3*) has been clarified by two recent studies (*4*, *5*). The two groups showed that, during initial transcription, the initial transcribing complex (ITC) remains bound at the promoter while downstream DNA is continually unwound and pulled into the enzyme. This results in expansion of the transcription bubble, and the unwound DNA has to be accommodated inside the enzyme in a scrunched form. As the transcription bubble continues to grow, stress continues to build up inside the enzyme. Rewinding the transcription bubble at the upstream boundary relieves the cumulative stress, and the energy thus released is sufficient to disengage RNA polymerase from the promoter and allow escape. Thus, promoter release and bubble collapse occur concurrently during promoter escape. On occasions when the downstream edge of the bubble rewinds, the DNA and DNA−RNA hybrid track back past the enzyme active site, causing the 3′ end of the nascent RNA to become unannealed, protrude into the secondary channel, and result in transcriptional arrest. The arrested RNA can be rescued in the presence of transcription factors GreA and GreB, which bind in the secondary channel and stimulate the internal hydrolysis of the nascent RNA to realign a 3′-hydroxyl end at the active site (*6*). The newly generated 3′-hydroxyl end can now be extended, and transcription resumes (*7*, *8*). In the absence of the Gre factors, RNA polymerase is forced to abort transcription, release the

† This work was supported by an NSF grant (RUI-0418316) to L.M.H.

* To whom correspondence should be addressed. E-mail: lhsu@mtholyoke.edu. Phone: (413) 538-2609. Fax: (413) 538-2327.

‡ These authors contributed to this work while undergraduates at Mount Holyoke College.

[1] Abbreviations: AP, abortive probability; APR, abortive to productive ratio; α-CTD, carboxy-terminal domain of the α subunit; DIS, discriminator; FLB, formamide loading buffer; ITC(s), initial transcribing complex(es); ITS(s), initial transcribed sequence(s); nt, nucleotide; NTP, nucleoside triphosphates; PRR, promotr recognition region; σ, σ subunit; σ4, σ domain 4; σ2, σ domain 2; σ1.2, σ subdomain 1.2; VLATs, very long abortive transcripts.

10.1021/bi701236f CCC: $37.00 © 2007 American Chemical Society
Published on Web 10/11/2007

nascent RNA, reassume the open complex conformation, and reinitiate transcription.

It is clear from the above description that the extent of abortive cycling and the efficiency of promoter escape are dependent on how well RNA polymerase is anchored by the promoter recognition region (PRR; −60 to −1 relative to the transcription start site). In general, the tighter the association of RNA polymerase to DNA, the less well it escapes the promoter (*3*). The efficiency of promoter binding is determined primarily by the −35 and −10 hexamers, optimally separated by 17 bp. These sequences are directly contacted by the σ subunit of the holoenzyme—the −35 element via σ domain 4 (σ4) and the −10 element via σ domain 2 (σ2). σ subdomain 1.2 (σ1.2) also contacts the nontemplate strand of the discriminator sequence (DIS; the region connecting the −10 hexamer to the transcription start site) to stabilize the open complex. This interaction is optimal if a G is positioned 2 bp downstream of the −10 element (*9*). Footprint analysis shows that promoter escape is accompanied by the release of σ factor contacts with promoter DNA (*10*, *11*); thus, σ1.2/DIS, σ2/−10, and σ4/−35 contacts all need to be disrupted before polymerase can move away from the promoter. In addition, the σ3.2 linker, which occupies the RNA exit channel in the holoenzyme, presents a physical barrier to the growing RNA chain during initiation and needs to be displaced to allow passage of the nascent transcript (*12*, *13*). Some promoters also contain an A/T-rich sequence, known as the UP element, which interacts with the carboxy terminal domain of the α subunits (α-CTD) of RNA polymerase to further increase the association of RNA polymerase with DNA (*14*).

The second factor that governs abortive initiation and promoter escape is the initial transcribed sequence (ITS; +1 to +20 relative to the start site). This was initially demonstrated by changing the ITS of the phage T5 N25 promoter to the "anti" sequence (A ↔ C, G ↔ T), which impaired promoter escape and rendered the mutant promoter 10-fold weaker in productive RNA synthesis (*15*, *16*). Although at greatly reduced levels, RNA polymerase achieves escape at the $N25_{anti}$ promoter at position +16 in contrast to the wild-type N25 promoter, which supports earlier escape at position +12. $N25_{anti}$ is not uniquely compromised in promoter escape properties. Analysis of ~40 N25 random-ITS variants, with mutations encompassing positions +3 to +20, showed that they all delayed promoter escape to predominantly the +16 position, after production of an abortive ladder of 2−15 nt transcripts (*17*). A subset of these random-ITS mutants produced longer RNA transcripts extending to 19 or 20 nt, which were unusual for two reasons. First, at lengths >15 nt, the nascent RNA transcript should have maximized its interaction with the protein and emerged from the interior of the polymerase (*18*). Second, the level of these transcripts was not diminished in the presence of GreB, which rescues backtracked ITCs with nascent transcripts ≤15 nt long by cleavage and reextension, to facilitate escape (*17*, *19*). The latter observation suggests that the 16−19 nt transcripts are not located in the secondary channel and therefore are unlikely to be the products of RNA polymerase backtracking.

Here, we investigate the nature of these GreB-refractive transcripts and show that they are abortive RNAs. We further show that these transcripts arise predominantly at variant N25 promoters with specific changes in the ITS region encom-

Table 1: Primers Used in This Study

Random-ITS Mutagenic Primers	
ITS2-d (KPH)	5′-GGGTACCTGCAGAAGCTTTCTGCGAG AACCAGCCATATTTAAA CTCCTCTCN NNNNNNNATGAATCTATTATACAGAA AAATTTTCC-3′
N25-u (SSX)	5′-AAGGCCACCTAGGCCTCGAGGGAAAT CATAAAAAATTTATTTG CTTTCAGGA AAATTTTTCTGTATAATAGATTCAT-3′
Template Primers, DG203	
203T	5′-ACCTGCAGAAGCTTTCTGCGAGAACC AGCCATATTTAAACTCCTCTCCCGGT CGC-3′
Lac	5′-ACCTGCAGAAGCTTTCTATAGCTGTT TCCTGTGTGAAAACTCCTCTCCCGGT CGC-3′
T7A1	5′-ACCTGCAGAAGCTTTCTGGTGTCGAT TGGGATGGCTAAACTCCTCTCCCGGT CGC-3′
PR	5′-ACCTGCAGAAGCTTTCTTCAGGGTTA TGCGTTGTTCCAACTCCTCTCCCGGT CGC-3′
Nontemplate Primers, DG203	
DG203	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGAAA ATTTTTCTGTATA ATAGATTCATGCGACCGGGAGAGGAG TT-3′
UPcon	5′-CCTCGAGAAATCATAAAAAAGTAATT TGCTTTCAGGAAA ATTTTTCTGTATA ATAGATTCATGCGACCGGGAGAGGAG TT-3′
UPwk	5′-CCTCGAGAAATCATAAACAATTTATT TGCTTTCAGGAAAATTTTTCTGTATA ATAGATTCATGCGACCGGGAGAGGAGT T-3′
UPdel	5′-CCTCGAGAAATCATCCTAGGAATATT TGCTTTCAGGAAA ATTTTTCTGTATA ATAGATTCATGCGACCGGGAGAGGAG TT-3′
−35wk	5′-CCTCGAGAAATCATAAAAAATTTATG TGCTTTCAGGAAA ATTTTTCTGTATA ATAGATTCATGCGACCGGGAGAGGAG TT-3′
SP16	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGAAAA TTTTTCTGTATAA TAGATTCATGCGACCGGGAGAGGAGT T-3′
SP18	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGGAAA ATTTTTCTGTAT AATAGATTCATGCGACCGGGAGAGGA GTT-3′
−10wk	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGAAA ATTTTTCTGGATA ATAGATTCATGCGACCGGGAGAGGAG TT-3′
DIS5	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGAAA ATTTTTCTGTATA ATGATTCATGCGACCGGGAGAGGAGT T-3′
DIS7	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGAAA ATTTTTCTGTATA ATAGAATTCATGCGACCGGGAGAGGA GTT-3′
DIS AT	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGAAA ATTTTTCTGTATA ATAAATTCATGCGACCGGGAGAGGAG TT-3′
DIS GC	5′-CCTCGAGAAATCATAAAAAATTTATT TGCTTTCAGGAAA ATTTTTCTGTATA ATAGCGGCATGCGACCGGGAGAGGAG TT-3′

```
      -85                                                  -43
5′    CCCTC GAGGAATTCC CGGGGATCCG TCGAGGGAAA TCATAAAAAA
                  -35                                 -10
            TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGATTC

            +1
N25         ATAAATTTGA GAGAGGAGTT TAAATATGGC TGGTTCTCGC
                                    +67
            AGAAAGCTTC TGCAGGTACC CGGGCGG 3′

            +1                  +20
N25anti     ATCCGGAATC CTCTTCCCGG
DG203       ATGCGACCGG GAGAGGAGTT
DG204       ATAAGAGCTC GAGAGGAGTT
DG208       ATACGCGTGC GAGAGGAGTT
DG211       ATAAGTAAGC GAGAGGAGTT
DG212       ATAAGGCAGC GAGAGGAGTT
DG229       ATAGGGCACG GAGAGGAGTT
DG239       ATGGCTCCTA GAGAGGAGTT
DG241       ATCCCTGCTG GAGAGGAGTT
DG243       ATGTCCTTGC GAGAGGAGTT
DG247       ATGCACCCAG GAGAGGAGTT
```

FIGURE 1: Nucleotide sequence of T5 N25 and related DG200 series promoters. The promoter fragments span −85 to +67 and share identical upstream sequences from −85 to −1 but differ from N25 in the downstream sequence from +3 to +20 in the case of $N25_{anti}$ and from +3 to +10 in the case of the DG200 series ITS mutants. The promoter recognition sequences highlighted in bold include the UP (centered at −43), −35, and −10 elements. The ITS region from +1 to +20 is indicated in italics for each of the promoters.

passing +3 to +10, and their formation depends on strong contacts between RNA polymerase holoenzyme and promoter DNA. We speculate that the 16−19 nt abortive RNAs are produced by ITCs during the promoter escape transition via a mechanism distinct from backtracking.

