

The Roles of Specific Template Nucleosides in the Formation of Stable Transcription Complexes by *Escherichia coli* RNA Polymerase*

(Received for publication, September 13, 1999, and in revised form, December 7, 1999)

Judith R. Levin^{‡§}, Julie J. Blake[¶], Ruth A. Ganunis^{¶*}, and Thomas D. Tullius^{¶‡§§}

From the [‡]Departments of Biological Sciences and Chemistry, Goucher College, Baltimore, Maryland 21204, the [¶]Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218, and the ^{‡‡}Department of Chemistry, Boston University, Boston, Massachusetts 02215

We have examined the effects of removing individual template nucleosides on promoter escape by *Escherichia coli* RNA polymerase *in vitro*. The ability of DNA templates containing random single nucleoside gaps generated by hydroxyl radical treatment to support the production of stable ternary transcription complexes was analyzed. On two templates containing different promoter and initial transcribed regions, we found that removal of nucleosides on the template strand in the region from –13 to at least +8 relative to the transcription start site interfered with ternary complex formation. The downstream border of this region varied for the two templates, suggesting an effect of the specific nucleotide sequence on the stability of intermediates in the promoter escape process. On the nontemplate strand, removal of nucleosides in the vicinity of the –10 consensus promoter element interfered with escape, whereas removal of nucleosides in the vicinity of the transcription start site actually enhanced the yield of ternary complexes. On one template, removal of nucleosides in an A-tract containing region upstream of the promoter caused a significant decrease in promoter escape, consistent with previous suggestions that contacts between this region and the RNA polymerase play a role in promoter binding and/or initiation.

The initiation of transcription by *Escherichia coli* RNA polymerase involves not only DNA binding, but also subsequent isomerization to a kinetically competent “open” promoter complex, onset of RNA synthesis in an abortive, unstable mode termed an “initial transcribing complex” (1), and eventually the loss of σ factor and transition to a stably elongating mode characterized by salt resistance and a reduced DNA footprint

* This work was supported by Public Health Service Grant GM41930 from the National Institutes of Health. We gratefully acknowledge the use of densitometry instrumentation maintained by the Institute for Biophysical Research on Macromolecular Assemblies at Johns Hopkins University, which was supported by an National Science Foundation Biological Research Centers Award and by a grant from the W. M. Keck Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Goucher College, 1021 Dulaney Valley Rd., Baltimore, MD 21204. E-mail: jlevin@goucher.edu.

¶ Present address: Dept. of Stomatology, Box 0422, Clinical Sciences C740, University of California, 521 Parnassus Ave., San Francisco, CA 94143-0422.

** Present address: FDA, 5600 Fishers Lane, Rockville, MD 20857.

§§ Present address: Dept. of Chemistry, Boston University, 590 Commonwealth Ave., Boston, MA 02215.

(2). Promoters vary greatly in their effectiveness at each of these steps, with any single step having the potential of becoming rate-limiting for transcription initiation (reviewed in Ref. 3). The variation in properties among promoters depends not only on the nucleotide sequence in the region of initial promoter binding, but also on the sequence comprising the initial transcribed region (1, 4–7).

Although the role of template DNA sequence in open complex formation has been extensively characterized through genetic, biochemical, and *in vitro* mutagenic analysis (3), the basis for variation in the steps subsequent to open complex formation is poorly understood. The current view of initiation supposes a series of conformational changes involving movement of flexible domains within the polymerase to accommodate a nascent RNA chain and translocate the active site to consecutive template positions, at first without relinquishing contacts established during open promoter formation. Presumably, the strain accumulated in this process eventually leads to a major transition in which σ factor is expelled, contacts with the promoter recognition elements are relinquished, and the enzyme assumes a conformation characteristic of a stable, processively elongating ternary complex. The properties of the elongating ternary complex have been the subject of recent controversy (reviewed in Refs. 8 and 9) as to whether the enzyme maintains a rigid structure such that ternary complexes at all template positions are analogous or whether continued relative motion of flexible parts within the enzyme leads to a cycling among multiple possible structures, resulting in an “inchworm-like” movement along the template. Regardless, the importance of enzyme flexibility in the process of initiation seems clear, and it seems likely that differences in interactions of specific template nucleotide sequences with the polymerase during the series of conformational changes involved in promoter escape must account for the promoter-specific variation in initiation properties.

A systematic analysis of the effects of specific template mutations on the process of promoter escape would be prohibitive in its scope and, barring the use of heteroduplex templates, would not distinguish between the roles of the template and nontemplate strands. As an initial approach to assessment of the role of nucleotide sequence in a biochemical process, it is desirable to be able to screen a large number of nucleotide changes, on each of the DNA strands, simultaneously. Chemical techniques for generating populations of such modified DNA molecules, such as methylation (10), ethylation (10), depurination (11), and cleavage with hydroxyl radical (12) have been used to define protein-DNA contacts necessary for specific binding events. In this case, we wished to apply such a technique to study the transition from an open promoter complex to a stable ternary transcription complex.

The missing nucleoside technique (12) uses native gel electrophoresis to discriminate among a population of DNA molecules, which have been randomly gapped by treatment with hydroxyl radical, according to their ability to form a specific protein-DNA complex. Native gel electrophoresis can resolve ternary complexes of RNAP, DNA, and RNA from binary complexes containing only enzyme and DNA (1, 13, 14), and ternary complexes themselves can exhibit varied mobilities based on a number of contributing factors that may include polymerase subunit composition or conformation, degree and position of DNA bending, and length of RNA product. We decided that these properties presented a rich opportunity to dissect the roles of various template nucleosides in the events leading from promoter binding to the formation of a productively elongating transcription complex.

EXPERIMENTAL PROCEDURES

RNA Polymerase—*E. coli* RNA polymerase was generously provided by Dr. Michael Chamberlin (University of California, Berkeley, CA). The C terminus of the β subunit of polymerase is histidine-tagged in the preparation of the enzyme (15).

Preparation of DNA Templates—Plasmids pAR1707 (16) and pT-T7 (17) were generously provided by Dr. Michael Chamberlin (University of California, Berkeley, CA). Fragments containing the promoters were prepared by PCR¹ as described previously (1). Template 1, a 242-base pair fragment spanning the region -149 to +93 of the T7 A1 promoter, was generated from pAR1707. Template 3, a 245 base pair fragment spanning the region -158 to +87 of the tac promoter, was generated from plasmid pT-T7. The template or nontemplate strand primers for template 3 or the template strand primer for template 1 were 5' radiolabeled prior to PCR using [γ -³²P]ATP and polynucleotide kinase (18). Following PCR and digestion with *Ban*I, the nontemplate strand of template 1 was 3' radiolabeled using [α -³²P]dGTP and DNA polymerase I Klenow fragment (18). The specific activity of the radiolabeled DNA was approximately 5×10^5 cpm/pmol. PCR products were purified on 6% native polyacrylamide gels (29:1 acrylamide:bisacrylamide). The DNA was recovered by the crush and soak procedure (18).

