Strong Natural Pausing by RNA Polymerase II within 10 Bases of Transcription Start May Result in Repeated Slippage and Reextension of the Nascent RNA

Mahadeb Pal and Donal S. Luse*

Department of Molecular Biology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195

Received 25 June 2001/Returned for modification 9 August 2001/Accepted 17 September 2001

We find that immediately following transcript initiation, RNA polymerase II pauses at several locations even in the presence of relatively high (200 μM) levels of nucleoside triphosphates. Strong pauses with half-lives of >30 s were observed at +7, +18/19, and about +25 on the template used in these experiments. We show that the strong pause at +7, after the synthesis of 5’-ACUCUCU, leads to repeated cycles of upstream slippage of the RNA-DNA hybrid followed by re-pairing with the DNA and continued RNA synthesis. The resulting transcripts are 2, 4, and 6 bases longer than predicted by the template sequence. Slippage is efficient when transcription is primed with the +1/+2 (ApC) dinucleotide, and it occurs at even higher levels with the +2/+3 primer (CpU). Slippage can occur at high levels with ATP initiation, but priming with CpA (−1/+1) supports very little slippage. This latter result is not simply an effect of transcript length at the point of pausing. Slippage can also occur with a second template on which the polymerase can be paused after synthesizing ACUCU. Slippage is not reduced by an ATP analog that blocks promoter escape, but it is inhibited by substitution of 5Br-U for U in the RNA. Our results reveal an unexpected flexibility of RNA polymerase II ternary complexes during the very early stage of transcription, and they suggest that initiation at different locations within the same promoter gives rise to transcription complexes with different properties.

An RNA polymerase II transcription cycle consists of several distinct steps: recruitment to assemble the preinitiation complex, open-complex formation, initiation, promoter escape and clearance, transcription elongation, and termination. It is now clear that regulation may be imposed at steps after initiation (reviewed in references 20, 26, 31, and 36). In particular, many transcription units contain newly initiated RNA polymerases that are stalled from 20 to 40 nucleotides downstream of transcription start (reviewed in reference 24). The molecular basis for the regulated pausing and release of RNA polymerase II from promoter-proximal locations has not been established, in part because the transition from the well-characterized preinitiation complex to the stable transcript elongation complex is poorly understood.

Rate-limiting steps early in transcription should be particularly interesting, since they correspond to likely regulatory targets and could identify important structural transitions in the transcription complex. This idea is strongly supported by the recent results of Kugel and Goodrich (18), who placed the rate-limiting step for in vitro transcription from a nonregulated promoter somewhere between the formation of the 4th and 14th bonds. In light of this report, we began the current work by preparing gently washed RNA polymerase II preinitiation complexes, which were then incubated with nonlimiting (i.e., 200 μM) levels of labeled nucleoside triphosphates (NTPs). This experiment revealed that transcription immediately downstream of initiation does not proceed uniformly. Instead, we observed a series of very discrete pauses with half-lives at least 20 times longer than the average bond formation rate under the conditions we employed. The first of these pauses occurs at +7, after the synthesis of the transcript 5’-ACUCU, on the template we are studying. We were surprised to discover that the +7 pause allowed the RNA-DNA hybrid to shift upstream by 2 bp and then re-pair with the template, leading to the synthesis of a transcript 2 bases longer than that predicted by the template sequence. Multiple rounds of this upstream slippage and reextension event can take place.

The transcript slippage reaction we describe here is interesting for a number of reasons. It can occur 7 bases downstream of initiation, at which point the transcription complex has almost completed assembly of the 8 to 9-bp RNA-DNA hybrid (16) that characterizes the final transcript elongation complex. Particularly in light of the tight pairing of the 3’-most bases of the transcript with the template (8), the ability of a 7-mer complex to disrupt the RNA-DNA hybrid and displace it upstream is surprising. Although transcript slippage at our promoter is affected by NTP concentration, it is most sensitive to the exact location at which transcription begins. Slippage is easily observed with the normal initiation substrate (ATP) or with a +1/+2 primer. However, priming with the −1/+1 dinucleotide does not support significant transcript slippage, while priming with the +2/+3 dinucleotide results in over half of the initiations going through at least one round of slippage even at near-physiological (800 μM) NTP levels. Our results with transcript slippage reveal an unexpected flexibility in the early RNA polymerase II transcription complex, and they also suggest that not all potentially available initiation sites give rise to transcription complexes with identical properties.