MATERIALS AND METHODS

Enzymes and Proteins. RNAP holoenzyme, with a hexahistidine tag at the C-terminus of the β' subunit, was purified from *Escherichia coli* strain RL721 (a gift from Dr. R. Landick) as described in Uptain (*20*) and contained ∼60% active molecules at the time of use (*21*). The *E. coli* GreA and GreB proteins were isolated from IPTG-induced JM109 cells harboring the plasmids pDNL278 and pGF296, respectively (*22*).

Generation of Random-ITS Mutant Promoters. Degenerate mutations in the ITS from +3 to +10 were generated as described by Hsu et al. (*17*) using primers ITS2-d (KPH) and N25-u (SSX) (sequences shown in Table 1). Single ITS mutants, designated as the DG200 series, were obtained by cloning and purifying the clonal strains. The region spanning −85 to +67 was PCR-amplified from plasmid templates and purified for use in in vitro transcription assays (*17*).

Generation of DG203 Mutant Promoters. One of the random-ITS mutant promoters, DG203, was chosen for detailed analyses. The DG203 promoter and mutant derivatives were generated by hybridizing two primers overlapping at +3 to +20 and extending by Klenow Pol I. Upstream promoter mutations were incorporated into the nontemplate strand primer, and downstream mutations were incorporated into the template strand primer. The sequences of the template and nontemplate primers are shown in Table 1.

Steady-State Transcription, PAGE Analysis, and Quantitation. Quantitative transcription reactions were performed at 37 °C for the indicated times as described by Hsu et al. (*17*). ^{32}P-labeled RNAs were fractionated in 7 M urea/23% (10:1) polyacrylamide gels and quantified as described by Hsu et al. (*17*).

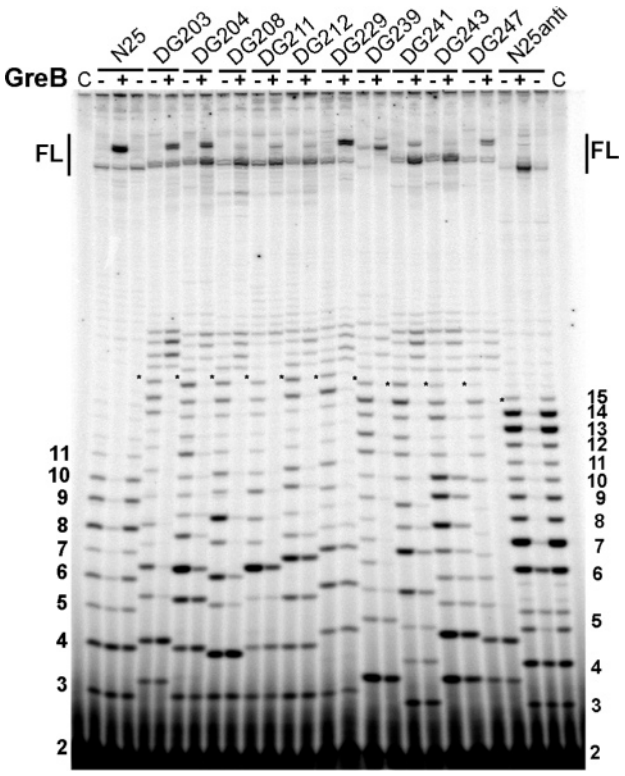


FIGURE 2: Gel profile of in vitro transcribed RNAs from DG200 promoters. Randomization of the N25 ITS spanning +3 to +10 results in the formation of very long abortive transcripts, 16−19 nt in length. The in vitro transcribed RNAs, each labeled once at the 5′-triphosphate, were prepared from 10 random-ITS promoters and fractionated on a 7 M urea, 23% (10:1) polyacrylamide gel. The promoters were transcribed with (lanes designated "+") or without (lanes designated "−") GreB for 10 min at 37 °C. Reactions contained 30 nM template DNA, 50 nM RNA polymerase, and 500 nM GreB (when present) in transcription buffer (50 mM Tris−HCl, pH 8, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, 200 mM KCl, 10 μg/mL acetylated bovine serum albumin), and 100 μM NTP with [γ-^{32}P]ATP added to ∼10 cpm/fmol. The abortive RNA ladder on the left is produced by N25 and the one on the right by $N25_{anti}$. "FL" indicates full-length RNA. Asterisks mark the 15 nt band in each lane to enable identification of the abortive RNAs. "C" indicates minus-enzyme control reactions to correct for [γ-^{32}P]-ATP background during quantitative analysis.

Solid-State Transcription. Solid-state transcription reactions were performed under conditions that allow RNA polymerase walking (*23*). Nickel−NTA agarose beads (Qiagen) were washed with five 1 mL aliquots of 1× transcription buffer/KCl (50 mM Tris−HCl, pH 8, 10 mM $MgCl_2$, 200 mM KCl, 10 mM β-mercaptoethanol, 10 μg/mL acetylated bovine serum albumin). When used in a transcription reaction, the washed beads accounted for 10% of the reaction volume. Steady-state transcription reactions (20 μL) were assembled by first forming open complexes immobilized on beads (5 min at room temperature with rotation), followed by the addition of nucleoside triphosphates (NTPs) to 100 μM and [γ-^{32}P]ATP to ∼10 cpm/fmol. Transcription was allowed to proceed at 37 °C for 30 min before termination with the addition of an equal volume of formamide loading buffer (FLB; 80% deionized formamide, 1× TBE, 10 mM Na_2EDTA, 0.08% (v/v) each of xylene cyanol and amaranth). For the supernatant and pellet fractions, a bead-based reaction was spun for 2 min in a capsule microfuge. The supernatant was withdrawn and mixed with an equal volume of FLB. The pellet was washed with 1 mL

Table 2: Quantitative Parameters for DG200 Series Promoters

	no GreB				with GreB		
template	PY[a] (%)	RIF[a] (%)	APR[a]	MSAT[a] (nt)	PY[a] (%)	APR[a]	MSAT[a] (nt)
N25	4.2 ± 0.4	100 ± 0	23 ± 2	11	13.9 ± 4.0	6 ± 1	7
$N25_{anti}$	1.1 ± 1.6	226 ± 111	120 ± 12	15	5.7 ± 2.5	17 ± 5	6
DG203	4.3 ± 0.4	146 ± 79	23 ± 2	19	13.0 ± 3.1	7 ± 2	4[b]
DG204	4.4 ± 1.3	225 ± 87	23 ± 8	19	10.9 ± 2.6	9 ± 2	6[b]
DG208	3.2 ± 0	174 ± 112	30 ± 0	19	9.4 ± 4.9	12 ± 6	6[b]
DG211	3.5 ± 0.1	219 ± 106	28 ± 1	19	8.7 ± 0.9	11 ± 1	6[b]
DG212	2.5 ± 0.4	143 ± 82	40 ± 7	19	8.7 ± 1.3	11 ± 2	6[b]
DG229	3.5 ± 1.4	103 ± 53	31 ± 13	19	20.0 ± 5.4	4 ± 1	7[b]
DG239	0.7 ± 0.1	208 ± 133	143 ± 20	19	6.7 ± 1.3	14 ± 3	5[b]
DG241	1.7 ± 0.2	97 ± 57	57 ± 6	19	13.0 ± 2.0	7 ± 1	8[b]
DG243	1.6 ± 0.8	256 ± 181	72 ± 36	19	5.8 ± 1.7	18 ± 6	10[b]
DG247	2.3 ± 0.7	61 ± 31	44 ± 14	19	6.4 ± 0.3	15 ± 1	4[b]

[a] The promoters were transcribed three times plus or minus GreB to obtain the quantitative parameters, which are reported as their mean value ± standard deviation. The productive yield (PY) is the productive RNA as a percentage of the total yield (productive + abortive RNAs). The relative initiation frequency (RIF) is the total number of RNA molecules (full-length and abortive) initiated from each promoter in the 10 min reaction period and is reported relative to that of N25. APR is the ratio of abortive to productive yield for a given promoter. MSAT is the maximum size of the abortive transcript. [b] These promoters showed GreB-mediated rescue at earlier template positions, but nevertheless produced very long abortive transcripts of 16−19 nt.

of 1× transcription buffer/KCl, and the wash was discarded. The final bead pellet was resuspended in 20 μL 1× transcription buffer/KCl and mixed with 20 μL of FLB. A control solution-based reaction was performed at the same time. Total RNAs affiliated with each fraction were analyzed on sequencing gels as above.

RESULTS

Mutations in N25 ITS Positions +3 to +10 Result in Very Long Abortive Transcripts. The ITS mutant derivatives of the N25 promoter reported here are designated as the DG200 series and deviate from the N25 promoter in the ITS region from +3 to +10 (Figure 1). In every case examined, these ITS mutants consistently produced transcripts 2−19 nt long (Figure 2). This was in contrast to the vast majority of the random-ITS mutants previously studied (the DG100 series), which produced abortive transcripts up to 15 nt in length (*17*). What might be the origin of the 16−19 nt RNAs? We ruled out ρ-independent termination, as there is no hairpin or terminator-like structure apparent in the first 20 nucleotides of any random-ITS promoter. We also ruled out the possibility that these are paused RNAs associated with elongating enzyme complexes since, unlike paused RNAs, the 16−19 nt RNAs were not eliminated by high NTP chase and were released by RNA polymerase into the supernatant (see below). We therefore conclude that the 16−19 nt transcripts produced on the DG200 promoters are aborted RNAs, and from here on out we shall refer to them as very long abortive transcripts (VLATs), in contrast to the short, medium, and long abortive transcripts previously characterized (*17*).