Templates used in the "singly gapped template" experiment (see Figs. 7 and 8) have the sequence of the template containing the T7 A1 promoter (template 1) but with a gap on the template strand at position +10 or +13. The templates were generated by annealing three oligonucleotides. The following oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA): 1) NT6040, the nontemplate strand from -60 to +40; 2) T6040, the template strand from -60 to +40; 3) T6012, the template strand from -60 to +12; 4) T1440, the template strand from +14 to +40; 5) T6009, the template strand from -60 to +9; and 6) T1140, the template strand from +11 to +40. The 5' ends of the oligonucleotides T6012 and T6009 and the 3' ends of the oligonucleotides T1440 and T1140 were phosphorylated so that both ends of the gaps would be terminated by phosphates.

For each annealing reaction 2 pmol (5×10^6 cpm) of 5' radiolabeled NT6040 DNA was used in 50 μ l of a buffer consisting of 10 mM Tris (pH 8.0), 10 mM MgCl₂, and 1 mM EDTA. To obtain the intact template, 4 pmol of oligonucleotide T6040 were added to 2 pmol of NT6040. Template with gap positioned at +13 (G13) was produced by annealing 2 pmol of NT6040, 4 pmol of T6012, and 8 pmol of T1440. The template with a gap positioned at +10 (G10) was produced by annealing 2 pmol of NT6040, 4 pmol of T6009, and 8 pmol of T1140. Reaction mixtures were denatured at 90 °C for 5 min and then cooled to room temperature over 3 h. The gapped templates were purified on a 12% native polyacrylamide gel.

Hydroxyl Radical Treatment—1.6 pmol of radiolabeled DNA or unlabeled DNA were hydroxyl radical treated as described previously (12), using the following final concentrations of reagents: 50 μ M ferrous ammonium sulfate, 100 μ M EDTA, 1 mM sodium ascorbate, and 0.03% H₂O₂.

Transcription Reactions—0.32 pmol of DNA (PCR product or singly gapped oligonucleotide construct) was dissolved in 30 μ l of transcription buffer consisting of 44 mM Tris (pH 8.0), 14 mM MgCl₂, 14 mM 2-mercaptoethanol, 20 mM NaCl, 0.5% glycerol, and 40 μ g/ml acetylated bovine serum albumin. On ice, 1.1 pmol of RNA polymerase was added

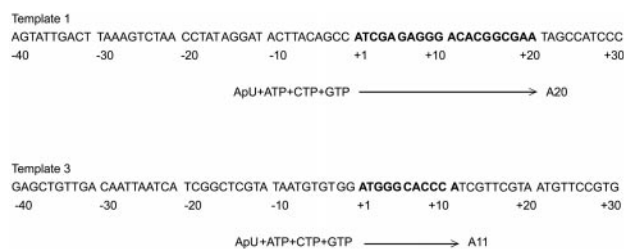


FIG. 1. Sequences of the nontemplate strands of the promoter regions of templates 1 and 3. The initial transcribed region under limited nucleotide conditions is shown in *bold type*.

to each reaction, followed by incubation in a water bath at 30 °C for 10 min. 10 μ l of this prebinding reaction was removed and stored at 30 °C for an additional 10 min, and then 3.5 μ l of sucrose loading buffer (60% sucrose, 0.01% bromphenol blue, and 0.01% xylene cyanol) was added. To the remaining reaction mixture nucleotides were added to final concentrations of 44 μ M ApU and 1.25 μ M each of ATP, GTP, and CTP for template 1, or 115 μ M ApU and 6.5 μ M each of ATP, GTP, and CTP for template 3. For reactions in which the RNA was to be labeled, [α -³²P]CTP was included at a specific activity of 10^5 cpm/pmol. For the RNA sequencing reactions shown in Fig. 8, 3'-deoxynucleoside triphosphates were included, as described by Arndt and Chamberlin (19). Following an additional 10-min incubation at 30 °C, the transcription mixture was divided into three samples: 1) 10 μ l was removed and added to 10 μ l of formamide loading buffer (80% (v/v) deionized formamide, 50 mM Tris, 50 mM boric acid, 1 mM EDTA, 0.1% (w/v) xylene cyanol, 0.1% (w/v) bromphenol blue); 2) 10 μ l of the reaction mixture was chased by the addition of 65 μ g/ml heparin and 0.9 mM each of ATP, GTP, UTP, and CTP. After 10 min at room temperature, 10 μ l of formamide loading buffer was added; and 3) to the remaining mixture, 4.5 μ l of sucrose loading buffer was added. For RNA size markers (see Fig. 8), radiolabeled RNA was generated in the presence of 3'-deoxynucleoside triphosphate chain terminators (19) or was subjected to hydrolysis using alkaline sodium phosphate (20).

Native Gel Electrophoresis—Samples in sucrose loading buffer were analyzed by gel electrophoresis on 4% native polyacrylamide gels (37.5:1 acrylamide:bisacrylamide, 0.8 mm thick, 14 cm long) run at 19 °C (template 1 and singly gapped oligonucleotide template samples) or 4 °C (template 3). After autoradiography, the complexes were recovered from the native gel by the crush and soak procedure (18) and subjected to denaturing gel electrophoresis.

Denaturing Gel Electrophoresis of DNA and RNA—Labeled DNA and/or RNA purified from the native gel or reserved from the reaction mixture in formamide loading buffer was electrophoresed on denaturing polyacrylamide gels (19:1 acrylamide:bisacrylamide, 8 M urea). The total polyacrylamide concentration was 8% for the PCR template samples and 20% for the singly gapped oligonucleotide template samples. The gels were dried and phosphorimaged using a Molecular Dynamics PhosphorImager.

Data Treatment—The program GelExplorer (21) was applied to the denaturing polyacrylamide gel image to obtain the intensities of the bands at each nucleotide position in the recovered samples. At each nucleotide position, the integral of a band from the sample of the DNA remaining in the open promoter complex after the addition of nucleotides (EP₀nuc) was subtracted from the integral of the corresponding band in the sample of DNA recovered from the ternary complex. To normalize to the starting open promoter population, this difference was divided by the band integral for the initial open promoter complex (EP₀) at each nucleotide position. The values were then divided by 2 so that the range of the data was 1. To correct for the partitioning behavior of intact DNA, a value representing the average partitioning for several bands near the bottom of the gel (typically around position +30), assumed to have gaps at positions that do not affect promoter escape, was subtracted from all data points. Therefore, the values for positions with no effect on the partitioning fluctuate around zero. A negative intensity indicates that a gap at that nucleotide position results in the polymerase remaining at the promoter after the addition of nucleotides more often than for intact DNA. A positive intensity indicates that a gap at that nucleotide position results in the polymerase leaving the promoter more often than for intact DNA.

RESULTS

For our experiments, we chose two templates (Fig. 1) for which the biochemical properties of ternary complexes with

¹ The abbreviations used are: PCR, polymerase chain reaction; EP₀, open promoter complex; EP₀nuc, EP₀ after the addition of nucleotides; ApU, adenylyl (3'→5')uridine.

