Promoter Escape by RNA Polymerase II

DOWNSTREAM PROMOTER DNA IS REQUIRED DURING MULTIPLE STEPS OF EARLY TRANSCRIPTION*

Received for publication, October 23, 2002, and in revised form, January 13, 2003
Published, JBC Papers in Press, January 13, 2003, DOI 10.1074/jbc.M210848200

Xiaoxue Wang, Lori Spangler, and Arik Dvir‡
From the Department of Biological Sciences, Oakland University, Rochester, Michigan 48309

Recent evidence, obtained in a reconstituted RNA polymerase II transcription system, indicated that the promoter escape stage of transcription requires template DNA located downstream of the elongating polymerase. In the absence of downstream DNA, very early elongation complexes are unable to synthesize transcripts longer than –10–14 nucleotides. In contrast, once transcripts longer than –15 nucleotides have been synthesized, an extended region of downstream DNA is no longer required (Dvir, A., Tan, S., Conaway, J. W., and Conaway, R. C. (1997) J. Biol. Chem. 272, 28175–28178). In this work, we sought to define precisely when, during the synthesis of the first 10–15 phosphodiester bonds, downstream DNA is required. We report that, for complete promoter escape, downstream DNA extending to position 40/42 is required. The polymerase can be forced to arrest at several points prior to the completion of promoter escape by removing downstream DNA proximally to positions 40/42. The positions at which the polymerase arrests appear to be determined by the length of available downstream DNA, with arrest occurring at a relatively fixed position of –28 nucleotides to the distal end of the template. A similar requirement is observed for transcription initiation, i.e. the formation of the first phosphodiester bond of nascent transcripts. In addition, we show that the requirement for a downstream region is independent of downstream DNA sequence, suggesting that the requirement reflects a general mechanism. Taken together, our results indicate (i) that downstream DNA is required continuously through the synthesis of the first 14–15 phosphodiester bonds of nascent transcripts, and (ii) that a major conformational change in the transcription complex likely occurs only after the completion of promoter escape.

Eukaryotic messenger RNA synthesis is a complex biochemical process that depends on RNA polymerase II and a variety of general and gene-specific transcription factors. Much information about the function of the RNA polymerase II transcription complex has been obtained in reconstituted, in vitro transcription systems, in which the contributions of individual cofactors to rate-limiting steps can be specifically evaluated (1, 2). Stable binding of RNA polymerase II to the promoter requires minimally the presence of the general transcription factors TFIIH (1) (TFIIB, TFIIF, and TFIIF). Before transcription can begin, the double-stranded structure of the DNA template surrounding the initiation site on the promoter needs to be melted into single-stranded DNA in a process that is referred to as open complex formation (3). Open complex formation and transcription initiation (i.e. the formation of the first phosphodiester bond (Refs. 4–6)) depend on the presence of two additional general transcription factors, TFIEH and TFIIFH, and are catalyzed by an ATP(dATP)-dependent DNA helicase activity associated with TFIIFH (6–9).

Transcription initiation is followed by a short phase that is referred to as promoter escape. This phase primarily includes the formation of the first 10–15 phosphodiester bonds of nascent RNA transcripts and is characterized by functional instability of the RNA polymerase II transcription complex (10). In the absence of either TFIIFH or an ATP(dATP) cofactor, early RNA polymerase II elongation intermediates are prone to premature arrest at –10 to –14 base pairs downstream of the transcriptional start site (11, 12). In contrast, further transcription elongation by very early RNA polymerase II elongation intermediates that have successfully synthesized transcripts 14 or 15 nucleotides long requires neither TFIIFH nor an ATP(dATP) cofactor. Finally, completion of promoter escape depends on the presence of an extended region of downstream DNA; digestion of a duplex template containing the AdML promoter with a restriction enzyme that cuts the template 35/39 nucleotides downstream of the DNA template has no effect on initiation but results in arrest by RNA polymerase II at a position 10–15 nucleotides downstream of the transcriptional start site. However, once polymerase has successfully synthesized –14 nucleotide transcripts, digestion of the template with the same restriction enzyme has no effect on further elongation by RNA polymerase II (13). Based on these characteristics, we operationally define early RNA polymerase II elongation complexes that have synthesized transcripts –15 nucleotides or longer as those that have successfully escaped the promoter.