* Corresponding author. Mailing address: Department of Molecular Biology, NC20, Lerner Research Institute, Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland OH 44195. Phone: (216) 445-7688. Fax: (216) 444-0512. E-mail: lused@ccf.org.
FIG. 1. Sequences of the nontemplate strands of the initially transcribed regions of the plasmids used in these studies. The single G residue which differs from the pML20-40 sequence in the 4G, 6G, and 9G templates is shown in each instance in lowercase type.

MATERIALS AND METHODS

Reagents. Ultrapure NTPs were obtained from Pharmacia Biotechnology. ApC and CpU dinucleotides were obtained from Sigma, [32P]-labeled NTPs (800 Ci/mmol) were obtained from NEN, Deep Vent DNA polymerase and restriction enzymes were obtained from New England Biolabs, RNase T1, RNase II, and RNase III were all obtained from Boehringer Mannheim, and streptavidin-coated magnetic beads were obtained from Promega Biotech.

Plasmids. All plasmids used in this study are based on the adenovirus major late (AdML) promoter. Plasmid pML20-40 is identical to the previously described pML20-42 construct (34) with the exception of a single nucleotide change at position +16 on the nontemplate strand (Fig. 1). Plasmids pML20-40(4G), pML20-40(6G), and pML20-40(9G) were constructed by replacing the BssStuI-HII fragment of pML20-40 with a synthetic DNA fragment carrying a G residue which differs from the pML20-40 sequence in the 4G, 6G, and 9G templates.

Template preparation for in vitro transcription reactions. Template DNAs were made by PCR using the forward primer 5'-GGGCTCAAGGAAAGTTGA TTG-3' (biotinylated at the 5' end) and the reverse primer 5'-CAGTGGCAGGGTTGCTGATG-3' (a HindIII site is underlined). The 190-bp product thus obtained has the transcription start site 96 bp downstream of the biotinylated end. After purification using the Concert cleanup kit (Bethesda Research Laboratories) and digestion with HindIII to reduce transcription arrest before the end of the template (12), the DNA was extracted with phenol-chloroform and precipitated with ethanol.

In vitro transcription reaction on attached templates. Preinitiation complexes were assembled using HeLa cell nuclear extract and biotinylated template DNA immobilized on streptavidin-coated magnetic beads as described previously (29, 34). Initial complex assembly was carried out at either 25 or 30°C for 25 min. Bead-attached complexes were washed twice in 150 μl of BC100 (20 mM Tris [pH 7.9], 8 mM MgCl2, 100 mM KCl, 1 mM dithiothreitol, 20% glycerol, 0.2 mM EDTA), resuspended in BC100, and stored on ice until used. Transcription reaction mixtures contained 1 mM dinucleotide primer (ApC, CpA, or CpU), dATP (usually 20 μM) to satisfy the energy requirement for transcript initiation, and other nucleotides at the concentrations indicated in the figure legends. The labeling nucleotides had specific activities of 800 Ci/mmol unless otherwise indicated. Reaction times were given in the legends. When ATP was used to initiate transcription instead of a dinucleotide, it was present at 100 μM and dATP was absent. Only 5 μM dATP was used for the reactions in Fig. 8B and whenever transcription was primed with CpU. Reactions were stopped by adding an equal volume of phenol-chloroform-isomyl alcohol (25:24:1). The aqueous phase from the initial extraction was reextracted with chloroform-isomyl alcohol (24:1). Reaction mixtures involving synthesis of 15-nucleotide or longer RNAs were usually ethanol precipitated before analysis. For shorter RNAs, the aqueous phase was lyophilized, resuspended in water, and treated with calf intestinal phosphatase (0.1 U/ml for 10 min at 30°C). RNase T1 digestions were performed in 10 μl at 37°C for 10 min with 1 or 2 U of T1 and 0.25 μg of yeast tRNA as carrier. RNAs were resuspended either in 50% formamide or 8 M urea before resolution on 25% acrylamide-3% bisacrylamide gels containing 7 M urea. Gels were imaged, and individual bands were quantified using a PhosphorImager and ImageQuant software (Molecular Dynamics). RNA length markers were made by partial resection of the transcript of a particular halted complex in the presence of 2 mM sodium pyrophosphate at 37°C for 2, 4, or 8 min.