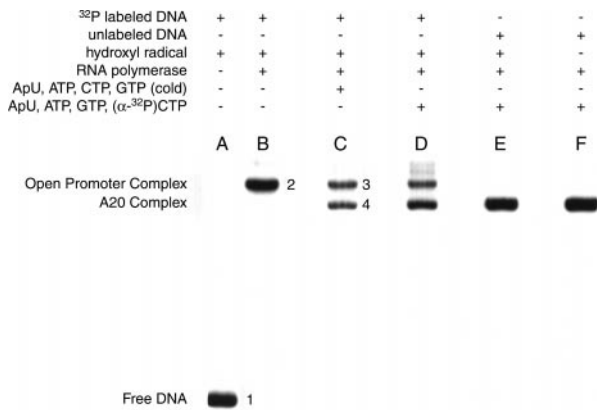


FIG. 2. Separation of template 1 transcription complexes by nondenaturing gel electrophoresis. Where indicated, the DNA was 5' end-labeled on the template strand. *Lane A*, free hydroxyl radical-treated DNA; *lane B*, hydroxyl radical-treated DNA bound by RNA polymerase; *lane C*, hydroxyl radical-treated DNA with RNA polymerase bound and ribonucleotides added; *lane D*, same as *lane C* except RNA is radiolabeled; *lane E*, same as *lane D* except the DNA is not radiolabeled; *lane F*, same as *lane E* but with control DNA (untreated with hydroxyl radical). Specific bands named in text are assigned numbers: *band 1*, free DNA; *band 2*, DNA bound by RNA polymerase (EP_o); *band 3*, DNA bound to RNA polymerase, with ribonucleotides added (EP_onuc); *band 4*, DNA, RNA polymerase, and transcript in a ternary complex (A20).

RNA polymerase have been well studied (1, 14). Template 1 contains the bacteriophage T7 A1 promoter recognition region and initial transcribed region and can support the production of a stable ternary complex containing RNA polymerase and a 20-nucleotide RNA (termed A20) when provided with a limited set of nucleotide substrates (ApU + ATP, GTP, and CTP; Ref. 22). Template 3 contains the *E. coli tac* promoter recognition region and an initial transcribed region that supports the production of a stable ternary complex containing an 11-nucleotide RNA (termed A11) under the same conditions (1). On both of these templates, the transition from an abortively transcribing complex to a stable ternary complex lacking σ factor is known to occur by the time a 10-nucleotide RNA has been synthesized (1, 23).

The ability to resolve binary open promoter complexes (EP_o) from stable ternary complexes and from free DNA using native gel electrophoresis is retained when the DNA template has been gapped by treatment with hydroxyl radical (Fig. 2). Consistent with previous reports (1, 13, 14, 24), the open promoter complex (Fig. 2, *lane B*) has greatly reduced mobility compared with the free DNA (Fig. 2, *lane A*). When EP_o complexes formed on hydroxyl radical treated template 1 are presented with nucleotide substrates appropriate for synthesis of A20 RNA, a significant proportion of the resulting complexes are seen to increase slightly in mobility (Fig. 2, *lane C*). Radiolabeling of the RNA in the complexes rather than the DNA (Fig. 2, *lane E*) reveals that only the new, increased mobility complex contains an RNA component and is therefore identified as a ternary complex. Analysis of RNA extracted from this band confirmed that it had the expected size of 20 nucleotides (with some contamination by a 26-nucleotide readthrough product; see Ref. 22) (data not shown). Control experiments confirmed that the A20 complexes assembled on hydroxyl radical treated DNA are capable of resuming elongation when presented with a "chase" of all four nucleotide substrates (data not shown). An analogous experiment was carried out with template 3 (data not shown).

The formation of ternary complexes and the retention of some complexes with a mobility of EP_o (Fig. 2, *lane C*) can be viewed as a partitioning of the starting population of EP_o

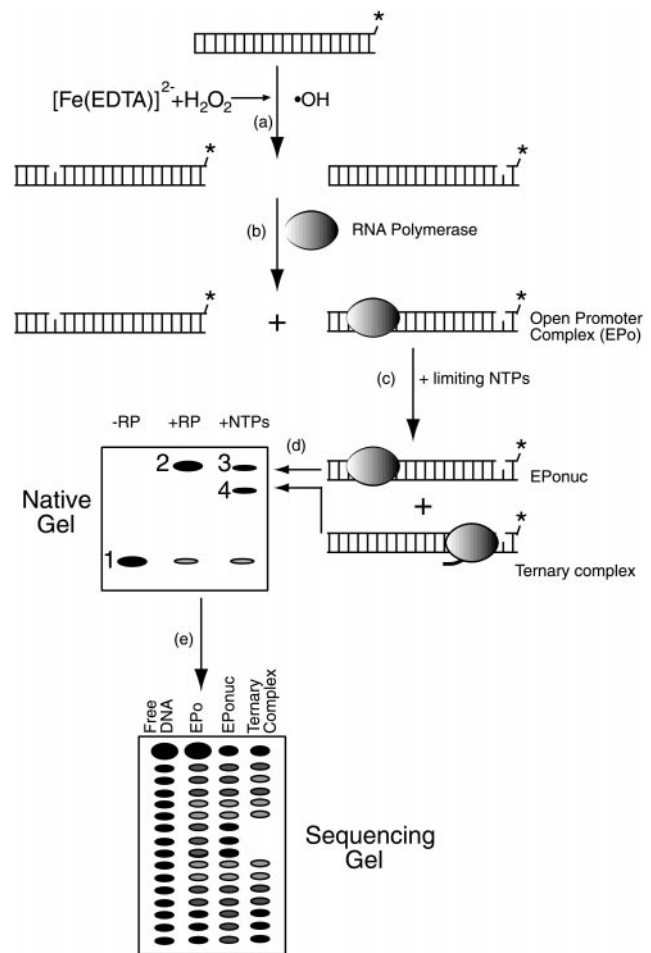


FIG. 3. Design of the missing nucleoside experiment applied to the study of promoter escape. *a*, radiolabeled promoter-containing template DNA is treated with hydroxyl radical to generate a pool of singly gapped DNA molecules. *b*, the DNA sample is then incubated with RNA polymerase. Although most of the gapped templates are still able to bind to RNA polymerase to form an open promoter complex (EP_o, right), some gaps interfere with polymerase binding (left). *c*, to a portion of the resulting mixture, a limited set of nucleotide substrates is added to allow transcription to initiate and proceed to form a ternary complex stalled for lack of the missing substrate (see Fig. 1). *d*, native gel electrophoresis is used to separate unbound DNA, ternary complexes, and open promoter complexes. EP_onuc designates the population of template molecules that retain the mobility of an open promoter complex even after exposure to NTP substrates. *e*, after elution of complexes from the native gel, denaturing electrophoresis is used to separate individual gapped DNA molecules from each other. Phosphorimaging is used to quantitate the partitioning of each gapped template from the starting EP_o population between the product EP_onuc and ternary complex populations.

complexes when presented with substrates for transcription. If gaps at different positions in the DNA affect this partitioning differently, then the representation of different gapped templates should vary between the two product complex populations. This concept formed the basis of our experimental design, which is depicted schematically in Fig. 3. After incubation of RNA polymerase with radiolabeled, hydroxyl radical treated DNA to form the starting EP_o population, nucleotide substrates appropriate for formation of a poised ternary complex are added. After sufficient time for transcription to occur, the resulting complexes are separated by native gel electrophoresis and eluted from the gel matrix. The DNA in these complexes is then analyzed by denaturing gel electrophoresis, where each band represents a template molecule gapped at a unique position on the radiolabeled strand. For each band, the quantity of that band found in the ternary complex population compared