During the transcription cycle, promoter escape follows immediately after initiation. There are a number of notable similarities between these two very early stages of transcription; both share a requirement for an ATP(dATP) cofactor and depend on the presence of TFIIE and TFIIFH (6, 11, 12, 14), and both are suppressed by mutations in the same DNA helicase subunit of TFIIE and TFIIFH, encoded by the Xeroderma pigmentosum complementation group B (XBP) gene (8, 9, 15, 16), arguing that promoter melting is required for both transcription initiation and promoter escape and is likely catalyzed by the same helicase activity (13).
Promoter Escape by RNA Polymerase II

mechanism in both steps. Finally, initiation, like promoter escape, depends on the presence of an extended downstream DNA region, which has been shown to extend to somewhere between 23 and 35 nucleotides downstream of the transcriptional start site (13).

Although previous studies have provided strong support for the model that interactions between component(s) of the transcription initiation and downstream DNA are important for both initiation and promoter escape, a variety of questions regarding the structure and function of downstream DNA remain unanswered. These include the following. (i) What are the precise boundaries of the regions required to support transcription initiation and promoter escape? (ii) Does the requirement for downstream DNA in promoter escape depend on downstream DNA sequence? (iii) Is the completion of promoter escape accompanied by a major rearrangement of the transcription complex? The experiments presented here provide evidence that the downstream DNA requirements do not change significantly through the synthesis of the first 14–15 bonds of nascent transcripts, that downstream DNA operates in a sequence-independent manner, and that a major conformational change in the transcription complex likely happens at the completion of promoter escape.

MATERIALS AND METHODS

DNA Templates—The primary DNA template used in this study is M13mp19-AdML, which contains original AdML promoter sequences from –50 to 10 (12). AdML promoter mutants were prepared by the uracil-containing DNA method for site-directed mutagenesis (17), using the Bio-Rad Mutagen 2 system. Resulting clones were verified by sequencing. For transcription, a 444-base pair fragment was produced by PCR from M13mp19-AdML. The primers were 5′-GAGGGCGGTGTGAATTCC and 5′-CCACGGTGGACCGCTTGC. The resulting DNA fragment, which contains sequences that extend 77 bp upstream and 367 bp downstream of the transcriptional start site, was gel-purified by agarose gel electrophoresis prior to use in transcription reactions.

RNA Markers—Hybrid templates were prepared containing the core T7 promoter fused to AdML promoter sequences with downstream HaeIII restriction sites. The T7 promoter sequence was introduced to the AdML templates by performing PCR using M13mp19-AdML derivative templates as templates. We used an upstream PCR primer consisting of the T7 promoter: 5′-25-ATGGTAACCTAATAGCTGCTCTTAGGAGAAGCTCCGCTCTGAGTCC. The underscored sequence is a 20-base sequence complimentary to the AdML template strand at promoter positions 1–20. At the 5′ end of the T7 promoter, a six-base “clamp” sequence has been added (18). The downstream primer was 5′-CCACGGTGGACCGCTTGC, identical to that used in the AdML template amplification reactions. The M13mp19-AdML templates used had HaeIII sites positioned at 34, 40, and 44 relative to the AdML start site. Because T7 begins transcription at –6 relative to the AdML insertion, after digestion of templates with HaeIII the resulting transcripts are 40, 46, and 50 nt long. T7-AdML template amplification was carried out in 100 μl with 5 units of Taq DNA polymerase, 20 μM Tris-HCl, pH 8.0, 1.5 mM MgCl2, 50 mM KC1, 0.2 mM each of dNTPs (dATP, dCTP, dGTP, dTTP), 500 μM of each primer, and 830 fmol of M13mp19-AdML template DNA. The length of the PCR product is 398 bp. PCR products were gel-purified by agarose gel electrophoresis.