RESULTS

RNA polymerase II exhibits distinct kinetic pauses in the early stage of RNA synthesis. The template we chose for our
study, called pML20-40 (Fig. 1), is based on the AdML promoter. In our earlier work we had used similar templates to purify and characterize transcription complexes stalled at various locations from 13 nt downstream of the transcription start (23, 33, 34). Our most recent studies revealed unexpectedly strong pausing when Sarkosyl-rinsed transcription complexes with RNAs as short as 15 nt were allowed to resume transcript elongation (29; see also reference 18). We were therefore interested in determining what other pauses might occur during the first steps in transcription on the AdML promoter. We assembled preinitiation complexes with HeLa nuclear extracts on bead-attached pML20-40 templates, rinsed them gently with transcription buffer (without Sarkosyl) and then initiated RNA synthesis with 1 mM ApC primer and 200 μM each NTP. This concentration of NTPs supports transcript elongation at 1 to 2 nt/s at 30°C with promoter-distal RNA polymerase II transcription complexes (29). The transcripts were uniformly labeled by including both [α-32P]CTP and [α-32P]UTP in the reaction mixtures. Total RNA was isolated from 5 to 300-s time points and resolved on a high-percentage polyacrylamide gel. Since all of the NTPs were present at the same relatively high concentration, one might have expected to observe a zone of transcripts moving uniformly up the gel. However, as shown in Fig. 2A, most of the RNA appeared at a series of very discrete pause sites, most prominently at +7, +18/19, and roughly +25. A less pronounced pause occurred at +13. We repeated the experiment in Fig. 2A with limiting NTP concentrations (1 mM ApC, 2 μM [α-32P]UTP and 20 μM each CTP and GTP); in this case ATP was left out to stall transcription at +20 (Fig. 2B). With the NTP levels reduced by either 10- or 100-fold relative to the reaction in Fig. 2A, the residence time at the pause sites in Fig. 2B was considerably longer but the locations of the major pauses were unchanged. These results strongly suggest that RNA polymerase II is undergoing a series of relatively slow conformational transitions during the synthesis of the first 25 bases of the transcript. We will return to this point in Discussion.

Transcript slippage by RNA polymerase II. A surprising outcome of these experiments can be seen in Fig. 2B, particularly in the 150- and 300-s lanes. In these reactions, no RNAs longer than 20 nt should have been made. However, a pair of bands reproducibly appeared above the 20-mer, whose length we verified with size markers (see Fig. 3A). The RNAs longer than 20 nt should not have resulted simply from readthrough of the A-stop at +20, since we saw no such readthrough on other, comparable templates at similar nucleotide concentrations (data not shown here, but see for example Fig. 4). When we ran RNAs from reactions like those in Fig. 2B along with length markers, it was apparent that the unexpected RNAs were 2, 4, and 6 nt longer than the 20-mer (see asterisk bands in Fig. 3A) whereas readthrough due to very low levels of contaminating ATP should have resulted in production of low levels of 21-nt and even lower levels of 22-nt RNAs (Fig. 1). Note that transcripts from the pML20-40 template begin

FIG. 3. Transcript slippage on the pML20-40 template. (A) In lanes 1 to 3, RNA was synthesized with 1 mM ApC, 2 μM [α-32P]UTP, and 20 μM each CTP, GTP, and dATP at 30°C for 4 min; 20 μM nonlabeled UTP was then added, and the reaction was continued for 4 min. Transcripts were purified (lane 1) and then digested with 1 U (lane 2) or 2 U (lane 3) of RNase T1 per 10 μl of reaction mixture. Length standards were generated by pyrophosphorolysis (PPI) of transcripts halted at C27 or U20 on template pML20-40 (4G). The single 4-nt T1 cleavage product (5’-CUUU-3’) from the 3’ termini of the RNAs is indicated (lanes 2 and 3). The slippage-reextension products are marked by asterisks. (B) Schematic of the first and second round of transcript slippage and reextension on the pML20-40 template. TS, template strand; base pairing of the transcripts before and after slipping is indicated.
5’-ACUCUCU (Fig. 1). This suggested to us that the unanticipated 22-nt RNA in Fig. 2B could result from upstream slippage of the initial RNA-DNA hybrid by 2 nt followed by re-pairing of the RNA with the template and resumption of RNA synthesis. The 24-nt RNA could result from two cycles of this process, and so forth. This hypothesis is diagrammed in Fig. 3B. We propose position +7 as the most likely location for slippage to begin, for two reasons. First, there is a strong natural pause at this location during transcription (Fig. 2), and second, this position marks the end of the three CU repeat elements and thus offers the maximum number of bases hybridized at the 3’ end (four) if slippage takes place.