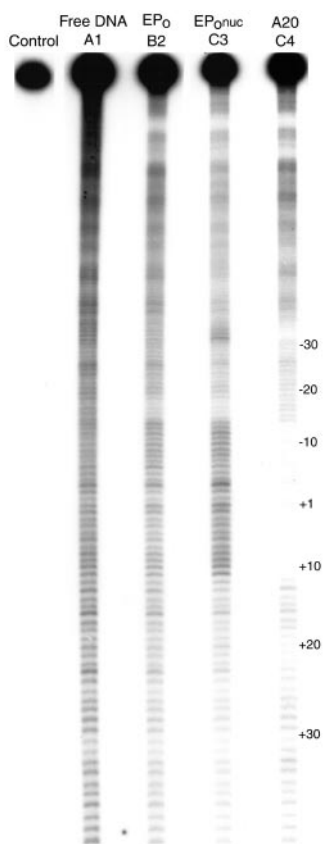


FIG. 4. Effect of template gaps on formation of a stable ternary complex. Shown is an autoradiogram of the denaturing polyacrylamide gel for the template strand of template 1. Lane numbers correspond to numbering of the bands on the native gel from which the DNA was recovered (Fig. 2). Control, radiolabeled DNA not hydroxyl radical-treated. Lane A1, hydroxyl radical-treated DNA; lane B2, EP_o complex; lane C3, EP_o,nuc complex; lane C4, A20 complex. Nucleotide positions relative to the start site of transcription are indicated on the right.

with the quantity of that band in the starting EP_o population reveals the ability of the corresponding gapped template to support production of a stable ternary complex.

It is important to note that under the experimental conditions used, the transition from open complexes is not complete, even with intact template molecules. This is due to a number of factors, including the slow kinetics of initiation at low nucleotide concentrations, and the presence of enzyme molecules that are catalytically inactive but competent in binding to the promoter (22). Therefore, to evaluate the importance of a particular template nucleoside in promoter escape, it is necessary to compare the fraction of open complexes that is able to make the transition to ternary complexes in the presence and absence of that nucleoside.

Effects of Missing Nucleosides on the Observed Partitioning between Open Promoter Complexes and Stable Ternary Complexes—The results of a typical experiment are shown in Fig. 4, which depicts the raw data for the template strand of template 1. Although the population of template molecules found in EP_o complexes (Fig. 4, lane B2) is partially depleted for certain positions where gaps adversely affect open promoter formation, there is no one template position where existence of a gap completely abolishes EP_o formation. Indeed, under the conditions employed, the vast majority of template molecules is bound by RNA polymerase (Fig. 2, lane B). Thus, every possible gapped template molecule is represented at least to some degree in the EP_o complex population that participates in the formation of a ternary complex. The amount of DNA representing each template gap position was calculated by peak fitting and integration of the

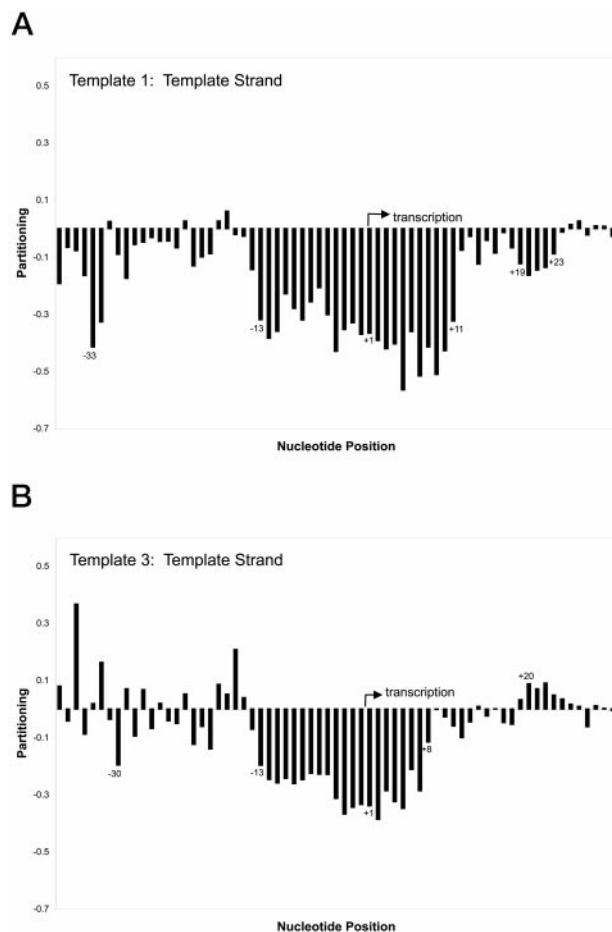


FIG. 5. The effect of a missing nucleoside in the template strand on escape from the promoter. Data for template 1 (A) and template 3 (B) are shown. Data were treated as described under “Experimental Procedures.”

phosphorimaged gel data using the GelExplorer program (21). For each experiment, the integrals for the two product populations (ternary complex and EP_o,nuc , the population of complexes that retains the mobility of the original EP_o complex) sum nearly exactly to the integral of the starting EP_o population (data not shown), indicating that it is valid to treat this process as a partitioning event between two possible outcomes.

The partitioning of each gapped template between the ternary complex population and the EP_o,nuc population was compared with the partitioning observed for full-length (*i.e.* un-gapped) template (Figs. 5 and 6, respectively, for gaps on the template and nontemplate strands). A negative value indicates a template position where a gap interferes with the transition of an EP_o complex to a stable ternary complex, whereas a positive value indicates a template position where a gap appears to enhance this process.

On the template strand (Fig. 5), there were few, if any, positions for either template where nucleoside removal enhanced formation of a stable ternary complex. Rather, an extensive region flanking the start site of transcription (-13 to $+11$ on template 1, and -13 to $+8$ on template 3) was observed in which missing nucleosides caused the complex to remain at the EP_o,nuc position. For each template, a large part of this region corresponds to positions where missing nucleosides have been observed to enhance the formation of the open promoter complex (24, 25).² In both cases, the inhibition is the strongest

² J. R. Levin, J. J. Blake, R. A. Ganunis, and T. D. Tullius, unpublished observations.

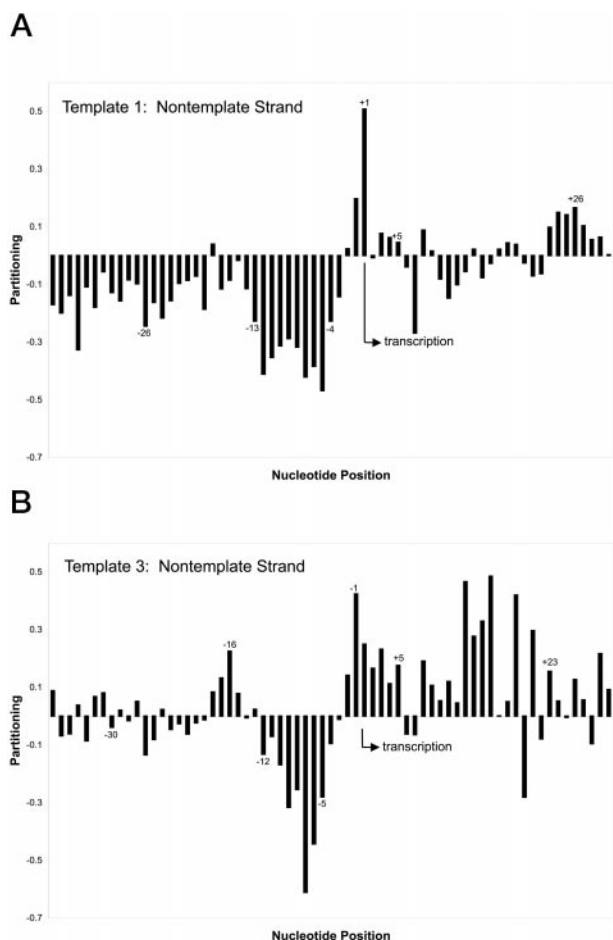


FIG. 6. The effect of a missing nucleoside in the nontemplate strand on escape from the promoter. Data for template 1 (A) and template 3 (B) are shown. Data were treated as described under "Experimental Procedures."

for gaps downstream and immediately upstream of the transcription start site. For gaps in this region (-4 to $+11$ on template 1 and -4 to $+8$ on template 3), we estimate that fewer than 20% of EP_0 complexes are able to form a stable ternary complex. For gaps in the region upstream of this (~ -13 to -5 on both templates), the inhibition is somewhat less severe although still quite significant.