RNA Polymerase II and Transcription Factors—RNA polymerase II, TFIIE and TFIIF, and TFIHII (19) and TFIHII (20) were purified from rat liver nuclear extracts as described. Recombinant yeast TBP (21, 22) and TFIIB (23) were expressed in Escherichia coli and purified as described. Recombinant TFIIE was prepared as described (24), except that the 56-kDa subunit was expressed in E. coli BL21 (DE3)-pLysS. Recombinant TFIIF was produced by a modified protocol from E. coli M15 (DE3) infected with M13mpET-RAP30 and M13mpET-RAP74.

Transcription Experiments—Transcription experiments are performed in vitro using a reconstituted transcription system that includes RNA polymerase II, the five general transcription factors (TBP, TFIIB, TFIIE, and TFIIF) and purified polymerase. RNA Templates—Hybrid templates were prepared containing the core T7 promoter: 5′-HaeIII site. Because T7 begins transcription at –6 to form the transcriptional start site, the resulting transcripts are 40, 46, and 50 nt long. T7-AdML template amplification was carried out in 100 μl with 5 units of Taq DNA polymerase, 20 μM Tris-HCl, pH 8.0, 1.5 mM MgCl2, 50 mM KC1, 0.2 mM each of dNTPs (dATP, dCTP, dGTP, dTTP), 500 μM of each primer, and 830 fmol of M13mp19-AdML template DNA. The length of the PCR product is 398 bp. PCR products were gel-purified by agarose gel electrophoresis.

RNA Polymerase II and Transcription Factors—RNA polymerase II, TFIIE and TFIIF, and TFIHII (19) and TFIHII (20) were purified from rat liver nuclear extracts as described. Recombinant yeast TBP (21, 22) and TFIIB (23) were expressed in Escherichia coli and purified as described. Recombinant TFIIE was prepared as described (24), except that the 56-kDa subunit was expressed in E. coli BL21 (DE3)-pLysS. Recombinant TFIIF was produced by a modified protocol from E. coli M15 (DE3) infected with M13mpET-RAP30 and M13mpET-RAP74.

Transcription Experiments—Transcription experiments are performed in vitro using a reconstituted transcription system that includes RNA polymerase II, the five general transcription factors (TBP, TFIIB, TFIIE, and TFIIF), and gel-purified promoter-DNA. Promoter DNA and proteins are combined in a final volume of 30 μl to form preinitiation complexes during a 30 °C, 45-min incubation (11). The incubation buffer contains (final concentrations are given) 0.3 mM HEPES-NaOH, pH 7.9, 25 mM Tris-HCl, pH 7.9, 25 mM KC1, 4 mM MgCl2, 0.2 mM EDTA, 1 mM dithiothreitol, 0.5 mM mg/ml bovine serum albumin, 2% (v/v) polyvinyl alcohol, and 6% (v/v) glycerol. Each reaction included 20 ng of AdML DNA fragment, 50 ng of recombinant TBP, 10 ng of recombinant TFIIB, 20 ng of recombinant TFIIF, 20 ng of recombinant TFIIE, 150 ng of highly purified TFIHII, and 0.01 units of RNA polymerase II. Digestion of template DNA by PstI or HaeIII endonuclease was included in various phases of the transcription experiment by adding 2 μl of a solution containing 0.1–0.5 units of the respective enzymes. Digestions were carried at 30 °C for the times indicated, mostly 20 min. Separate experiments were carried out to verify that these conditions allowed for complete digestions.

Transcription was initiated by a labeling mix containing [α-32P]CTP (3,000 Ci/mmol), a CpU dinucleotide primer, ATP or dATP cofactor, and other ribonucleoside triphosphates as described in the figure legends. The volume of reaction mixtures was 35 μl. Transcription was stopped at 30 °C for the time indicated in the figure legends. Transcription was stopped by addition of 15 μl of stop solution containing 100 mM EDTA. A 55-μl loading dye containing 10.0 μM urea, 0.025% bromophenol blue, and 0.025% xylene cyanole was added to each sample. Samples were then heated to 90 °C for 3 min, briefly centrifuged to remove insoluble particles, and separated on a 2% acrylamide, 3% bisacrylamide, 6.0 μM urea gel as described (6) and visualized by autoradiography. For run-off transcription experiments, reactions were stopped by addition of 0.05 M EDTA, 0.1 M NaCl, and 0.5% SDS to the reaction mixture. Precipitation is accomplished by ethanol followed by a 70% ethanol wash. Dried pellets were resuspended in 27 μl of solution of 10 μM urea, 0.025% bromophenol blue, and 0.025% xylene cyanole FF, heated at 90 °C for 5 min, then loaded on urea-containing acrylamide gels for electrophoretic separation of transcripts.