If hybrid slipping and reextension occurred as diagrammed in Fig. 3B, cleavage of the transcripts in Fig. 3A with RNase T1 should lead to the production of CUUU from the 3’ ends of all of the RNAs, with bands of 16, 18, 20, and 22 nt generated from the 5’ ends. The results of T1 digestion (Fig. 3A) were consistent with this expectation. Also, if our slippage model is correct, changes in the template that interrupt the CU repeats should prevent stable extended hybrid formation after slippage and thus block the slippage-reextension reaction. We generated three variants of our initial template, which we designated pML20-40(4G), pML20-40(6G), and pML20-40(9G).

As shown in Fig. 1, each template differed from the original by a single C-to-G change on the nontemplate strand. None of these changes alter the strength of the RNA-DNA hybrid compared to the original template. We confirmed (data not shown) that all of these templates support the identical pausing pattern during early RNA synthesis, including the prominent pause at +7, which we observed with the pML20-40 template. Note that when the polymerase pauses at +7 on either the 4G or 6G template, the RNA-DNA hybrid cannot slip 2 nt upstream while retaining 4 bp at the 3’ end. However, slippage from the 7-mer should be unaffected on the 9G template.

When we tested these templates under reaction conditions like those used in the experiment in Fig. 3, we obtained the expected results. RNAs longer than 20 nt were produced as efficiently on the 9G template as on pML20-40, but these RNAs were obtained at a much lower level on the 4G and 6G templates (Fig. 4). RNase T1 digestion of the RNA produced on the 9G template gave, in addition to the expected 4-, 7-, and 9-nt products, RNAs of 11 and 13 nt, consistent with two rounds of slippage and re-pairing as diagrammed in Fig. 3B.

We observed a strong pause at +7 even under free-running transcription conditions with 200 μM NTPs (Fig. 2A). However, the detailed characterization of the slippage reaction was done with more limiting NTP concentrations, particularly for the 32P-labeled NTP. To explore the dependence of transcript slippage on transcription conditions, we transcribed the pML20-40 and 9G templates with ApC primers and different concentrations of CTP, UTP, and GTP, ranging from the labeling NTP at 1 μM to all NTPs at 200 μM. We resolved the RNAs on high-percentage acrylamide gels as in the experiment in Fig. 4 and determined the relative yields of 20-, 22-, 24-, and 26-nt RNA. We also performed ApC-primed transcriptions on pML20-40 with CTP and UTP at 815 μM. In this last case, the expected product was 15 nt long and the slippage products were 17- and 19-mers. The results are shown in Table 1. Not surprisingly, the amounts of slippage-reextension products were larger when transcription was performed at low NTP levels. However, slipped RNAs represented 16% of the total transcript with all NTPs at 200 μM, and this percentage was essentially the same (18%) when 815 μM NTPs were used. Thus, transcript slippage is not an artifact of low NTP concentrations. As expected, slippage-reextension transcripts were made with equal efficiency on the pML20-40 and 9G templates when reaction conditions were identical.