Although the abortive initiation program is lengthened at every one of the DG200 promoters examined, the efficiency of promoter escape is not necessarily compromised. This is apparent from the ratio of abortive to productive yield (APR) shown in Table 2, which is indicative of the ease or difficulty at promoter escape; the higher the APR value, the more difficult it is for RNA polymerase to clear the promoter. For N25, 1 in every 23 initiation events results in the synthesis of full-length RNA, while $N25_{anti}$, which is severely escape-impaired, only produces 1 full-length RNA in every 120 initiations (Table 2). Seven of the ten DG200 promoters examined showed APR values comparable to that of N25 (less than 2-fold difference), indicating that although escape is delayed at these promoters, RNA polymerase is nevertheless able to negotiate escape with the same efficiency as at the N25 promoter.

There is no apparent sequence feature/pattern in the ITSs of the DG200 promoters that would indicate why promoter escape is delayed at these promoters. The ITS changes in positions +3 to +10 only affect the promoter escape step without affecting the earlier steps in initiation. This is evident in the relative initiation frequency (Table 2), which indicates that, with the exception of DG247, the DG200 promoters support the same number of initiation events as the N25 promoter, which was also previously observed with the DG100 promoters (*17*). With the DG100 promoters, however, an inverse correlation was noted between productive yield and relative initiation frequency, so that RNA polymerase reinitiated less frequently on promoters with facilitated escape (*17*). The current set of DG200 promoters does not display any such correlation, positive or negative.

The abortive probability profiles of the ten DG200 promoters also showed no correlation between ITS composition and the abortive potential of RNA polymerase at the initial positions (Figure 3). The most highly abortive position for DG203, for example, is +2, for DG211 +6, and for DG241 +14 (Figure 3). In general one notes four major barriers to escape—the first at positions +2 to +3, the second at positions +6 to +8, the third at positions +13 to +15, and the fourth at positions +18 to +19 (Figure 3). The first three abortive blocks were also observed for the DG100 random-ITS mutants that mostly aborted to position +15 (*17*). The abortive barrier at positions +2 to +3 was attributed to inherently unstable ITCs with a short RNA−DNA heteroduplex, while the abortive blocks at +6 to +8 and +13 to +15 were credited to tight σ2/−10 and σ4/−35 polymerase−promoter interactions, respectively (*3*, *17*, *24*, *25*). Thus, when the nascent RNA grows to a length of 6−8 nt, it probes the strength of the σ2/−10 contacts and must break these contacts to transcribe further downstream (*3*, *24*). We now know that the first contact of a 6−8 nt RNA is to the N-terminal loop of the σ3.2 linker blocking the RNA

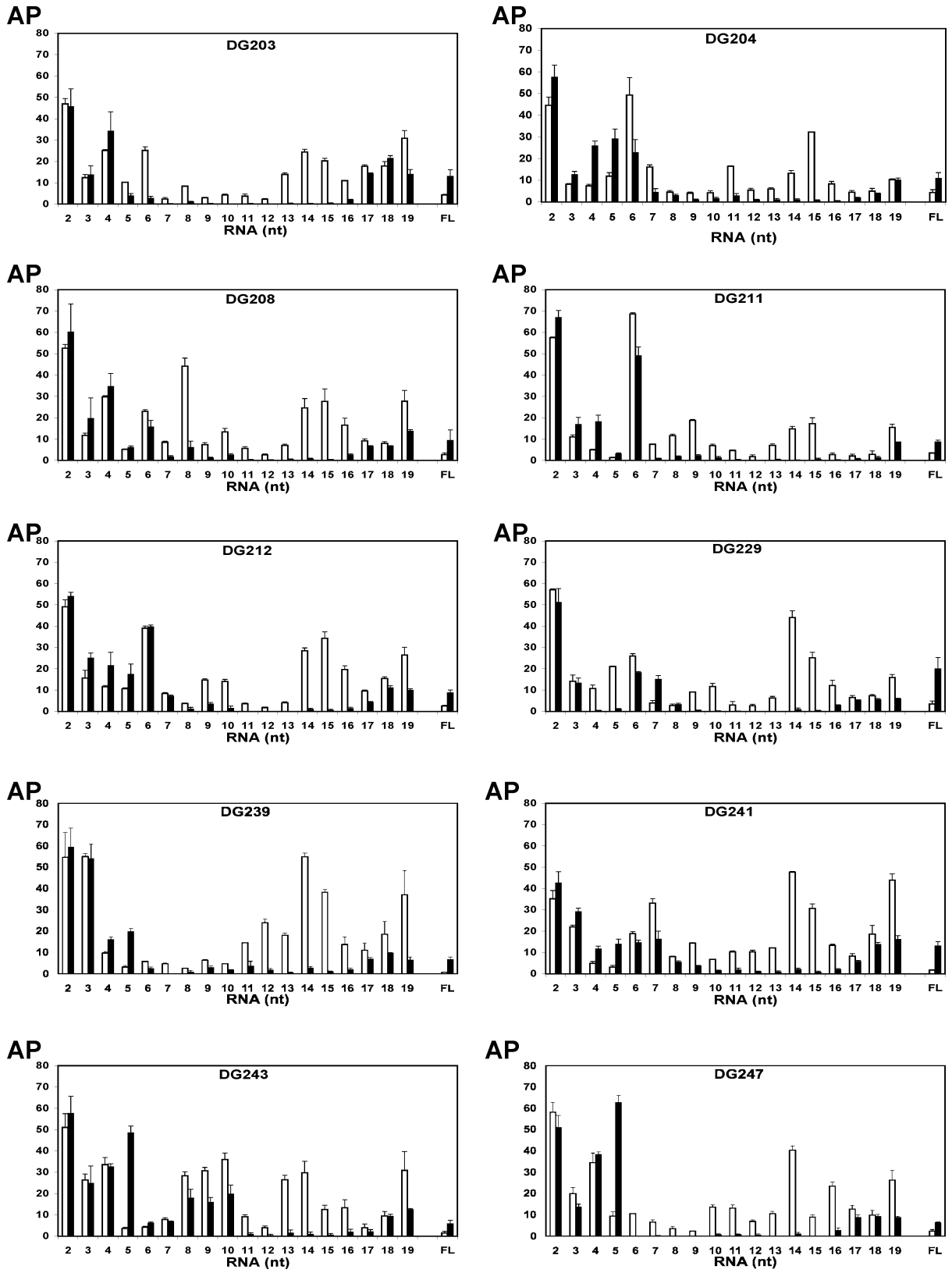


FIGURE 3: Abortive probability profiles for DG200 series promoters in the absence or presence of GreB: clear bars, without GreB; solid bars, with GreB. Abortive probability is calculated for each position as the yield (%) of an abortive RNA corrected by the fraction of RNA polymerase molecules reaching that position (*53*). The higher the abortive probability, the greater the barrier to escape, and the more likely the nascent RNA will be released at that position.

exit channel (*26*). When the linker N-terminus is displaced by the nascent RNA, its movement is propagated to probe the $\sigma 2/-10$ contacts. Similarly, a 13−15 nt long nascent RNA needs to disrupt the tight $\sigma 4/-35$ contacts before escape can occur (*3*, *25*). We reasoned that the fourth abortive block at positions +18 to +19, peculiar to the DG200 subset of ITS mutants, could potentially be due to strong interactions between α-CTD and the UP element upstream of the −35 hexamer. Notwithstanding that the same UP element is present in N25 and all the random-ITS

promoters, it may only create an additional hurdle for the polymerase when combined with particular ITSs—those present in the DG200 promoters. Later, we investigate this possibility by probing the effects of distinct promoter element−RNA polymerase interactions on VLAT production.

VLATs Are Refractive to GreB-Mediated Cleavage and Rescue. The Gre factors GreA and GreB facilitate the rescue of RNAs in backtracked transcription complexes. Previous analysis of the random-ITS promoters showed that abortive transcripts ≤15 nt long are products of backtracked ITCs since they were rescued by 3′-cleavage and reextension in the presence of GreB (*17*). However, there is a minimum requirement of 4−5 bp of RNA−DNA heteroduplex in the backtracked RNA (*17*). Thus, very short abortive transcripts (≤5 nt) are not rescued in GreB-supplemented reactions because they are only weakly held in a short DNA−RNA heteroduplex and rapidly released before GreB can act on them (*17*). On the other hand, the medium and long transcripts (6−15 nt), when backtracked, are stably held in an RNA−DNA hybrid of sufficient length (>4−5 bp) that allows for GreB-mediated cleavage and rescue. A similar analysis with the DG200 promoters also showed a dramatic reduction (or complete absence) of medium and long abortive RNAs in the presence of GreB, while RNAs ≤5 nt long were unaffected (lanes indicated by +GreB in Figure 2). GreB-stimulated promoter escape from positions +6 to +15 led to an increase in the synthesis of full-length product (Figure 3 and Table 2). Interestingly, like the short abortive RNAs ≤5 nt long, the 16−19 nt long VLATs were also undiminished in the presence of GreB; if anything, the level of these RNAs appeared to increase (compare −GreB and +GreB lanes in Figure 2). That the VLATs are refractive to GreB cannot be explained using the same reasoning applied to the short abortive transcripts since the DNA−RNA hybrids formed in the case of the VLATs would be considerably more stable. (For example, even if the 16 nt RNA backtracked by 10 positions, it would still be held in a stable 6 bp heteroduplex). Instead, the VLATs must arise from a mechanism that is distinct from backtracking and are thus impervious to GreB-mediated cleavage and rescue.