TT-AdML templates obtained by PCR above were digested with HaeIII prior to transcription to combine TT-AdML PCR product, 1 unit of HaeIII, and 1 × enzyme buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl2, and 50 mM NaCl) in a 10-μl reaction incubated for 15 min at 37 °C. The RNA markers were synthesized in vitro as follows; reactions of 50 μl included 50 units of TT RNA polymerase, 1 × enzyme buffer (40 mM Tris-HCl, pH 7.9, 6 mM MgCl2, 2 mM spermidine, 10 mM dithiothreitol), 0.1 mg/ml bovine serum albumin, 100 μM ATP, 100 μM UTP, 100 μM GTP, 5 μM cold CTP, 67 μM [α-32P]CTP, and pre-digested TT-AdML DNA as template. Reactions were incubated for 15 min at 37 °C, then heated to 90 °C for 5 min, cooled on ice, and loaded on 7.5% urea-containing acrylamide gels for electrophoretic separation of transcripts.

Results

Experimental Strategy and Assays for Initiation and Promoter Escape—To map precisely the extent of downstream DNA required at specific stages of early transcription, we generated a series of AdML promoter-containing constructs that contain restriction sites for the endonuclease HaeIII at various distances downstream of the transcription initiation site. The locations of the newly inserted cleavage sites correspond to promoter positions 26–50 from the start site in the AdML promoter (Fig. 1). When treated with HaeIII, these DNA templates acquire new, shorter ends that differ by 2-nucleotide increments from each other. In addition, following cleavage by HaeIII, templates are left with blunt ends, and the 3′ end of the templates differs from the sequence of the original, parental plasmid by at most 2 nucleotides.

In our studies of early transcription by RNA polymerase II, we utilized a transcription system reconstituted with recombinant TBP, TFIIB, TFIIE, and TFIIF and purified polymerase and TFIIHII from rat liver. Promoter-specific initiation was assayed by measuring synthesis of abortive, dinucleotide-primed trinucleotide transcripts. As shown previously, transcription initiation by RNA polymerase II from the AdML promoter can be primed by a variety of dinucleotides. These dinucleotides must be complementary to template DNA surrounding the transcriptional start site (26). We assayed the synthesis of the first phosphodiester bond of nascent transcripts by measuring synthesis of trinucleotide transcripts in reactions containing the initiating dinucleotide CpU and [α-32P]CTP. These nucleotides support synthesis by polymerase of radioactively labeled C5P5C transcripts initiated at a position 3 base pairs upstream of the normal AdML transcriptional start site.

To measure promoter escape, we monitor the synthesis of...
short transcripts in reactions containing the initiating dinucleotide CpU, ATP, UTP, [α-32P]CTP, as well as the RNA chain-terminating nucleotide 3′-O-methylguanosine triphosphate (3′-O-MeGTP). The maximal transcript length under these conditions is 18 nucleotides, determined by the insertion of 3′-O-MeGTP at the first G downstream of the initiation site. Because early RNA polymerase II elongation intermediates that have synthesized transcripts of −15 nucleotides or longer are considered to have successfully escaped the promoter, formation of the 3′-O-MeGTP-terminated transcripts is a useful assay for the completion of promoter escape (11–13).