**Effect of the initiating nucleotide and site of initiation on transcript slippage.** A number of earlier studies have reported slippage and resumption of RNA synthesis during transcript initiation by *Escherichia coli* RNA polymerase (2, 35). It was suggested that the lack of slippage for certain transcripts reflected a stabilization of the ternary complex by interaction of
TABLE 1. Relative yields of normal transcripts and slippage-reextension products as a function of NTP concentrationa

<table>
<thead>
<tr>
<th>Template</th>
<th>[NTP] (μM)b</th>
<th>15 nt</th>
<th>17 nt</th>
<th>19 nt</th>
<th>20 nt</th>
<th>22 nt</th>
<th>24 nt</th>
<th>26 nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>pML20-40</td>
<td>20 C, 20 G, (1 + 20) U</td>
<td>57.9 ± 5.3</td>
<td>20.6 ± 0.9</td>
<td>19.3 ± 3.1</td>
<td>3.0 ± 1.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 C, 20 G, U</td>
<td>69.1 ± 6.6</td>
<td>18.3 ± 2.4</td>
<td>11.6 ± 4.6</td>
<td>1.4 ± 0.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>200 C, 200 G, 200 U</td>
<td>83.9 ± 1.7</td>
<td>10.4 ± 1.6</td>
<td>5.3 ± 1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9G</td>
<td>63.6 ± 2.1</td>
<td>20.4 ± 1.2</td>
<td>15.6 ± 0.1</td>
<td>1.6 ± 0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pML20-40</td>
<td>815 C, 815 U</td>
<td>82</td>
<td>8</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Transcription was performed on the indicated templates with 1 mM ApC and the indicated concentrations of CTP, UTP, and GTP at 30°C for 20 min (top four rows) or with 815 μM CTP and UTP for 10 min (last row).

b (1 + 20) U indicates that the reaction mixture contained 1 μM 32P-labeled UTP for 5 min followed by addition of nonlabeled 20 μM UTP and incubation for an additional 15 min.

The intensities of nonslipped (20 or 15 nt) and slipped (22, 24, and 26 nt, or 17 and 19 nt) transcripts were quantified with a PhosphorImager (Molecular Dynamics) and expressed as a percentage of the total transcript of 20 nt and longer. Values were corrected for different numbers of U residues in the various transcripts. Standard deviations for the top four rows were calculated from at least three independent experiments. The experiment in the last row was done only once under these conditions; it was repeated at 25°C with very similar results.

We therefore compared the extent of slippage on the pML20-40 template in the ApC-initiated reaction (priming at +1/1 + 2) with that seen with CpA priming (−1/−1 + 1). The results are shown in Fig. 5. In lanes 1 and 2, RNA was generated with ApC priming, CTP, UTP, and GTP, as in lane 7 of Fig. 4. Slippage products of 22, 24, and (very faint) 26 nt were obtained (Fig. 5, lane 1). These RNAs were shortened by 4 nt, as expected, after treatment with RNase T1 (lane 2). When a similar reaction was done without GTP, the G-stop band at 15 nt and the slippage bands at 17, 19, and (faintly) 21 nt were also observed (lane 3). These latter bands all corresponded to stops before the incorporation of the G at +16, since they were unaffected by RNase T1 (lanes 3 and 4). Interestingly, when G-less reactions were carried out using a CpA primer, much less slippage was observed (lane 5). Since we have hypothesized that slippage with the ApC primer depends on the strong pause at +7, at the end of the three CU repeats, it was important to establish that this pause also occurred with CpA-mediated initiation. Kinetic assays (as in Fig. 2B) demonstrated that for CpA initiations, a pause of similar duration occurred at the same template location as seen with ApC priming (data not shown). That is, an 8-nt RNA transiently accumulated in the CpA-primed reactions.

We also assayed transcription of the pML20-40 template without dinucleotide primers, using ATP, CTP, and UTP to support RNA synthesis. This resulted in detectable slippage products, but the level of slippage was lower than we observed with ApC priming. However, when we performed kinetic experiments like those in Fig. 2 with ATP-initiated reactions, we found that the pause at +7 was significantly longer (data not shown). We therefore attribute the lower level of slippage in ATP-initiated reactions, relative to similar reactions primed with ApC, to the shorter residence time at the critical +7 position. On other templates which permitted pause times to be manipulated, high levels of slippage after initiation with ATP could be observed (see Fig. 7), but we were unable under any experimental conditions to obtain high levels of slippage in CpA-primed reactions (see Fig. 7).