VLATs Are Abortively Released by ITCs. Seeing that the default position for promoter escape at the vast majority of the N25 random-ITS mutant promoters is +15/+16, we questioned whether transcripts >15 nt long produced on the DG200 promoters were truly products of abortive initiation or merely associated with paused elongating complexes. At lengths of >15 nt, the nascent RNA should have emerged from RNA polymerase and displaced polymerase−promoter contacts, and the enzyme should have undergone promoter escape and entered the elongation phase of transcription. Paused RNAs similar in length to the VLATs have been detected on the λPR′ and *lac*UV5 promoters, where RNA polymerase undergoes a promoter-proximal pause early in elongation after the synthesis of 16−18 nt long transcripts (*27*, *28*). Unlike aborted RNAs, paused RNAs diminish with time, can be chased with a high concentration (∼1 mM) of NTP, and remain associated with the enzyme (*28*−*31*). We took advantage of these differences in properties to distinguish whether the VLATs were aborted or paused RNAs.

Steady-state transcription assays performed with a representative VLAT-producing promoter, DG203, showed that the 16−19 nt transcripts accumulated continuously over a 60 min period (data not shown). This was in contrast to the paused RNAs on the λPR′ and *lac*UV5 promoters, which were present early in transcription but undetectable after 2−4 min (*28*). Furthermore, the levels of all transcripts (2−19 nt) were unaffected by NTP chase on DG203, indicating that these transcripts did not arise as a result of substrate limitation during elongation (compare lane b to lane c and lane d to lane e in Figure 4A). The NTP chase was also performed in reactions supplemented with either GreA or GreB, which were shown to rescue paused RNAs at the λPR′ and *lac*UV5 promoters (*28*, *32*). As noted earlier, supplementation with GreB had no effect on the 16−19 nt RNAs (compare lane a to lanes d and e in Figure 4A), and a similar observation was made with GreA (compare lane a to lanes b and c in Figure 4A). These results reaffirm that the 16−19 nt RNAs are not products of backtracking. In addition to the Gre factors, we also assessed the effect of NusA on VLAT production. NusA enhances an early pause in elongation at the λPR′ promoter (*29*, *33*). Thus, if the VLATs were associated with paused elongating RNA polymerase, one would expect their levels to increase in the presence of NusA. This, however, was not the case, and NusA had no effect on the VLATs even when present at a 10-fold molar excess over RNA polymerase (data not shown).

To further confirm that RNAs >15 nt long are indeed abortive transcripts, we showed by solid-state transcription on nickel agarose beads that 2−19 nt long RNAs produced on the DG203 promoter are all released into the supernatant (lane S in Figure 4B). GreB supplementation to immobilized RNA polymerase gave the same results as the soluble enzyme (Figure 4B, compare lanes B and R). The NTP chase and solid-state transcription assays conducted on a different VLAT-producing template, DG241, produced similar results (data not shown) and conclusively demonstrate that RNAs 16−19 nt long are not associated with paused elongating complexes, but are abortively released by RNA polymerase during initiation.

VLAT Formation Depends on Strong Polymerase−*Promoter Contacts*. During initiation, RNA polymerase makes specific contacts with various promoter elements including the UP element, −35 and −10 hexamers, the spacer connecting them, the DIS, and the ITS. These contacts have to be displaced for the enzyme to move away from the promoter and commence elongation. Previous work that systematically analyzed the influence of individual promoter elements on abortive initiation and escape at select phage promoters showed that the most significant impediment to promoter escape was the strength of the σ^{70} interactions with promoter DNA (*3*). More specifically, σ2/−10 element interactions present a barrier at positions +6 to +8, while σ4/−35 element contacts present another barrier at positions +13 to +15 (*3*, *24*). Sequential disruption of these contacts by the growing nascent RNA, starting with σ2/−10 contacts and extending upstream, dislodges the σ factor from the DNA (*25*), so that by the time the nascent RNA is 14−15 nt long, the σ subunit should have been completely dislodged from the DNA and the RNA should have emerged from the interior of the enzyme (*18*). If this holds true at all promoters, what accounts for the escape barrier at positions +18 to +19

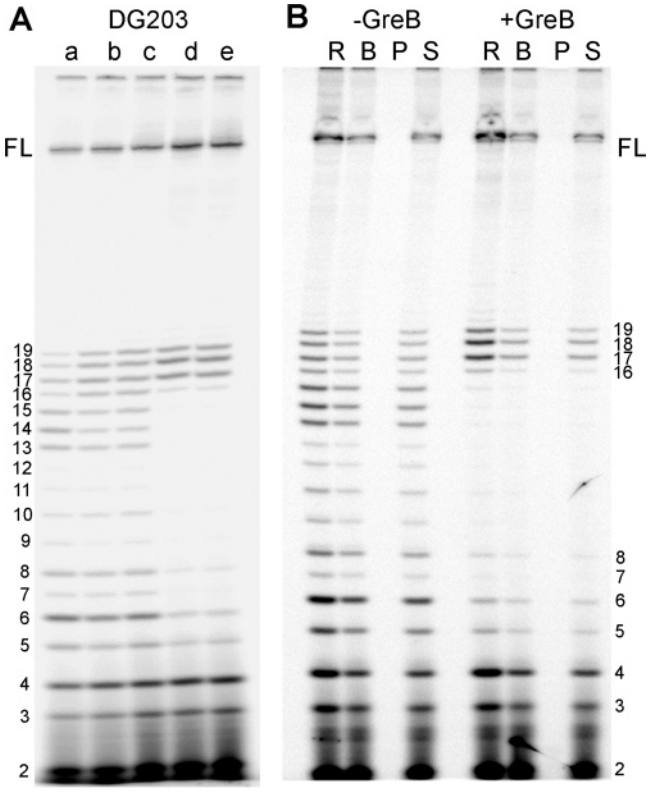


FIGURE 4: Characterization of the 16−19 nt transcripts. The formation of the 16−19 nt RNAs was examined on the DG203 promoter, which differs from N25 only at the +3 to +10 positions of the ITS as shown in Figure 1. (A) High NTP chase to distinguish between released versus paused RNAs. The DG203 promoter was transcribed for 10 min at 37 °C by RNA polymerase alone (lane a) and RNA polymerase supplemented with a 10-fold molar excess of GreA (lanes b and c) or GreB (lanes d and e) followed by a 10 min chase with 1 mM NTP (lanes c and e). (B) Solid-state transcription shows the 16−19 nt RNAs are released from the enzyme complex. DG203 was transcribed with [γ-^{32}P]ATP label by RNA polymerase alone (left panel) or RNA polymerase supplemented with a 10-fold molar excess of GreB (right panel), either in solution (designated R) or immobilized on nickel agarose beads (designated B) for 30 min at 37 °C. Subsequently, an equivalent volume of the bead-based reaction was centrifuged and washed to yield the pellet (P) and supernatant (S) fractions.

on the DG200 promoters? We considered the possibility that strong interactions between the α subunit and the UP element may introduce a further impediment to promoter escape, thereby lengthening the abortive initiation program at the DG200 promoters. The DG200 promoters contain a fairly strong proximal half of the UP element as defined by Estrem et al. (*34*). The same UP element, also present in N25, does not noticeably influence promoter escape at this promoter, but when combined with the ITSs of the DG200 promoters may interact in a different manner with RNA polymerase to delay escape and lengthen abortive initiation.

There are several reports of UP element sequences inhibiting promoter escape at consensus promoters (*35*, *36*). To investigate whether a similar situation could account for the formation of VLATs on the DG200 promoters, we made three changes to the UP element on the DG203 promoter as shown in Figure 5A. In addition to the UP element, we also made changes to other promoter elements, including the −35 and −10 hexamers, the spacer DNA connecting the two, and the DIS region. The resulting promoter fragments all contain the same ITS and downstream transcribed sequence (from +1 to +57)—that of DG203 (Figure 5A). The transcription profile from the various promoters is shown in Figure 5B, the abortive probability profiles calculated for each ITS position are displayed in Figure 5C, and other quantitative initiation parameters are shown in Table 3. Our findings are summarized in the following sections.