In addition to the region flanking the transcription start site, removal of template strand nucleosides in the vicinity of the -35 region of the promoter appeared to interfere with escape from the promoter. This effect was more evident for template 1 (positions -32 and -33) than for template 3, where the gel data contained significant noise in this region.

Removal of nucleosides at positions $+19$ to $+23$ of the template strand on template 1 resulted in a modest inhibition of ternary complex formation. No corresponding result was observed on template 3, where if anything, removal of nucleosides in a similar region (positions $+19$ to $+24$) caused a slight enhancement of the ternary complex product. It should be noted, however, that this region may lie downstream of the enzyme's leading edge in the template 3 A11 complex and thus may have little bearing on the formation of that complex.

The results for the nontemplate strand for the two templates also showed some qualitative similarities (Fig. 6). In each case, removal of nucleosides in the -10 promoter consensus region or the adjacent 3–5 nucleotides downstream (-13 to -4 on template 1; -12 to -4 on template 3) resulted in a significant decrease in promoter escape. In contrast, removal of nucleosides immediately flanking the transcription start site (-1 to

$+5$ on template 1 and -2 to $+5$ on template 3) in general actually enhanced the production of stable ternary complexes. Outside of the region encompassing the -10 promoter element and transcription start site, the patterns for the two templates were qualitatively different. Although removal of nucleosides throughout the upstream region of template 1 caused varying degrees of inhibition (with the exception of positions -38 and -49), the corresponding region on template 3 showed minor effects, with the exception of the spacer region (-18 to -15), where removal of nucleosides modestly enhanced promoter escape. Downstream of the transcription start site, template 1 showed significant inhibition when nucleoside $+7$ was removed, mild inhibition in the region of $+10$ to $+13$, and moderate enhancement of promoter escape with removal of nucleosides in the region of $+23$ to $+29$. Template 3, on the other hand, showed mild to moderate enhancement in the region of $+8$ to $+12$ and significant enhancement when the nucleoside removed was in the region of $+13$ to $+16$. It is perhaps noteworthy that the downstream region of strongest enhancement on each template ($+23$ to $+29$ on template 1 and $+13$ to $+16$ on template 3) is in a similar position relative to the transcription halt point in this experiment (position A20 for template 1 and position A11 for template 3).

The region upstream of the promoter (~ -40 to ~ -60) on template 1 is extremely A/T-rich ($>85\%$) and marked by several runs of four or more adenines. The corresponding region on template 3 is somewhat less A/T-rich ($\sim 70\%$) and contains no runs of more than three adenines. This region has been implicated in affecting both promoter binding and subsequent transcription through contacts with the α subunit of RNA polymerase (26–28). Although our data for this region were not amenable to quantitation with GelExplorer, visual examination of line graphs of the gel data reveals some intriguing differences between the two templates in this region. Significant and periodic differences between the populations of gapped molecules found in the EP_0 nuc band and the ternary complex band are observed for template 1 (Fig. 7), especially in the vicinity of an A-tract found at positions -53 to -56 , where missing nucleosides severely interfere with promoter escape. Removal of nucleosides in the same region of template 3 shows only minor effects (data not shown).

Effect of Template Strand Gaps on Progress of the Initiating Complex—A discontinuity in the template strand during transcription would be expected to pose a challenge, if not a barrier, to an elongating polymerase. It at first seemed surprising, therefore, that there were several positions within the transcribed region on the template strand of each template (Fig. 4; positions $+12$ to $+18$ on template 1 and $+9$ to $+11$ on template 3) whose removal did not seem to impede the formation of a stable ternary complex. Analysis of the RNA in the ternary complex band for template 1 revealed only the expected A20 RNA (plus a small amount of readthrough A26) (data not shown). Therefore, it seemed possible that the polymerase was able to bypass discontinuities at these positions. Although bypass of template strand gaps by RNA polymerases has been reported (29), it has not been found to occur at gaps flanked by backbone phosphate ends, as would be a predominant product of hydroxyl radical treatment (30). Alternatively, it was possible that gaps at these positions indeed posed a barrier to the polymerase, but the halted complexes, with their truncated transcripts, were stable to electrophoresis and had mobility identical to the A20 complex. In this case we would assume that the lack of detection of the truncated transcripts was due to the small number of these particular gapped templates relative to the remainder of the population that was capable of supporting A20 production.

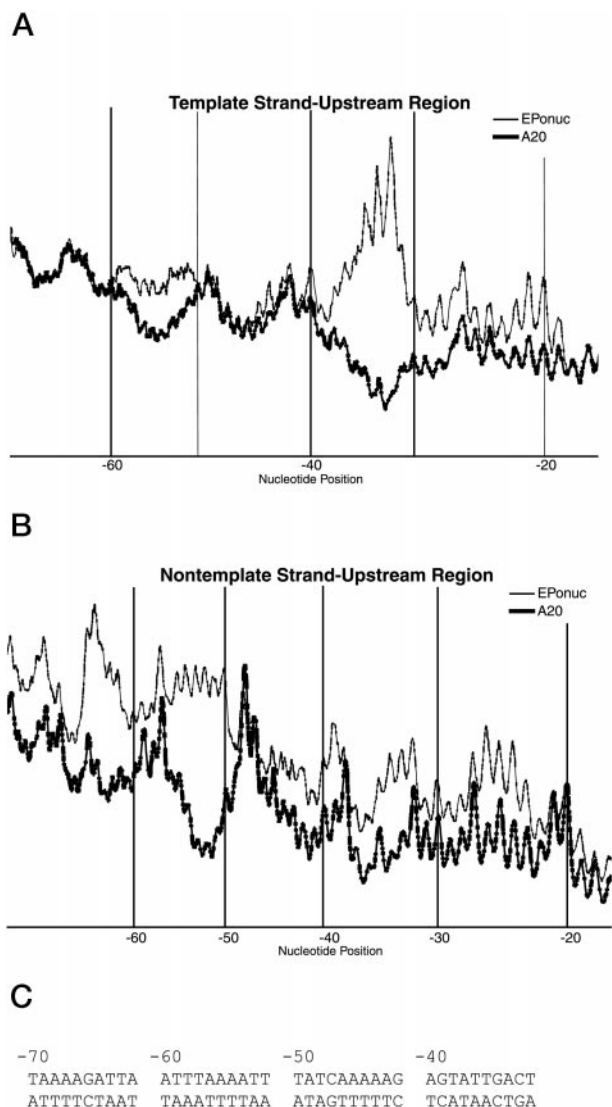


FIG. 7. A comparison of gaps represented in the open promoter and the A20 complex in the upstream region of template 1. Line scans of denaturing gel data for the template strand (A) and nontemplate strand (B) are shown. Light line, EP_{nuc} population; heavy line, A20 population. To allow comparison of these line scans, the data were normalized to the quantity of full-length DNA in each lane. The sequence of the upstream region is shown in (C).