The Region of Downstream DNA Required for Initiation Extends to 34 Base Pairs Downstream of the Transcriptional Start Site—Previous experiments (13) have shown that initiation is strongly inhibited by cleaving an AdML promoter-containing template (pDN-AdML) with PstI, which cuts at position 29/27 and to a lesser degree by cleaving with SphI or HindIII, which cut at positions 29/33 and 35/39, respectively. However, these results only roughly defined the downstream border of the DNA required for initiation. In addition, because HindIII leaves a 5′ overhang and PstI leaves a 3′ overhang, the results of these earlier experiments could have been affected by the difference in DNA ends left by the different restriction enzymes used.

In the experiments presented here, we used the abortive initiation assay described above to compare the activities of the Ad+26 through Ad+32 templates, with or without prior cleavage by HaeIII. As shown in Fig. 2, the promoters on all templates were capable of supporting abortive initiation when not cleaved with HaeIII, although several (Ad+30 and Ad+34) appeared somewhat less efficient. Following HaeIII cleavage, however, very little initiation occurred on the Ad+26 and Ad+28 templates. An intermediate level of initiation was observed following HaeIII cleavage of the Ad+30, Ad+32, and Ad+34 templates, whereas HaeIII cleavage had little or no effect on initiation from the Ad+36 and Ad+40 templates. Thus, DNA downstream of position 34 is largely dispensable for initiation from the AdML promoter, and the region of downstream DNA most critical for initiation extends to 28 base pairs downstream of the transcriptional start site.

During Promoter Escape, Efficient Elongation by RNA Polymerase II Requires the Presence of DNA Extending −28 Base Pairs Downstream of the Transcript 3′ End—To determine the extent of downstream DNA required for efficient promoter escape, templates that can be cleaved with HaeIII at positions from 34 to 46 base pairs downstream of the site of transcription initiation were tested for their abilities to support promoter escape, and with and without prior treatment with HaeIII. As shown in Fig. 3, both transcription initiation and promoter escape were supported on all templates not treated with HaeIII. Cleavage of the Ad+42, Ad+44, and Ad+46 templates with HaeIII prior to transcription reactions had very little or no effect on the efficiency of promoter escape, as measured by synthesis of 18 nucleotide, 3′-O-MeG-terminated transcripts. Thus, DNA downstream of 42 is dispensable for promoter escape. In contrast, the efficiency of promoter escape was substantially reduced when the Ad+34, Ad+36, Ad+38, and Ad+40 templates were treated with HaeIII prior to the reactions, indicating that DNA extending to 40 is needed for efficient promoter escape. Notably, the length of the longest major transcripts synthesized following HaeIII cleavage depended on the position of the HaeIII cleavage sites. Thus, the lengths of the longest major transcripts were 8, 11, and 13–14 nt following cleavage with HaeIII at 34, 36, and 38, respectively, whereas cleavage at 40 allowed synthesis of a reduced level of 18-nt, 3′-O-MeG-terminated transcript.

In Fig. 3, HaeIII digestion is performed before transcription initiation. To examine the effect of downstream DNA on promoter escape alone, we utilized a two-step, pulse-chase transcription protocol. In the first stage of the reaction, RNA synthesis was carried out in the presence of limiting concentrations of radioactive nucleotides, resulting in the formation of “pre-escaped” transcription complexes containing 3–9-nucleotide-long RNA transcripts. This initiation stage was followed by a chase phase, in which a large excess of unlabeled nucleotides was added, allowing further extension of transcripts and completion of promoter escape. Reaction conditions can be selectively changed between the two stages, allowing for the assessment of specific template and cofactor requirements for promoter escape.

In the experiment of Fig. 4, we included a 15-min incubation with HaeIII to allow cleavage of downstream DNA between the pulse and chase stages of the reactions. This experiment was performed using promoters with restriction sites at positions 36, 38, 40, 42, and 44. Four transcription reactions were performed with each of the mutant templates; transcription in two of the reactions included only the labeling step, whereas in the other two it also included the cold chase phase. In each pair, one reaction was incubated with the HaeIII endonuclease at the end of the labeling stage. The other reaction served as a “no HaeIII” control. Transcripts at the end of the labeling stage of the reaction were 3–9 nucleotides long. The 3-nucleotide-long product is abortive, and the remaining bands corresponded to the labeled transcripts.
primarily to 5- and 7-nucleotide-long transcripts. Addition of HaeIII to the “pulse-only” reactions did not significantly affect the length or amount of transcripts detected.