Transcript slippage may occur upstream of +7 if pausing is induced at the upstream location. As noted above, our results are consistent with a model in which the transcription complex paused at +7 is the primary precursor of transcript slippage.
after upstream slippage for transcription to resume may explain the low level of slippage products we see with the 6G template (Fig. 4). Even though there is no discernable pause at +5, the residence time at this position may be long enough for some slippage to take place. Upstream slippage by 2 bases from the normal pause site at +7 seems less likely for 6G transcripts, since this would result in an attempt to pair two G residues. On the other hand, slipping upstream from the pause at +7 may account for the low level of slippage we observed with the 4G template. In this case, translocation upstream by 2 bases would result in the pairing of UCU in the RNA with ACA in the DNA. Only 2 bp are formed, but the middle position in this case would not result in a potential purine-purine clash.

The results in Fig. 6 indicated that only two of the CU repeats need be present at the 3' end of the transcript at the point of pausing for slippage to take place. This finding allowed us to test another dinucleotide primer, CpU, for its ability to support slippage. CpU can prime transcription on our templates from two locations, −3/−2 and +2/+3 (Fig. 1). Preinitiation complexes on the 6G template incubated with CpU, CTP, and radiolabeled UTP should give rise to only one labeled product, the CUCU tetramer initiated from the +2/+3 position (Fig. 1). Note that this RNA could slip upstream by 2 nt and re-pair in a manner analogous to that just described for ApC initiation in the absence of GTP on the 6G template. As shown in Fig. 7A, CpU initiation from +2/+3 yielded not only the expected 4-nt RNA but also RNAs that are 2, 4, 6, and 8, and, at a low level, even 10 nt longer. None of these RNAs represented readthrough of the G stop, since none of them were sensitive to RNase T1 (lanes 7 and 8). High levels of the slippage-reextension products did not accumulate until the reaction had proceeded for at least 40 s. Most of the slippage products remained in ternary complex after 4 min because they could be chased into longer RNAs (lane 9). Thus, the slippage reaction not only occurs with a +2/+3 primer but is even more efficient than with the +1/+2 primer, ApC.

The ability to stall the RNA polymerase indefinitely at +5 on the 6G template allowed us to assay for slippage with all of the initiation substrates under comparable reaction conditions (Figs. 7B and C). In Fig. 7B, each 2-min reaction mixture contained CTP and radiolabeled UTP in addition to the initiating nucleotide or dinucleotide. Transcripts attributable to multiple rounds of slippage were observed with both ApC and CpU priming, but not with CpA priming, as shown above. (Note that the CpA-primed 6-mer is the normal G-stop transcript and not a slippage product.) Interestingly, when the RNA polymerase was forced to pause at +5 for an extended period, products consistent with at least three rounds of slippage could also be observed in ATP-initiated reactions (lanes 7 and 8). All of the RNAs marked with asterisks in lanes 1 to 8 of Fig. 7B should represent slippage products and not readthrough of the G stop, since these RNAs were not sensitive to RNase T1. If readthrough had taken place, the resulting transcripts should have accumulated at +15, before the next G stop. Low levels of such RNAs were observed in the ATP and ApC-initiated reactions (lanes 3 and 7); these transcripts were sensitive to RNase T1 (lanes 4 and 8). As a second positive control for activity of the T1, a parallel ApC-primed reaction

The results in Fig. 6 indicated that only two of the CU repeats need be present at the 3' end of the transcript at the point of pausing for slippage to take place. This finding allowed us to test another dinucleotide primer, CpU, for its ability to support slippage. CpU can prime transcription on our templates from two locations, −3/−2 and +2/+3 (Fig. 1). Preinitiation complexes on the 6G template incubated with CpU, CTP, and radiolabeled UTP should give rise to only one labeled product, the CUCU tetramer initiated from the +2/+3 position (Fig. 1). Note that this RNA could slip upstream by 2 nt and re-pair in a manner analogous to that just described for ApC initiation in the absence of GTP on the 6G template. As shown in Fig. 7A, CpU initiation from +2/+3 yielded not only the expected 4-nt RNA but also RNAs that are 2, 4, 6, and 8, and, at a low level, even 10 nt longer. None of these RNAs represented readthrough of the G stop, since none of them were sensitive to RNase T1 (lanes 7 and 8). High levels of the slippage-reextension products did not accumulate until the reaction had proceeded for at least 40 s. Most of the slippage products remained in ternary complex after 4 min because they could be chased into longer RNAs (lane 9). Thus, the slippage reaction not only occurs with a +2/+3 primer but is even more efficient than with the +1/+2 primer, ApC.