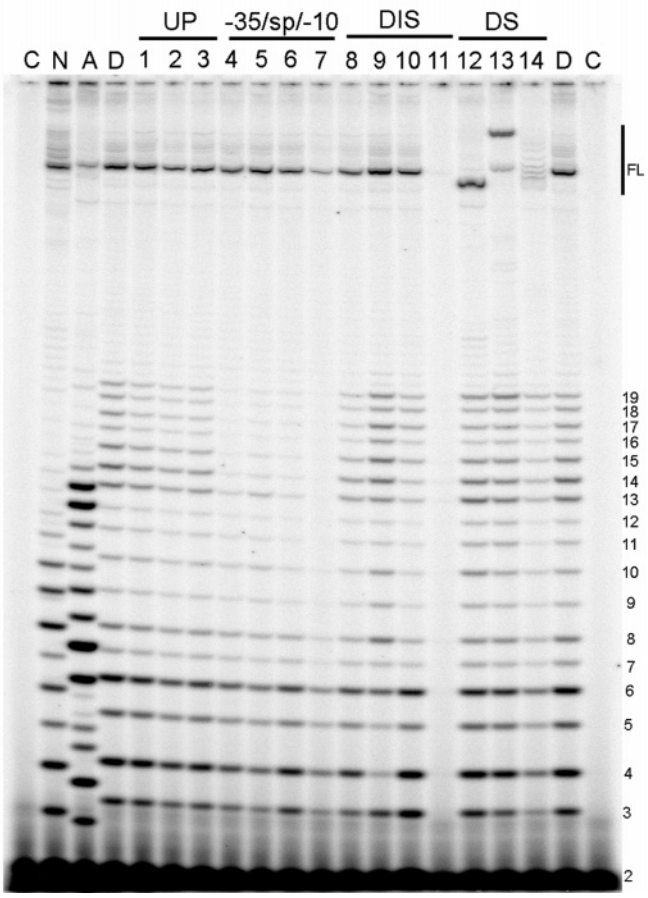
The UP Element Has No Effect on Abortive Initiation and Promoter Escape at DG203. The AT-rich UP element promotes RNA polymerase binding and open complex stability via interactions with the α-CTD (*14*). By virtue of this interaction, the UP element has been shown to inhibit promoter escape and increase abortive initiation at strong promoters—i.e., those with strong polymerase binding signals (*3*, *35*, *36*). This effect was particularly pronounced at late ITS positions +14 and +15 (*3*). We constructed three templates, UPcon, UPwk, and UPdel to test whether UP element−αCTD contacts could account for VLAT production at the DG200 promoters and found that they produced levels of abortive and productive transcription essentially comparable to those at the DG203 promoter (Figure 5B, lanes 1−3, and Figure 5C). For the UPcon construct, changing the UP element to the consensus sequence (*34*) showed a negligible effect. We reasoned that this might be because the UP element on the DG203 promoter is already fairly strong. We were surprised by the results for the UPwk and UPdel constructs, however. Weakening the UP element (UPwk) or completely eliminating the AT-rich region (UPdel) also had no effect on VLAT production (Figure 5B). In fact, for these three constructs, the abortive probability at every ITS position from +2 to +19 was the same as that at the DG203 promoter (Figure 5C). These results argue that UP element−αCTD interactions do not affect promoter escape and abortive initiation on the DG203 promoter.

Mutations in the −35 and −10 Elements Facilitate Promoter Escape at Earlier ITS Positions. The primary determinants of holoenzyme binding to the promoter are the −10 and −35 hexamers, which make specific contacts with σ2 and σ4, respectively. Previously, a T to G substitution at the first base of the consensus −35 or −10 element was sufficient to decrease abortive initiation and increase promoter escape (*3*). We adopted similar changes on the DG203 promoter to investigate their effect on VLAT production. As shown in Figure 5B, the single-base-pair alteration in the −35 or −10 element of the DG203 promoter resulted in a dramatic decrease in the length of the abortive ladder. The longest abortive transcript produced on the −35 and −10 mutant constructs was 14 and 10 nt, respectively (Figure 5B, lanes 4 and 7). The shortened abortive ladder suggests that both mutants achieved promoter escape earlier than on the DG203 promoter. However, the decrease in longer abortive products did not translate into a corresponding increase in full-length product, and both mutants showed productive yields similar to that of the DG203 promoter (Table 3, Figure 5C). Two explanations can account for this observation. First, both the −35wk and −10wk templates initiate at only half the level of DG203 (RIF in Table 3); thus, overall transcription at the mutant promoters is lower than at the DG203 promoter. Second, only a small fraction of RNA polymerase molecules (~7%) abort at positions +11 to +19 on the DG203 promoter. Therefore, even though RNA polymerase is able to undergo escape early at the mutant promoters, the overall effect on the full-length yield is not dramatic.

A

```
                -60              -43          -35                -10
DG203           CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGATTC
                +1                                                          +57
                ATGCGACCGG GAGAGGAGTT TAAATATGGC TGGTCCTCGC AGAAAGCTTC TGCAGGT

                -60              -43          -35                -10
UPcon           CCTCGAGAAA TCATAAAAAA GTAATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGATTC

                -60              -43          -35                -10
UPwk            CCTCGAGAAA TCATAAACAA TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGATTC

                -60              -43          -35                -10
UPdel           CCTCGAGAAA TCATCCTAGG AATATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGATTC

                -60              -43          -35                -10
-35wk           CCTCGAGAAA TCATAAAAAA TTTATGTGCT TTCAGGAAAA TTTTTCTGTA TAATAGATTC

                -60              -43          -35  ΔG            -10
SP16            CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGAAAAT TTTTCTGTAT AATAGATTC

                -60              -43          -35        +A      -10
SP18            CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA ATTTTTCTGT ATAATAGATT C

                -60              -43          -35                -10
-10wk           CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA TTTTTCTGGA TAATAGATTC

                -60              -43          -35                -10
DIS AT          CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAAATTC

                -60              -43          -35                -10
DIS GC          CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGCGGC

                -60              -43          -35                -10     +A
DIS7            CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATAGAATT C

                -60              -43          -35                -10    ΔA
DIS5            CCTCGAGAAA TCATAAAAAA TTTATTTGCT TTCAGGAAAA TTTTTCTGTA TAATGATTC

                +1                                                          +57
DSLac           ATGCGACCGG GAGAGGAGTT TTCACACAGG AAACAGCTAT AGAAAGCTTC TGCAGGT

                +1                                                          +57
DST7A1          ATGCGACCGG GAGAGGAGTT TAGCCATCCC AATCGACACC AGAAAGCTTC TGCAGGT

                +1                                                          +57
DSPR            ATGCGACCGG GAGAGGAGTT GGAACAACGC ATAACCCTGA AGAAAGCTTC TGCAGGT

                +1                                                          +57
N25             ATAAATTTGA GAGAGGAGTT TAAATATGGC TGGTCCTCGC AGAAAGCTTC TGCAGGT

                +1                                                          +57
N25anti         ATCCGGAATC CTCTTCCCGG TAAATATGGC TGGTCCTCGC AGAAAGCTTC TGCAGGT
```

B



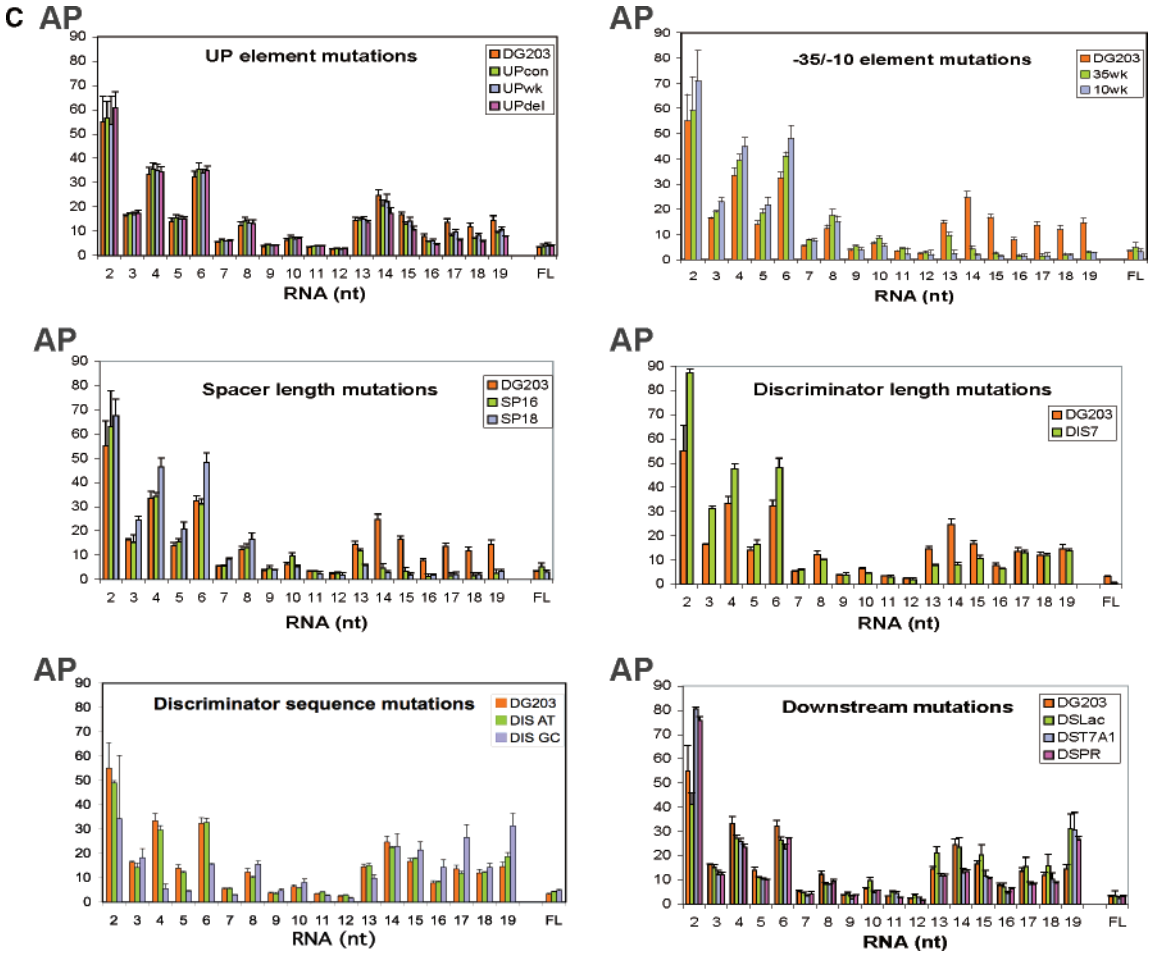


FIGURE 5: Strong interactions between RNA polymerase and promoter DNA promote the production of very long abortive transcripts (16−19 nt). (A) Nucleotide sequence of DG203 and related mutant promoters. The first 12 promoters share identical ITSs and downstream sequences from +1 to +57 as shown for DG203, but differ in upstream promoter sequence as indicated. The last five promoters share identical upstream sequences from −60 to −1 as shown for DG203, but differ in the downstream transcribed sequence as indicated. Bold, underlined sequences denote changes in promoter or downstream transcribed sequences that deviate from the original DG203 template. The UP (centered at −43), −35, and −10 elements are indicated in bold type. (B) Gel profile of in vitro transcribed RNAs from DG203-derived promoters. Each promoter (30 nM) was transcribed for 10 min at 37 °C with 50 nM RNA polymerase in 200 mM KCl and with 100 μM NTP. The transcripts were ^{32}P-labeled at the 5′-triphosphate and fractionated on 7 M urea, 23% (10:1) polyacrylamide gels. The numbers on the right border indicate the size of the abortive RNAs. The promoters are as follows: C, minus-enzyme control reactions; N, N25; A, $N25_{anti}$; D, DG203; 1, UPcon; 2, UPwk; 3, UPdel; 4, −35wk; 5, SP 16; 6, SP 18; 7, −10wk; 8, DIS AT; 9, DIS GC; 10, DIS7; 11, DIS5; 12, DSLac; 13, DST7A1; 14, DSPR. (C) Abortive probability profiles for the DG203-derived promoters.