The other two lanes shown for each template are reactions in which a cold nucleotide chase phase was added following the

Due to the presence of ATP. As in the first two lanes, one reaction included a HaeIII treatment and the other was a “no HaeIII” control. In all the templates in the control lanes, promoter escape was supported to a similar level, as judged by the level of 18-nucleotide-long, 3’-O-MeGTP-terminated transcripts. In reactions that included a HaeIII treatment, significant differences appeared between the various templates. On templates with the HaeIII restriction sites at positions 36 and 38, promoter escape was substantially suppressed. The 40 promoter showed an intermediate level of inhibition, and the 42 promoter showed only slight inhibition. Promoter escape on the template with a HaeIII site at 44 was not affected by the endonuclease. Therefore, in the two-step promoter escape experiment, DNA up to position 40/42 from the transcription initiation site appears critical for efficient promoter escape. This result is in full agreement with the result obtained in the single-step promoter escape experiment (Fig. 3).

The experiment presented in Fig. 4 provides additional information regarding transcription complexes that became arrested at promoter-proximal positions. The reduction in the level of promoter escape is accompanied by an increase in the level of promoter proximal arrest. This is evident by a new pattern of arrested RNA transcripts 11–15 nucleotides in length formed in reactions containing templates cut at positions 36 through 42. Small, but highly reproducible differences in the size distribution of arrested transcripts formed on the various templates can be recognized. Arrested transcripts were primarily 11 and 12 nucleotides in length on the 36 promoter, 12 and 13 nucleotides for the 38 template, and 13 and 14 nucleotides for the 40 template. On the 42 template, which is only slightly affected by HaeIII treatment, arrested transcripts were primarily 14 and 15 nucleotides in length.

Because transcription is primed with CpU at position −3 relative to the in vivo initiation site, the observed RNA length less 3 nucleotides corresponds to the precise location of the RNA polymerase catalytic site at the time when its progress was stopped. The distance between the end of the downstream DNA and the size of paused transcripts, as seen in Fig. 4, is 

Downstream DNA Functions in a Sequence-independent Manner—During promoter escape, the downstream DNA region does not function as a template for transcription as it is at a distance of more than 20 nucleotides downstream from the catalytic site of the polymerase. Instead, it is likely that the downstream region forms critical contacts with component(s) of the transcription complex. These interactions might directly involve TFIIH, based on evidence from UV-induced protein-DNA cross-linking studies (27, 28) and functional studies (29).

To begin to address the question whether downstream DNA functions in a sequence-dependent or sequence-independent manner, we introduced a series of sequence mutations in the downstream region of M13mp19-AdML, substituting every base pair in promoter positions 25 through 42 (Fig. 5A). To increase the odds of perturbing sequence-specific interactions, we have consistently substituted purines for pyrimidines and vice versa.

To determine whether ATP-dependent promoter escape is sensitive to downstream DNA sequence, we utilized the two-step promoter escape protocol described in Fig. 4. Transcription complexes were first assembled on the various promoters (three identical reactions for a single promoter mutant). Fol-
Substituted nucleotides in each derivative are highlighted in the mutation relative to the transcription initiation site (1). The H11001plexes and short RNA chains, 100 mM UTP, CTP, 150 mM GTP, and ATP. The formation of preinitiation complexes and Methods.

As mentioned above, published functional (29) and structural (27, 28) data suggest that TFIIH physically occupies DNA downstream of the active site of the polymerase. An explanation for our observation that RNA polymerase II that has not yet escaped the promoter tends to arrest transcription at 28 nt proximally to the 3' end of the available template may be that TFIIH bound to the downstream DNA blocks further elongation by the enzyme. If this is the case, we should be able to determine when TFIIH (or other template-associated components of the transcription complex) are released by determining when RNA polymerase II gains the capacity to synthesize transcripts that extend to the end of the DNA template.