That the latter case was true was determined by the following experiment. We used synthetic oligonucleotides to construct an artificial template 1 gapped specifically at position +13 on the template strand and containing phosphate moieties at the termini on either side of the single nucleoside gap. Transcription of this template (referred to as the G13 template) revealed that the gap indeed posed a barrier to the polymerase such that a 12-nucleotide transcript was formed (Fig. 8, lane 7) and that this 12-mer containing complex comigrated with the A20 complex on a nondenaturing gel (Fig. 9, compare lanes 2 and 4; Fig. 8, lane 16). A small amount of gap bypass to produce a 19-nucleotide RNA (under conditions for A20 synthesis) was observed (Fig. 8, lanes 7 and 16); however, this could be attributed to a small proportion of oligonucleotides that had not been successfully phosphorylated during synthesis. We infer that the presence of templates gapped at positions +12 to +18 on the template strand among the ternary complex population of template 1 is primarily not due to gap bypass but indicates that ternary complexes on this template containing RNAs as small as 11 nucleotides are stable to electrophoresis.

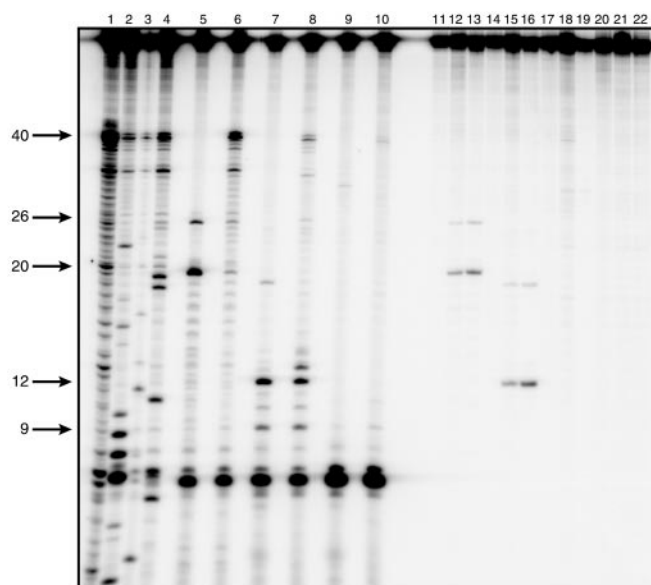


FIG. 8. Effect of specific gaps in the DNA template on transcription. Shown is a phosphorimage of a denaturing polyacrylamide gel on which was separated the RNA products of transcription. See "Experimental Procedures" for experimental details. Lane 1, ladder; lane 2, 3'-dGTP sequencing reaction; lane 3, 3'-dCTP sequencing reaction; lane 4, 3'-dATP sequencing reaction; lane 5, A20 transcription conditions with intact template; lane 6, run-off transcription conditions with intact template; lane 7, A20 transcription conditions with G13 template; lane 8, run-off transcription conditions with G13 template; lane 9, A20 transcription conditions with G10 template; lane 10, run-off transcription conditions with G10 template; lanes 11-19, band shift bands A-I (see Fig. 9); lane 20, intact template control; lane 21, template G13 control; lane 22, template G10 control. The lengths of the transcripts are indicated to the left of the gel image.

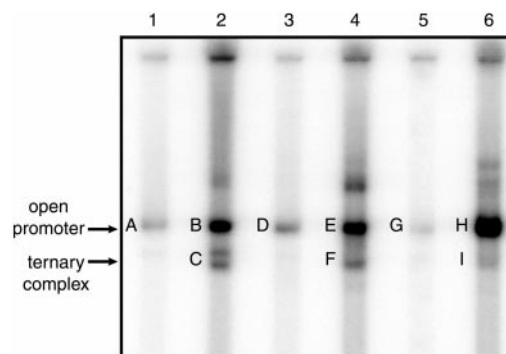


FIG. 9. Native gel electrophoresis of transcription complexes on intact and gapped oligonucleotide templates. Both DNA and RNA are radiolabeled. Lane 1, intact template plus RNA polymerase; lane 2, intact template, A20 transcription conditions; lane 3, G13 template plus RNA polymerase; lane 4, G13 template, A20 transcription conditions; lane 5, G10 template plus RNA polymerase; lane 6, G10 template, A20 transcription conditions. The reactions shown in lanes 5, 7, and 9 of Fig. 8 were loaded in lanes 2, 4, and 6, respectively, of this gel. Labeled bands A-I were excised and electrophoresed in lanes 11-19 of the gel in Fig. 8.

Interestingly, templates gapped at positions +9 through +11 of the template strand were represented significantly in the ternary complex population for template 3 but not for template 1 (Fig. 5). We wished to determine whether templates gapped at these positions supported incorporation of RNA nucleotides up to the position immediately preceding the gap (as suggested by the G13 template results above), producing a complex that was unstable to electrophoresis (and hence not isolated with the ternary complex population), or whether gaps at these positions interfered with some earlier step of transcription. To address this question, we used synthetic oligonucleotides to

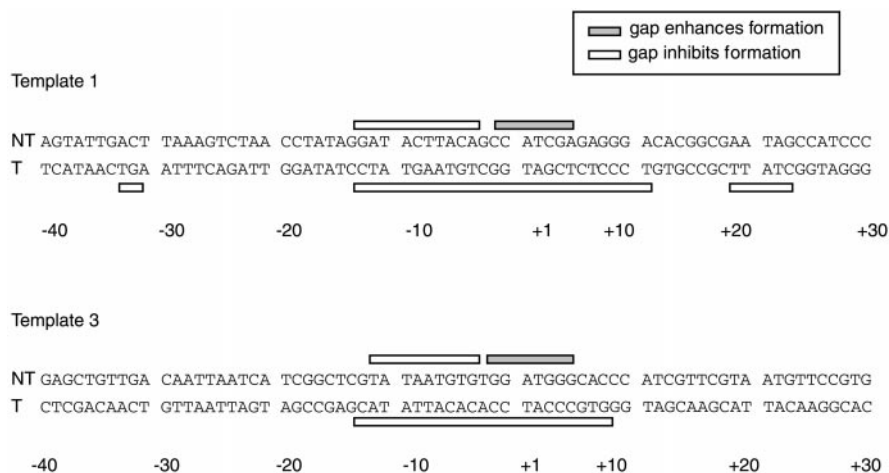


FIG. 10. Schematic summary of the effects of missing nucleosides on promoter escape on templates 1 and 3.

construct an artificial template 1 gapped specifically at position +10 on the template strand, again flanked by phosphate termini at the gap. Transcription and native gel electrophoresis of this specifically gapped template (referred to as G10 template) revealed that, as expected, no RNA-containing complex that comigrated with ternary complexes on intact templates was formed. Rather, the vast majority of DNA after transcription comigrated with EP_0 complexes and lacked associated RNA (Fig. 9, lane 6; Fig. 8, lane 18). Electrophoretic analysis of the RNA produced during transcription of this template revealed that the primary product, rather than the 9-nucleotide RNA that might have been expected, was an RNA of 5 or 6 nucleotides in length (Fig. 8, lane 9), which was produced in amounts somewhat greater than abortive transcripts produced on the intact or G13 templates. Therefore, complexes formed on the G10 template, unlike those formed on the G13 template, were unable to transcribe to a point that would have brought their active site within one nucleotide of a template strand gap and potentially allowed the complex to achieve the stable conformation of an elongating complex. The contrast between template 1 and template 3 in this respect underlines the importance of the promoter and/or initial transcribed sequence in the process of promoter escape (see "Discussion").