Changes in Spacer Length Affect the Abortive Probability at Late ITS Positions. At consensus promoters, the −35 and −10 core elements are separated by 17 bp, which allows RNA polymerase to optimally position itself on the promoter (*37*). Variations in spacer length can have dramatic effects on promoter binding and transcription efficiency (*3*, *38*−*40*). We made two variants with spacer lengths of 16 bp (SP16) and 18 bp (SP18) (Figure 5A) and found that both profoundly affected VLAT production (Figure 5B, lanes 5 and 6; Figure 5C). Although both spacer mutants showed an abortive ladder that extended to 19 nt, the levels of 14−19 nt transcripts were significantly reduced (Figure 5B, lanes 5 and 6). In the case of SP16, the overall decrease in abortive initiation was accompanied by a 1.5-fold increase in full-length production and a corresponding decrease in the APR value from 30 to 19 (Table 3). It should be noted that the SP16 spacer variant seemed to initiate with lower efficiency (∼2-fold less) than DG203 with a 17 bp spacer (Table 3). Thus, weakening RNA polymerase−promoter contacts may have facilitated promoter escape, leading to a significant increase in productive synthesis, but at the cost of reducing overall transcription.

The SP18 spacer mutant was different in that, although it too showed facilitated escape at positions +14 to +19, other initiation parameters were very similar to those of DG203, including productive synthesis, relative initiation frequency, and APR (Table 3). More specifically, the decrease in abortive initiation at positions +14 to +19 did not translate into an increase in full-length production. The abortive probability profile of the SP18 mutant promoter (Figure 5C) indicates that although the ITC is stabilized at late ITS positions (+14 to +19) on this template, it is destabilized at early ITS positions (+2 to +6) compared to that of the DG203 promoter. The decrease in abortive synthesis at positions +14 to +19 is compensated by an increase in abortive synthesis at positions +2 to +6, resulting in similar levels of total abortive synthesis, productive synthesis, and APR values (Table 3).

Changes in the Discriminator Length Increase Abortive Initiation at Proximal ITS Positions with No Effect on the

Table 3: Quantitative Parameters for DG203-Derived Promoters

template	PY[a]	APR[a]	MSAT[a] (nt)	RIF[a] (%)
N25	3.1 ± 1.2	34 ± 13	11	136 ± 10
$N25_{anti}$	0.3 ± 0.3	341 ± 252	15	181 ± 23
DG203	3.3 ± 0.3	30 ± 3	19	100 ± 0
UPcon	3.7 ± 1.0	28 ± 8	19	79 ± 5
UPwk	3.9 ± 1.3	27 ± 11	19	84 ± 3
UPdel	3.9 ± 0.3	25 ± 2	19	81 ± 7
−35wk	4.8 ± 2.0	23 ± 9	14	54 ± 20
SP16	5.2 ± 1.5	19 ± 6	19	42 ± 29
SP18	3.0 ± 0.9	35 ± 12	19	96 ± 16
−10wk	3.1 ± 1.1	34 ± 11	10	56 ± 14
DIS AT	4.2 ± 0.2	23 ± 1	19	141 ± 79
DIS GC	4.8 ± 0.4	20 ± 2	19	102 ± 79
DIS7[b]	0.7 ± 0.1	149 ± 26	19	381 ± 57
DIS5	initiates at G(+3), thus not labeled with [γ-^{32}P]ATP			
DSLac	3.5 ± 2.1	35 ± 18	19	189 ± 27
DST7A1	2.5 ± 0.7	41 ± 14	19	63 ± 6
DSPR	3.1 ± 0.5	32 ± 5	19	70 ± 5

[a] The results are the average ± standard deviation from three independent experiments. PY represents the full-length transcript obtained as a percentage of total transcription. APR is the ratio of abortive to productive synthesis. MSAT is the maximum size of the abortive transcript. RIF is the relative initiation frequency as a percentage of that obtained with DG203. [b] DIS7 initiates at both A(+1) and C(−1); the data shown are for A-initiating transcripts.

VLATs. The region between the −10 element and the +1 transcription start site, termed the discriminator or DIS, is melted open in the open complex. The length of the discriminator influences the choice of start sites by *E. coli* RNA polymerase, which shows a preference for a 7 bp discriminator, although 6 and 8 bp discriminators are also commonly seen (*41*, *42*). To determine the effect of the DIS on abortive initiation and promoter escape from the DG203 promoter, we either shortened or lengthened the 6 bp sequence by 1 bp to give constructs DIS5 and DIS7, respectively (Figure 5A).

Using [γ-^{32}P]ATP labeling, we found a complete absence of transcription with DIS5 (Figure 5B, lane 11). With [α-^{32}P]-ATP labeling, however, an abortive ladder extending from 2 to 17 nt was obtained, which we determined to correspond to transcription starting at the G at position +3, eight bases downstream of the −10 element (data not shown). Thus, shortening the DIS to 5 bp necessitated a new transcriptional start site, but promoter escape still occurred at the same position (+19) as on the DG203 promoter (data not shown).

Lengthening the DIS to 7 bp also affected start site selection. The DIS7 variant showed two alternative start sites—one at the normal A at +1, and a second at the C at −1 (data not shown). The A-initiating transcripts from the DIS7 variant were analyzed quantitatively to assess the effect of a lengthened DIS on VLAT production (Figure 5B, lane 10). While VLAT production was unaffected by the 1 bp addition in the DIS (Figure 5C), this mutation appeared to have slightly stabilized ITCs at positions +13 to +15 while simultaneously destabilizing ITCs at positions +2 to +6 (Figure 5C). The increased abortive initiation at the earlier ITS positions resulted in a 5-fold decrease in productive yield with a corresponding increase in the APR value (Table 3).

A GC-Rich Discriminator Increases the Probability of Aborting at VLAT Positions. Because the DIS region is melted in the open complex, the sequence composition of this region can influence the kinetics of open complex formation. Thus, the *rrnB* P1 promoter has a notoriously unstable open complex partially due to a GC-rich DIS which requires higher energy to melt open and keep apart than AT-rich sequences (*9*, *43*). Once an open complex is formed, however, transcription is pulled forward by mass action (i.e., at high NTP concentration), and little abortive initiation is observed at the *rrnB* P1 promoter (*43*, *44*). The DG203 DIS is relatively AT-rich, with four ATs and two GCs (Figure 5A). To assess whether compositional changes in this DIS would alter the abortive initiation pattern at this promoter—specifically as it relates to the VLATs—we made two DIS variants—DISAT, which contains five ATs and one GC, and DISGC, which contains five GCs and one AT (Figure 5A).

Promoter variant DISAT did not show a significant change in abortive initiation profile compared to the DG203 promoter (Figure 5B, lane 8; Figure 5C). This was not surprising, since the DG203 discriminator is very AT-rich to begin with. The productive yield, however, was slightly higher at the DISAT promoter, suggesting that the promoter escape and overall transcriptional efficiency were enhanced at this promoter (Table 3). Recently, Haugen et al. (*9*) demonstrated that σ1.2 directly contacts the DIS region 2 bp downstream from the −10 element and the interaction is strongest when a G occupies this position. Interestingly, this is the case for DG203, which has a G 2 bp downstream of −10 and therefore presumably makes optimal contacts with σ1.2 (Figure 5A). In the DISAT promoter variant, this G was replaced with an A (Figure 5A), which may not associate as strongly with σ1.2, thus allowing for easier escape.

In contrast to the AT-rich DIS, increasing the GC content notably altered the pattern of abortive initiation. Short abortive transcripts 4−6 nt long were significantly decreased, while VLATs were significantly increased (Figure 5B, lane 9; Figure 5C). Nevertheless, the productive yield increased 1.5-fold (presumably due to the decrease in abortive initiation at the early ITS positions), with a corresponding decrease in APR from 30 to 20 (Table 3).

VLAT Production Depends Only on the ITS from +1 to +20 and Not on Sequences Further Downstream. To confirm that the production of VLATs depended solely on the DG203 ITS region encompassing +1 to +20 and not on sequences downstream of +20, we constructed three promoters which retained the upstream promoter region (−60 to −1) and ITS region (+1 to +20) of DG203, but in which the downstream transcribed sequence from +21 to +40 was replaced with the corresponding sequence from *lac*, T7 A1, or λPR as shown in Figure 5A. Although the full-length product for all four promoters was 57 nt, the RNAs migrated differently on the gel due to differences in composition (Figure 5B). As shown in Figure 5B (lanes 12−14) and Figure 5C, these three constructs showed abortive transcription profiles essentially identical to that of DG203, indicating that the sequence downstream of +20 does not have a notable effect on VLAT production. These templates further confirmed that there was no peculiar downstream sequence or structure in the DG200 series mutants that likely caused the premature termination of transcription at 16−19 nt.