In the experiment shown in Fig. 6, DNA templates with HaeIII endonuclease restriction sites inserted incrementally between positions 40 and 50 downstream of the AdML promoter (Fig. 1). As shown earlier in Figs. 3 and 4, the removal of downstream DNA proximally to 40 causes arrest of early transcription, whereas those cut downstream of 40 have little or no effect on the formation of 18-nt transcripts. Therefore, constructs with restriction sites in the 40–50 range would be suitable to determine whether the polymerase can utilize the entire downstream region as a template.

In the experiment shown in Fig. 6, DNA templates with HaeIII restriction sites at 34, 44, 46, and 50 were digested with HaeIII, and incubated with the general transcription factors (Materials and Methods). The addition of a nucleotide mixture containing CTP, ATP, UTP, GTP, and [α-32P]CTP, allowing for the formation of a full-length runoff transcripts. To accurately determine whether transcripts produced under the experimental conditions correspond to the full runoff length of the template, we prepared a series of sequence-specific RNA size markers. These markers were created by using a bacterial T7 RNA polymerase promoter fitted with AdML sequence downstream of 1, including the HaeIII restriction sites as in the M13mp19-AdML series. Lanes 1–3 show the run-off products generated by bacterial T7 RNA polymerase and the size indi-
cated by arrows, corresponding to 40, 46, and 50 nucleotides in length. Lanes 4–7 show transcripts synthesized by RNA polymerase II from AdML promoters that were cleaved at 50, 46, 44, and 34, respectively. In each of lanes 4, 5, and 6, there is a prominent transcript that corresponds to the maximal runoff size allowed by the shortened template. The template cleaved at 34 (lane 7) lacks any apparent transcripts, because of promoter proximal arrest of the transcription complex prior to the formation of transcripts 14/15 nt long (see Figs. 2 and 3). The short transcripts produced under these conditions migrated all the way through the gel to the lower tank (data not shown). This lane provides an internal control for the products that resolve on lanes 4–6.

Several premature stops in elongation can be seen in lanes 4–6, indicated by shorter RNA transcripts, with the most prominent one ~36 nt in length. For the most part, these are of the same size in all three lanes and do not appear to display an obvious regularity of recurrence. The nature of these stops remains to be determined, although these may indicate natural pause sites in the template (30, 31). More importantly, an RNA band of a size that would indicate a fixed spacing of 25–28 nt from the distal edge of the three templates cannot be identified. These results suggest that downstream DNA is not physically occupied after the formation of the 14th or 15th phosphodiester bond as it is prior to that point, because full runoff transcripts can be produced. These results are consistent with the hypothesis that a conformation change occurs after the completion of promoter escape at 14/15 and that the change likely involves the release of TFIIH.

**DISCUSSION**

In the experiments described here, we utilized linear DNA templates that differ in length by 2-nucleotide increments to obtain an accurate measure of the downstream DNA that is required during very early transcription. First, we determine that downstream DNA required to carry out transcription initiation (first phosphodiester bond formation) extends to ~34 nucleotides (Fig. 2) and that downstream DNA required for complete promoter escape extends to ~42 nucleotides from the transcription start site (Figs. 3 and 4). These numbers are consistent with the previously estimated 35–50 range for promoter escape and 23–35 range for initiation (13). The polymerase is not committed to elongation of transcripts until the promoter escape stage is completed; therefore, the 42-nucleotide boundary defines the distal end of the AdML core promoter.

The experiments presented here demonstrate that the early-elongating polymerase requires template DNA that extends 28 nucleotides downstream of its catalytic site, and that this requirement is most likely continuous through the entire phase of promoter escape. Therefore, the distal boundary of required DNA is dynamic, moving further downstream as the polymerase synthesizes the first 14–15 phosphodiester bonds of nascent transcripts. At the last step of promoter escape, downstream DNA extending to positions 40/42 is required.