The ability to stall the RNA polymerase indefinitely at +5 on the 6G template allowed us to assay for slippage with all of the initiation substrates under comparable reaction conditions (Figs. 7B and C). In Fig. 7B, each 2-min reaction mixture contained CTP and radiolabeled UTP in addition to the initiating nucleotide or dinucleotide. Transcripts attributable to multiple rounds of slippage were observed with both ApC and CpU priming, but not with CpA priming, as shown above. (Note that the CpA-primed 6-mer is the normal G-stop transcript and not a slippage product.) Interestingly, when the RNA polymerase was forced to pause at +5 for an extended period, products consistent with at least three rounds of slippage could also be observed in ATP-initiated reactions (lanes 7 and 8). All of the RNAs marked with asterisks in lanes 1 to 8 of Fig. 7B should represent slippage products and not readthrough of the G stop, since these RNAs were not sensitive to RNase T1. If readthrough had taken place, the resulting transcripts should have accumulated at +15, before the next G stop. Low levels of such RNAs were observed in the ATP and ApC-initiated reactions (lanes 3 and 7); these transcripts were sensitive to RNase T1 (lanes 4 and 8). As a second positive control for activity of the T1, a parallel ApC-primed reaction
was run which did include GTP. These transcripts gave the expected fragments on treatment with T1 (lanes 9 and 10). In light of the efficient slippage supported by the CpU primer, we decided to test for slippage in CpU-initiated reactions with high levels of NTPs. For these experiments, we transcribed the pML20-40 template with 1 mM CpU, 2 μM [α-32P]UTP, 5 μM dATP, and 20 μM CTP; reactions were stopped at the indicated times (lanes 3 to 8). Reactions in lanes 1 and 2 were done without the CpU primer. An aliquot of the 240-s reaction mixture was digested with RNase T1 (lane 8). Another aliquot of the 240-s reaction mixture was chased with 200 μM each NTP (lane 9); RNA from an equivalent reaction mixture was digested with T1 (lane 10). Slippage products are marked with asterisks, and the lengths of these RNAs are indicated to the left of the gel. (B) Transcription on the 6G template was initiated either with 1 mM dinucleotide or with 0.1 mM ATP, as indicated on the figure, along with 2 μM [α-32P]UTP, 20 μM CTP, and, in dinucleotide-containing reactions, 5 μM dATP, at 30°C for 2 min. One-half of each reaction mixture was digested with RNase T1 as indicated. For lanes 9 and 10, an ApC-primed reaction mixture otherwise identical to that in lane 3 was supplemented with 20 μM GTP; one-half of this reaction mixture was also digested with T1, as indicated (lane 10). Slippage products are marked by asterisks; transcripts resulting from readthrough of the G stop are marked by #. The lengths of the nonslipped G-stop RNAs for the CpA-, ApC-, and ATP-initiated reactions are marked on the gel. (C) The pML20-40 template was transcribed with 1 mM dinucleotide or 1 mM ATP plus 800 μM each CTP and UTP and 15 μM each [α-32P]CTP and [α-32P]UTP at 30°C for 10 min. Reaction products were ethanol precipitated, and the indicated samples were digested with RNase T1. Slippage products and RNAs resulting from readthrough of the G stop are marked by asterisks and #, respectively.

Roles of RNA-DNA hybrid strength and helicase activity in transcript slippage. Slippage requires the RNA-DNA hybrid to be transiently disrupted in order to allow re-pairing at a more upstream location. This predicts that strengthening the RNA-DNA hybrid would inhibit the slippage reaction. To test this, we substituted 5Br-UTP for UTP in ApC-primed transcription reactions with the pML20-40 template. As shown in Fig. 8A, the slippage products longer than 20 nt observed with 5Br-UTP were completely absent (even at long exposure) when 5Br-UTP was substituted. The presence of 5Br-U also reduced the pause time at +7 compared to that in the UTP reaction, but the 40 to 60-s residence time at this location was sufficient to support easily detectable slippage in other experiments (data not shown) using UTP.