Summary of Mutations That Affect VLAT Production. Mutations in the core promoter elements (−35, spacer, −10) that changed the sequence/spacing away from consensus features dramatically reduced the production of VLATs. By

contrast, a GC-rich DIS sequence significantly increased the production of VLATs. Neither the UP element nor the region downstream of +20 had any notable effect on the production of VLATs.

DISCUSSION

Previous work analyzing the role of the ITS in promoter escape led to the discovery of unusually long transcripts (16−19 nt) that were produced on the phage T5 N25 promoter with mutations spanning +3 to +20 of the ITS (*17*). However, ITS changes from +3 to +10 invariably led to synthesis of the 16−19 nt long transcripts. In this paper we investigated the nature of these very long nascent transcripts and demonstrated that, despite their unusual lengths that are reminiscent of paused RNAs produced during elongation, the 16−19 nt RNAs are abortively released by RNA polymerase during initiation. These very long abortive RNAs are also notable for their resistance to the GreB factor, which facilitates cleavage and rescue of abortive RNAs ≤15 nt in length that are products of backtracked RNA polymerases. This suggests that, unlike their shorter counterparts, abortive RNAs ≥16 nt in length are produced via an alternate mechanism that involves RNA polymerase hyper forward translocation during the promoter escape transition. We discuss this mechanism further below.

The formation of the VLATs depends on the two DNA elements with which RNA polymerase interacts during initiation—the PRR and the ITS—which together dictate the efficiency of promoter escape at any promoter. The influence of the PRR on abortive initiation and promoter escape has been clear for some time now, and the general trend is that RNA polymerase has more difficulty clearing promoters with which it forms strong interactions. These interactions include those between the σ factor and the core hexamers and the UP element (where present) and α-CTD. In addition, the σ3.2 linker, which occupies the RNA exit channel in the holoenzyme, acts as a physical barrier to the growing RNA chain (*12*, *13*) and, if not successfully displaced by the RNA, can lead to abortive initiation (*12*, *26*, *45*−*48*). In analyzing the specific promoter−polymerase contacts that influence VLAT production at the DG203 promoter, we found that the most important interactions are among σ, the core hexamers, and the spacer separating them. Weakening these contacts by mutating the −35 and −10 hexamers and the spacer length resulted in a significant reduction of abortive initiation at positions +16 to +19. Thus, as expected, abortive initiation at the VLAT positions also requires the formation of a highly stable open complex. Aside from σ−promoter contacts, we had also anticipated that UP element interactions with α-CTD would play a role in VLAT production, but found this not to be the case at the DG203 promoter. This result was unexpected since it had been previously shown that UP element−α-CTD contacts played a significant role in increasing abortive initiation at distal ITS positions (+14 and +15) on a consensus σ^{70} promoter (*3*). It is possible that the lack of effect seen in the case of DG203 is because σ−promoter contacts are already optimized at this promoter, resulting in a very stable open complex, and UP element−α-CTD interactions contribute negligibly to this stability. In addition to probing polymerase−promoter interactions, we assessed the effect of the σ3.2 linker on VLAT production. Transcription reactions carried out with a S506L linker mutant (*48*), which is presumed to make weaker contacts with the core enzyme (*47*), did not significantly increase or decrease abortive initiation at positions +16 to +19 (data not shown), suggesting that the linker may have already been displaced from the RNA exit channel by transcripts <16 nt in length.

The role that the ITS region plays in VLAT formation has been less straightforward to define. The sequence composition of the ITS is clearly important in demarcating where abortive initiation ceases and promoter escape occurs, although exactly how it does so remains a mystery. The small number of DG200 promoters analyzed precludes the definition of a motif that would clarify which ITS features lead to VLAT production. We do note certain differences between the ITSs of non-VLAT-producing promoters, N25 and the DG100 series, and the VLAT-producing DG200 series, which allows us to suggest a hypothesis. The ITS can be divided into two 10 nt blocks—the first encompassing +1 to +10 and the second +11 to +20. The DG100 promoters, which in general do not produce VLATs, differ from N25 in positions +3 to +20, while the DG200 promoters, which do produce VLATs, differ from N25 in positions +3 to +10. Thus, there must be some information encoded in the second 10 nt block (+11 to +20) that specifies whether VLATs will be produced. Why then does N25, which also contains this 10 nt block from +11 to +20 not produce VLATs? We believe that the key difference lies in the composition of the ITS from +3 to +10, which is very AT-rich in the case of N25, but relatively GC-rich in the case of DG200 promoters (Figure 1). We incorporate these features into the DNA scrunching mechanism described in the introduction to propose a model for VLAT production.

During initial transcription, the 13 bp open complex bubble (spanning −11 to +2) expands as RNA polymerase unwinds downstream DNA to begin RNA synthesis. On the DG200 promoters, where promoter escape occurs at the +19/+20 position, the initial bubble must expand to at least +19, and according to Revyakin et al. (*4*), this would require at least 17 bp of DNA to be compacted within the enzyme, resulting in a highly stressed ITC that is poised to undergo promoter escape. At this point, the upstream edge of the transcription bubble starts to rewind and the energy released dislodges RNA polymerase from the promoter, resulting in promoter escape. The transcription complexes that forward translocate by 1 nt go on to produce full-length RNA. In a small number of cases, however (estimated at ∼7% on DG203), the upstream edge of the initial bubble collapses with such force that RNA polymerase is propelled downstream by several nucleotides. This hyper forward translocation places the growing end of the nascent RNA upstream and out of register with the active site, and the RNA cannot be further elongated; thus, transcription stalls, and the nascent RNA is released via the exit channel. The outlined mechanism is analogous to one proposed for transcription termination at intrinsic terminators which, by forward translocation, would lead to shortening of the RNA−DNA hybrid and subsequent release of the RNA (*49*, *50*). However, whereas forward translocation of RNA polymerase is driven by formation of an RNA hairpin during termination (*50*), we believe that the force of bubble collapse powers forward movement of the enzyme

during promoter escape (and abortive initiation) at the VLAT positions.

The force associated with bubble collapse can conceivably explain why RNA polymerase hyper forward translocates only on the DG200 promoters and not on N25. Earlier we noted that the ITS regions of the DG200 promoters are relatively GC-rich from +3 to +10, while the corresponding region in N25 is very AT-rich. The energy stored in GC base pairs is significantly higher than that stored in AT base pairs. Therefore, when the GC-rich ITS of the DG200 promoters is scrunched inside the ITC, the free energy stored is much higher than in the case of N25. The liberation of this energy upon bubble collapse can in turn propel the polymerase either upstream (i.e., backtracking) or downstream (i.e., hyper forward translocation). Hsu et al. (*17*) have noted that RNA polymerase can reverse translocate up to ~10 nt at the promoter escape junction, and there is no reason to believe that the enzyme cannot forward translocate over ~10 nt as well. Thus, after synthesis of a 16−19 nt long nascent transcript and upon collapse of the upstream edge of the transcription bubble, forward propulsion of the RNA polymerase over ~10 nt would result in the nascent transcript being held only in the RNA exit channel, lacking any DNA/RNA hybrid interactions. Such a loosely held RNA would be readily released. However, to attain RNA release, it is sufficient to hyper forward translocate just ~3−4 bp to generate a weakened hybrid (≤ 5 bp) that spontaneously releases its RNA (*17*, *49*). Importantly also, the hyper forward translocation model can account for the observation that VLATs do not appear to be cleaved and re-extended in the presence of GreB. As detailed above, forward movement of the polymerase along the DNA would result in the relocation of the 3′-end of the RNA upstream, out of the proximity of the enzyme active site. As such, the enzyme would be unable to either continue polymerization or carry out the intrinisic endonucleolytic cleavage, regardless of the presence of GreB. Although we favor this model, we cannot ignore alternative explanations for the lack of GreB-mediated cleavage rescue of VLATs. For instance, GreB-mediated cleavage of 2−18 nt of backtracked RNA has been observed (*51*). Such long backtracked RNA can be accommodated in the secondary channel along with GreB (*6*). Thus, it is theoretically possible that VLATs are also produced as a consequence of RNA polymerase backtracking. However, like the short abortive transcripts that are refractive to GreB cleavage, RNA polymerse complex containing a VLAT must reverse translocate over a large distance (i.e., ~14−15 nt) to leave an RNA−DNA hybrid shorter than 5 bp. As a result of the weak hybrid, the backtracked RNA is quickly released before GreB-stimulated cleavage can occur (*17*). Another source of GreB-refractive VLATs can arise if the nascent RNA is directed onto a path outside of the active site/secondary channel. The 18−20 nt primer RNA (pRNA) associated with the M13 replication origin is such an example where the pRNA occupies the downstream double-stranded DNA binding site (*52*). Unlike VLATs, however, the pRNA remains bound to the template DNA as an extended RNA−DNA hybrid and is not readily released. Experiments are currently under way to test the hyper forward translocation model of abortive release.

ACKNOWLEDGMENT

We thank Dr. J. Burt at the University of California, Berkeley, for his expert advice on the solid-state transcription experiments.