DISCUSSION

We have examined the effects of removing individual template nucleosides on the transition of RNA polymerase from a binary open promoter complex to a stable ternary elongation complex. The results for the two templates analyzed (schematized in Fig. 10) suggest some common features that may reflect the basic mechanism for this process, as well as some differences that may help reveal the influence of nucleotide sequence on the properties of individual promoters.

Perhaps the most obvious potential disruption to the transcription process resulting from a discontinuity in the DNA is a complete blockage of progress because of absence of a template strand base in the active site of the enzyme. Our results are consistent with previous reports that, at least when the gap is flanked by phosphate moieties, *E. coli* RNA polymerase is unable to bypass a single nucleoside gap in the template strand (29). However, the position of the gap affected how close to the gap the polymerase was able to approach. With a gap at position +13 on template 1, transcription proceeded all the way to the last available template base to produce a 12-nucleotide transcript, and this halted complex had the electrophoretic mobility of a mature elongating complex. In contrast, when the template was gapped at position +10, the longest observed RNA had a length of 5 or 6 nucleotides and was not stably associated with the enzyme-DNA complex, and the enzyme-

DNA complex comigrated with an open promoter complex, implying that σ factor had not been released. These results suggest that prior to σ release (*i.e.* for complexes in the "abortive initiation" phase of transcription), contacts with the template strand downstream of the active site are critical to the progress of the transcribing complex. Contacts upstream of the active site may of course also be important for these complexes. Within the transcribed region, however, we are unable to distinguish such effects because gaps are always encountered first by the leading edge of the polymerase.

It has been observed (24, 25) that removal of nucleosides from either strand of the DNA in the vicinity of the -10 consensus sequence enhances formation of the open promoter complex. Presumably, gaps at these positions cause a decrease in the free energy required to melt open the transcription bubble that more than compensates for any loss of enzyme-DNA contacts, thereby stabilizing the complex. Such stabilization of the open promoter complex might be expected to interfere with the process of promoter escape, because this requires disruption of interactions at the promoter (6). Indeed, for both templates studied, the region of the template strand in which removal of nucleosides has the most deleterious effect on promoter escape coincides with the region of enhanced EP_0 formation (-13 to +5). Whether this effect is solely due to stabilization of the EP_0 complex or whether contacts at some of these positions are actually required in the process of promoter escape is unclear. Perhaps the fact that within the -13 to +5 region, the strongest interference with escape occurs with gaps in the immediate vicinity of the start site (-4 to +5) suggests that these nucleosides are involved in contacts in intermediate complexes during the transition from abortive initiation to the elongating mode, whereas the effect at the upstream positions (-13 to -5) is only due to the above-mentioned EP_0 stabilization.

On the nontemplate strand, both templates also displayed reduced promoter escape when gapped at some, but not all, of the positions that enhance promoter binding (-6 to -3). (Note that the region of enhanced promoter binding does not extend into the -10 consensus region on the nontemplate strand of template 3²; this presumably reflects the requirement for nontemplate strand contacts in this region in the open complex (31)). Surprisingly, gaps in the remainder of this region (-2 to +5) resulted in enhanced formation of a stable ternary complex. Roberts and co-workers (32) have observed pausing of RNA polymerase early in the transcribed region of the λ pR' promoter that is attributed to movement of σ away from its initial contacts with the nontemplate strand in the -10 region to new contacts at similar sequences in the transcribed region.

If such contacts were to occur early in the initial transcribed region such that they interfered with the transition from an abortive to an elongating complex, then perhaps removal of the nucleosides responsible for these interactions could favor promoter escape. However, the closest match to the -10 consensus found in this region is a 2/6 fit from positions -3 to $+3$ for template 3, and a 2/6 fit from positions -1 to $+5$ for template 1. An alternative explanation might be that gaps in this region provide greater DNA flexibility, making the conformational changes necessary for promoter escape more favorable. If not for the obvious deleterious effect of a discontinuity in the template strand, gaps in the template strand in this region might also enhance escape. It is noteworthy that the nontemplate strand is important for lateral stability of elongating complexes (33); perhaps the flexibility in complex positioning caused by removal of contacts between the enzyme and the nontemplate strand is favorable for the process of promoter escape. Yet another possibility is that, in an effect analogous to the stabilizing of the open complex by gaps in the promoter melting region, gaps in the early transcribed region might stabilize an intermediate in promoter escape by favoring the melted form of the DNA. In such a view, reformation of duplex DNA in this region might expel short RNA and reset the open complex, thereby promoting abortive initiation. Again, though, the deleterious effects of gaps in the template strand argue for a dominant effect of templating over the role of DNA stability in the early transcribed region.

Although gaps that stabilize the open complex might interfere with promoter escape, gaps that remove important contacts in the open complex could be envisioned to have either a positive or a negative effect. If a protein-DNA contact provides binding energy to stabilize the complex but plays no functional role in subsequent events, then removal of that contact might enhance escape from the promoter. On the other hand, if a contact in the EP_o complex plays a role in a subsequent conformational change, lack of that contact may prevent the complex from functioning. Nucleosides on the nontemplate strand in the -10 consensus region of template 3 and on the template strand in the -35 region of both templates, may fall in this category. Because σ contacts both the -10 and -35 regions in the open complex (34), it is reasonable to expect that both of these regions could play a role in subsequent rearrangements of the complex and/or the exit of σ from the complex.

The notion that stabilization of the open promoter complex is counterproductive to the initiation process may be relevant to our observations in the region upstream of the -35 region of template 1. The region from -39 to -59 on template 1 contains an 8/9 and a 10/12 fit, respectively, to the proximal and distal subsites of the "UP" element consensus sequence defined recently by Estrem *et al.* (35), with the subsites positioned nearly identically to those observed by Estrem *et al.* The UP element has been implicated as a third promoter recognition element in some promoters whose strengths do not correlate solely with their homologies to the -10 and -35 core recognition elements (26); the T7 A1 promoter found on template 1 is one such promoter (36). There is substantial evidence that DNA upstream of the -35 element wraps around the enzyme (37–39) and that this wrapping may be facilitated by contacts in this region with the α subunit (28). Removal of nucleosides in this region of template 1 interfered significantly with promoter escape. Perhaps the increased flexibility of DNA gapped in this region caused a stabilization of the EP_o complex, thereby interfering with subsequent events. This would be consistent with the proposal by Rivetti *et al.* (39) that the energy cost of wrapping the upstream DNA around the enzyme during promoter binding is actually important to allow subsequent pro-

motor escape. Alternatively, perhaps the removal of an important enzyme-DNA contact in this region prevented a subsequent isomerization event. In either case, these results reveal an important difference between templates 1 and 3 and will be important to pursue to further elucidate the role of the α subunit in promoter binding and escape.