A particularly intriguing question is the nature of the biochemical mechanism underlying the downstream DNA requirement. The ~28-nucleotide spacing between the position of the polymerase and the distal end of required DNA could be used for physical contact between RNA polymerase or other components of the transcription complex and the DNA template. A likely hypothesis is that downstream DNA is required to bind TFIIH, allowing its DNA helicase subunits to facilitate promoter melting required for initiation and promoter escape. This hypothesis is supported by evidence regarding the direct involvement of the TFIIH helicase in initiation and promoter escape (8, 9, 12, 16, 29) and with cross-linking experiments showing that the TFIIH helicase binds the template downstream to RNA polymerase (27, 28). Recently it was shown that TFIIH action in both transcription initiation and promoter escape requires the presence of specific downstream DNA regions (29).

These findings are consistent with the idea that TFIIH binds to downstream DNA and suggest that it slides along the template ahead of the polymerase until promoter escape is completed. In this light, the arrest of early elongation complexes 28 nt from the distal end of the template can be explained in terms of the inability of the bound TFIIH to physically translocate further downstream, thereby forming a physical impediment to the translocation of the polymerase.

Moreover, our current and recently published results (29) are also consistent with a model by which the conformation of the very early transcription complex remains relatively unchanged through the completion of promoter escape. The downstream spacing required for initiation is very similar to that observed for the promoter escape stage, but not identical, as some inhibition of initiation is seen when the templates are cleaved at promoter positions 32 and 34. This difference might be indicative of a rearrangement shortly after the synthesis of the first phosphodiester bond and before nascent RNA transcripts reach 8–9 nt in length. New information regarding possible changes in composition of the early elongation complex during initial extension of RNA transcripts would be extremely helpful in correlation with these findings.

As it becomes apparent that downstream interactions play an important role in early transcription, it becomes essential to find out whether these interactions reflect a sequence-specific mechanism. To address this question, we have taken a minimalist approach: introducing strong perturbations in template sequence using a relatively small set of substitutions. Because our starting point is a template where there is strong dependence on downstream DNA in early transcription, it is likely that such changes will have an effect if the mechanism is indeed sequence-specific. Clearly our results shown in Fig. 5 do not indicate that. Therefore, it is our conclusion that the downstream DNA requirement in initiation and promoter escape reflects a sequence-independent mechanism. This possibility is
consistent with a sliding model for TFIIH during these steps, indicating a general mechanism that likely functions in other promoters as well.

After carefully characterizing the downstream DNA requirements during promoter escape, an interesting follow-up question was when, during early transcription, does the DNA ahead of the polymerase become fully available as a template? Based on the evidence from cross-linking studies that TFIIH directly contacts the downstream DNA region, this information should help determine when TFIIH has left the early elongation complex. Failure of the polymerase to use the entire downstream template for transcription by the end of the promoter escape stage could indicate a continued association of TFIIH with the transcription complex. We have addressed this question in Fig. 6, and observe no indication for a “fixed spacing” impediment as is the case in the steps that precede the completion of promoter escape. These results are consistent with a model in which TFIIH is occupied at the early elongation complex when the promoter escape stage is complete, i.e. the DNA-helicase activities associated with TFIIH are no longer required.

Fig. 7 is a schematic illustration of a model describing the continuous role of downstream DNA and TFIIH in promoter escape. According to this model, TFIIH slides forward during the transition from initiation through the completion of promoter escape, while occupying a template space downstream of the polymerase and assisting in promoter melting. Because this TFIIH action is continuously required until the completion of promoter escape, removal of downstream DNA, ATP, or TFIIH from the transcription complex at any of these steps is expected to stall the transcription complex. At the ending point of promoter escape, the nascent mRNA is 15 nucleotides in length or larger, and ATP, downstream DNA, and TFIIH are no longer required for its further extension. When transcription is carried out using linear templates, the polymerase can utilize the entire template length for transcription with no spacing required to its distal end. Therefore, it could be expected that the early elongation complex undergo a major conformation change at the completion of promoter escape, probably involving the release of TFIIH and TFIIIE.

Acknowledgments—We give special thanks to Lori Spangler and Kathleen Schmitz-Lesich for extensive help with the graphics and to Michele Ross and Meredith Nusbaum for invaluable help with text editing. We also thank Dr. Zachary Burton, Drs. Joan and Ron Conaway, and Dr. Charles Lindemann for very helpful discussions.

REFERENCES