REFERENCES

1. deHaseth, P. L., Zupancic, M. L., and Record, M. T., Jr. (1998) RNA polymerase-promoter interactions: the comings and goings of RNA polymerase, *J. Bacteriol. 180*, 3019−3025.
2. Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R. H. (2000) Structural organization of the RNA polymerase-promoter open complex, *Cell 101*, 601−611.
3. Vo, N. V., Hsu, L. M., Kane, C. M., and Chamberlin, M. J. (2003) *In vitro* studies of transcript initiation by *Escherichia coli* RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape, *Biochemistry 42*, 3798−3811.
4. 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.
5. 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.
6. Opalka, N., Chlenov, M., Chacon, P., Rice, W. J., Wriggers, W., and Darst, S. A. (2003) Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase, *Cell 114*, 335−345.
7. Borukhov, S., Sagitov, V., and Goldfarb, A. (1993) Transcript cleavage factors from *E. coli*, *Cell 72*, 459−466.
8. Fish, R. N., and Kane, C. M. (2002) Promoting elongation with transcript cleavage stimulatory factors, *Biochim. Biophys. Acta 1577*, 287−307.
9. Haugen, S. P., Berkmen, M. B., Ross, W., Gaal, T., Ward, C., and Gourse, R. L. (2006) rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase, *Cell 125*, 1069−1082.
10. Carpousis, A. J., and Gralla, J. D. (1985) Interaction of RNA polymerase with *lac*UV5 promoter DNA during mRNA initiation and elongation. Footprinting, methylation, and rifampicin-sensitivity changes accompanying transcription initiation, *J. Mol. Biol. 183*, 165−177.
11. Zaychikov, E., Denissova, L., and Heumann, H. (1995) Translocation of the *Escherichia coli* transcription complex observed in the register 11 to 20: "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble", *Proc. Natl. Acad. Sci. U.S.A. 92*, 1739−1743.
12. Murakami, K. S., Masuda, S., and Darst, S. A. (2002a) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution, *Science 296*, 1280−1284.
13. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution, *Nature 417*, 712−719.
14. Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase, *Science 262*, 1407−1413.
15. Kammerer, W., Deuschle, U., Gentz, R., and Bujard, H. (1986) Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process, *EMBO J. 5*, 2995−3000.
16. Hsu, L. M., Vo, N. V., Kane, C. M., and Chamberlin, M. J. (2003) *In vitro* studies of transcript initiation by *E. coli* RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters, *Biochemistry 42*, 3777−3786.
17. Hsu, L. M., Cobb, I. M., Ozmore, J. R., Khoo, M., Nahm, G., Xia, L., Bao, Y., and Ahn, C. (2006) Initial transcribed sequence mutations specifically affect promoter escape properties, *Biochemistry 45*, 8841−8854.
18. Komissarova, N., and Kashlev, M. (1998) Functional topography of nascent RNA in elongation intermediates of RNA polymerase, *Proc. Natl. Acad. Sci. U.S.A. 95*, 14699−14704.
19. Hsu, L. M., Vo, N. V., and Chamberlin, M. J. (1995) *Escherichia coli* transcript cleavage factors GreA and GreB stimulate promoter

escape and gene expression *in vivo* and *in vitro*, *Proc. Natl. Acad. Sci. U.S.A. 92*, 11588−11592.
20. Uptain, S. (1997) Structural and functional characterization of *Escherichia coli* RNA polymerase ternary complexes during transcript elongation and termination, Ph.D. thesis, University of California, Berkeley, CA.
21. Chamberlin, M., Nierman, W. C., Wiggs, J., and Meff, N. (1979) A quantitative assay for bacterial RNA polymerases, *J. Biol. Chem. 254*, 10061−10069.
22. Feng, G., Lee, D. N., Wang, D., Chan, C. L., and Landick, R. (1994) GreA-induced transcript cleavage in transcription complexes containing *Escherichia coli* RNA polymerase is controlled by multiple factors, including nascent transcript location and structure, *J. Biol. Chem. 269*, 22282−22294.
23. Kashlev, M., Nudler, E., Severinov, K., Borukhov, S., Komissarova, N., and Goldfarb, A. (1996) Histidine-tagged RNA polymerase of *Escherichia coli* and transcription in solid phase, *Methods Enzymol. 271*, 326−334.
24. Chan, C. L., and Gross, C. A. (2001) The anti-initial transcribed sequence, a portable sequence that impedes promoter escape, requires σ^{70} for function, *J. Biol. Chem. 276*, 38201−38209.
25. Nickels, B. E., Garrity, S. J., Mekler, V., Minahin, L., Severinov, K., Ebright, R. H., and Hochschild, A. (2005) The interaction between σ^{70} and the β-flap of *Escherichia coli* RNA polymerase inhibits extension of nascent RNA during early elongation, *Proc. Natl. Acad. Sci. U.S.A. 102*, 4488−4493.
26. Kulbachinskiy, A., and Mustaev, A. (2006) Region 3.2 of the σ 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.
27. Roberts, J. W., Yarnell, W., Bartlett, E., Guo, J., Marr, M., Ko, D. C., Sun, H., and Roberts, C. W. (1998) Antitermination by bacteriophage λ Q protein, *Cold Spring Harbor Symp. Quant. Biol. 63*, 318−325.
28. Nickels, B. E., Mukhopadhyay, J., Garrity, S. J., Ebright, R. H., and Hochschild, A. (2004) The σ^{70} subunit of RNA polymerase mediates a promoter-proximal pause at the *lac* promoter, *Nat. Struct. Mol. Biol. 11*, 544−550.
29. Grayhack, E. J., Yang, X. J., Lau, L. F., and Roberts, J. W. (1985) Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site, *Cell 114*, 259−269.
30. Levin, J. R., and Chamberlin, M. J. (1987) Mapping and characterization of transcriptional pause sites in the early genetic region of bacteriophage T7, *J. Mol. Biol. 196*, 61−84.
31. Artsimovich, I., and Landick, R. (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals, *Proc. Natl. Acad. Sci. U.S.A. 97*, 7090−7095.
32. 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.
33. Yarnell, W. S., and Roberts, J. W. (1992) The phage λ gene Q transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase, *Cell 69*, 1181−1189.
34. Estrem, S. T., Ross, W., Gaal, T., Chen, S. Z. W., Niu, W., Ebright, R. H., and Gourse, R. L. (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit, *Genes Dev. 13*, 2134−2147.
35. Ellinger, T., Behnke, D., Knaus, R., Bujard, H., and Gralla, J. D. (1994) Context-dependent effects of upstream A-tracts, *J. Mol. Biol. 239*, 466−475.
36. Strainic, M. G., Jr., Sullivan, J. J., Velevis, A., and deHaseth, P. L. (1998) Promoter recognition by *Escherichia coli* RNA polymerase: effects of the UP element on open complex formation and promoter clearance, *Biochemistry 37*, 18074−18080.
37. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002b) Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex, *Science 296*, 1285−1290.
38. Stefano, J. E., and Gralla, J. D. (1982) Spacer mutations in the *lac* promoter, *Proc. Natl. Acad. Sci. U.S.A. 79*, 1069−1072.
39. Mulligan, M. E., Brosius, J., and McClure, W. R. (1985) Characterization *in vitro* of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the TAC promoter, *J. Biol. Chem. 260*, 3529−3538.
40. Warne, S. E., and deHaseth, P. L. (1993) Promoter recognition by *Escherichia coli* RNA polymerase. Effects of single base pair deletions and insertions in the spacer DNA separating the −10 and −35 regions are dependent on spacer DNA sequence, *Biochemistry 32*, 6134−6140.
41. Hawley, D. K., and McClure, W. R. (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences, *Nucleic Acids Res. 11*, 2237−2255.
42. Lisser, S., and Margalit, H. (1994) Determination of common structural features in *Escherichia coli* promoters by computer analysis, *Eur. J. Biochem. 223*, 823−830.
43. Barker, M. M., Gaal, T., and Gourse, R. L. (2001) Mechanisms of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP, *J. Mol. Biol. 305*, 689−702.
44. Gourse, R. L. (1988) Visualization and quantitative analysis of complex formation between *E. coli* RNA polymerase and an rRNA promoter *in vitro*, *Nucleic Acids Res. 16*, 9789−9809.
45. Hernandez, V. J., and Cashel, M. (1995) Changes in conserved region 3 of *Escherichia coli* sigma 70 mediate ppGpp-dependent functions *in vivo*, *J. Mol. Biol. 252*, 536−549.
46. Hernandez, V. J., Hsu, L. M., and Cashel, M. (1996) Conserved region 3 of *Escherichia coli* sigma 70 is implicated in the process of abortive initiation, *J. Biol. Chem. 271*, 18775−18779.
47. Sen, R., Nagai, H., Hernandes, V. J., and Shimamoto, N. (1998) Reduction in abortive transcription from the λPR promoter by mutations in region 3 of the σ^{70} subunit of *Escherichia coli* RNA polymerase, *J. Biol. Chem. 273*, 9872−9877.
48. Cashel, M., Hsu, L. M., and Hernandez, V. J. (2003) Changes in conserved region 3 of *Escherichia coli* σ^{70} reduce abortive transcription and enhance promoter escape, *J. Biol. Chem. 278*, 5539−5547.
49. Komissarova, N., Becker, J., Solter, S., Kireeva, M., and Kashlev, M. (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination, *Mol. Cell 10*, 1151−1162.
50. Santangelo, T. J., and Roberts, J. W. (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator, *Mol. Cell 14*, 117−126.
51. Kulish, D., Lee, J., Lomakin, I., Nowicka, B., Das, A., Darst, S., Normet, K., and Borukhov, S. (2000) The functional role of basic patch, a structural element of *Escherichia coli* transcript cleavage factors GreA and GreB, *J. Biol. Chem. 275*, 12789−12798.
52. Zenkin, N., Naryshkina, T., Kuznedelov, K., and Severinov, K. (2006) The mechanism of DNA replication primer synthesis by RNA polymerase, *Nature 439*, 617−620.
53. Hsu, L. M. (1996) Quantitative parameters for promoter clearance, *Methods Enzymol. 273*, 59−71.

BI701236F