What about the role of nucleosides at the downstream edge of the advancing polymerase? As noted above, our results with singly gapped template 1 molecules suggest that interactions with the template strand downstream of the active site may be important for promoter escape. In this region we again noted significant differences in the behavior of the two templates studied. Although templates gapped at positions $+9$, $+10$, and $+11$ on the template strand were poorly represented in the ternary complex population with template 1, template 3 molecules gapped at these positions were well represented in the ternary complex band, implying that σ factor had been lost from these complexes. Krummel and Chamberlin (1) showed that with template 3, σ is still present in complexes that have transcribed to position 7 but is released after incorporation of the 8th or 9th nucleotide. Our results therefore imply that for template 3, the polymerase active site is able to progress to position 8 and undergo σ loss even when the template strand is gapped one nucleotide downstream at position 9. On template 1, in contrast, a continuous template strand out to position 11 is required for σ release. This result adds to the considerable evidence that sequences in the initial transcribed region as well as the promoter region itself influence the properties of the initial transcribing complex (1, 5, 7). It has been suggested (7, 40) that the nucleotide sequence in the initial transcribed region plays a role in keeping the polymerase active site in register with the 3' end of the nascent transcript and that release of abortive transcripts is triggered by a backsliding of the active site. Our results would predict that interactions with the template strand immediately downstream of the active site play an important role in this process.

The strength of the experimental approach reported here is in the ability to screen, simultaneously, a large number of template positions for interesting roles in the dynamic process of transcription initiation. Because the transition from open promoter complex to the final complex studied here is a multi-step process, we cannot be sure at which step any particular nucleoside gap exerts its effect; rather, we refer to an effect on the composite process, calling it "promoter escape." Furthermore, because the criterion used to define promoter escape in these experiments is based on a change in electrophoretic mobility of an enzyme-DNA complex rather than a functional definition, the properties of the final complexes in the "ternary complex" band are likely not uniform. Indeed, it is clear that this population includes not only ternary complexes with the expected 20 nucleotide transcript (for template 1) but also complexes halted at earlier positions because of gaps in the template strand and perhaps even complexes that have lost σ factor but whose transcripts are not stably associated. Because of these limitations, the effects identified here must be viewed as a starting point for more detailed studies focusing on specific template positions. Templates mutated or specifically gapped at positions of interest must be constructed and characterized structurally and functionally to elucidate the mechanisms accounting for the effects observed here.

REFERENCES

1. Krummel, B., and Chamberlin, M. J. (1989) *Biochemistry* **28**, 7829–7842
2. Carposis, A. J., and Gralla, J. D. (1985) *J. Mol. Biol.* **183**, 165–177
3. Record, M. T., Jr., Reznikoff, W. S., Craig, M. L., McQuade, K. L., and Schlax, P. J. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., Curtis, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. R., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umberger, H. E., eds) pp. 792–820, American Society for Microbiology,

Washington, D. C.

4. Gralla, J. D., Carpousis, A. J., and Stefano, J. E. (1980) *Biochemistry* **19**, 5864–5869
5. Kammerer, W., Deuschle, U., Gentz, R., and Bujard, H. (1986) *EMBO J.* **5**, 2995–3000
6. Ellinger, T., Behnke, D., Bujard, H., and Gralla, J. D. (1994) *J. Mol. Biol.* **239**, 455–465
7. Hsu, L. M., Vo, N. V., and Chamberlin, M. J. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 11588–11592
8. Nudler, E. (1999) *J. Mol. Biol.* **288**, 1–12
9. Uptain, S. M., Kane, C. M., and Chamberlin, M. J. (1997) *Annu. Rev. Biochem.* **66**, 117–172
10. Siebenlist, U., and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 122–126
11. Brunelle, A., and Schleif, R. F. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6673–6676
12. Hayes, J. J., and Tullius, T. D. (1989) *Biochemistry* **28**, 9521–9527
13. Straney, D. C., and Crothers, D. M. (1985) *Cell* **43**, 449–459
14. Krummel, B., and Chamberlin, M. J. (1992) *J. Mol. Biol.* **225**, 221–234
15. Kashlev, M., Martin, E., Polyakov, A., Severinov, K., Nikiforov, V., and Goldfarb, A. (1993) *Gene (Amst.)* **130**, 9–14
16. Studier, F. W., and Rosenberg, A. (1981) *J. Mol. Biol.* **153**, 503–527
17. Telesnitsky, A. P. W., and Chamberlin, M. J. (1989) *J. Mol. Biol.* **205**, 315–330
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
19. Arndt, K. M., and Chamberlin, M. J. (1990) *J. Mol. Biol.* **213**, 79–108
20. Zhou, W., and Doetsch, P. W. (1994) *Biochemistry* **33**, 14926–14934
21. Shadle, S. E., Allen, D. F., Guo, H., Pogozielski, W. K., Bashkin, J. S., and Tullius, T. D. (1997) *Nucleic Acids Res.* **25**, 850–860
22. Levin, J. R., Krummel, B., and Chamberlin, M. J. (1987) *J. Mol. Biol.* **196**, 85–100
23. Metzger, W., Schickor, P., Meier, T., Werel, W., and Heumann, H. (1993) *J. Mol. Biol.* **232**, 35–49
24. Werel, W., Schickor, P., and Heumann, H. (1991) *EMBO J.* **10**, 2589–2594
25. Li, X.-Y., and McClure, W. R. (1998) *J. Biol. Chem.* **273**, 23558–23566
26. Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. (1993) *Science* **262**, 1407–1413
27. Ellinger, T., Behnke, D., Knaus, R., Bujard, H., and Gralla, J. D. (1994) *J. Mol. Biol.* **239**, 466–475
28. Aiyar, S. E., Gourse, R. L., and Ross, W. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14652–14657
29. Liu, J., and Doetsch, P. W. (1996) *Biochemistry* **35**, 14999–15008
30. Balasubramanian, B., Pogozielski, W. K., and Tullius, T. D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9738–9743
31. Roberts, C. W., and Roberts, J. W. (1996) *Cell* **86**, 495–501
32. Ring, B. Z., Yarnell, W. S., and Roberts, J. W. (1996) *Cell* **86**, 485–493
33. Wang, D., and Landick, R. (1997) *J. Biol. Chem.* **272**, 5989–5994
34. Dombroski, A. J., Walter, W. A., Record, M. T., Jr., Siegele, D. A., and Gross, C. A. (1992) *Cell* **70**, 501–512
35. Estrem, S. T., Ross, W., Gaal, T., Chen, Z. W. S., Niu, W., Ebright, R. H., and Gourse, R. L. (1999) *Genes Dev.* **13**, 2134–2147
36. Deuschle, U., Kammerer, W., Gentz, R., and Bujard, H. (1986) *EMBO J.* **5**, 2987–2994
37. Craig, M. L., Suh, W.-C., and Record, M. T., Jr. (1995) *Biochemistry* **34**, 15624–15632
38. Nickerson, C. A., and Achberger, E. C. (1995) *J. Bacteriol.* **177**, 5756–5761
39. Rivetti, C., Guthold, M., and Bustamante, C. (1999) *EMBO J.* **18**, 4464–4475
40. Feng, G., Lee, D. N., Wang, D., Chan, C. L., and Landick, R. (1994) *J. Biol. Chem.* **269**, 22282–22294