The location from which slippage initiates, as far downstream as +7, raises the question whether the TFIIH initiation factor might be involved in this process, since it has been established that inhibiting the XPB helicase activity of TFIIH with ATPγS blocks promoter escape (5–7) and inhibits the accumulation of transcripts longer than 9 or 10 nt. Earlier work (38) indicated that TFIIH should still be present in our early elongation complexes, which are substrates for slippage. We performed a series of transcription reactions on the 6G tem-
plate with an ApC primer, CTP, and radiolabeled UTP, plus 5 μM dATP to support promoter unwinding. After 40 s, the expected 5-mer had accumulated, plus very low levels of slippage products (Fig. 8B, lane 1). When transcription was continued for 2 min without any additions (lanes 2 and 3), 7-, 9-, and 11-nt RNAs accumulated, as expected from the results in Fig. 7B. Also as expected from the results in Fig. 4, continuation of the 40-s reaction in the presence of GTP (lanes 6 and 7) led primarily to the accumulation of 20-nt RNA with only low levels of slippage. When ATP was added at 30 s along with 7, 9-, and 11-nt RNAs accumulated, as expected from the results in Fig. 7B.

DISCUSSION

We have studied the kinetics of the early postinitiation phase of transcription by native RNA polymerase II complexes. For the promoter used in this study, strong pauses were seen at +7, +18/19, and about +25, with a weaker stop at +13 (Fig. 2A). The average rate of bond formation by RNA polymerase II was 1 to 2 nt/s under the conditions used for the experiment in Fig. 2A, but the polymerase paused for 30 s or more at the three major promoter-proximal pause sites. Note that in the Fig. 2A experiment all of the NTPs were present at 200 μM, which is at or close to the $K_v$ values determined for the four NTPs when calf thymus RNA polymerase II transcribes dC-tailed templates (180 μM for CTP, 200 μM for GTP, 400 μM for ATP, and 480 μM for UTP [15]). This suggests that these pauses are not simply the result of NTP limitation but instead mark important structural transitions within the early postinitiation transcription complex. It is useful to consider what these transitions might be, as an introduction to the discussion of the remarkable capacity for newly initiated RNA polymerase II complexes to allow upstream slipping and re-pairing of the transcript with the DNA template.

Cross-linking and biochemical studies have identified three essential features of the multisubunit RNA polymerases responsible for the continued elongation competence of ternary complexes: an RNA-DNA hybrid of about 8 bp (16, 28), the interaction of RNA immediately upstream of the hybrid with the RNA-DNA hybrid (19, 26, 27). RNase protection studies with RNA polymerase II showed that 17 to 19 nt of transcript resists digestion in ternary elongation complexes (9), consistent with the presence of 8 to 9 nt of RNA in the RNA exit channel (1, 26, 37), and the interaction of the polymerase with about 9 bp of double-stranded DNA downstream of the transcription bubble (19, 26, 27). RNase protection studies with RNA polymerase II showed that 17 to 19 nt of transcript resists digestion in ternary elongation complexes (9), consistent with the presence of 8 to 9 nt of RNA in the RNA exit channel immediately upstream of the hybrid with an RNA binding site on the RNA polymerase (17, 39) and yeast RNA polymerase II (8).

In light of these findings, it is tempting to suggest that the +7 pause is related to the breaking of the RNA-DNA hybrid and
the initial threading of the transcript into the exit channel. The RNA-DNA hybrid reported for the yeast RNA polymerase II ternary complex is 9 bp long (8), but it has been suggested that this complex is in a paused configuration (21). Perhaps the approach of the upstream end of the growing RNA-DNA hybrid to the "rudder," which is thought to be at least partially responsible for separating the transcript from the template (8), provides some resistance to translocation along the template. Such resistance could be particularly deleterious to elongation when consecutive U residues are being added to the transcript, as is the case at +7 on the pML20-40 template. This explanation is consistent with the reduction in the +7 pause which occurred when 5Br-U was substituted for U (Fig. 8). The +18/19 pause also corresponds to a location in which consecutive U residues are being added to the growing RNA chain, and this pause is also reduced in the presence of 5Br-U (Fig. 8). As with the +7 pause, there may be some resistance to continuing transcription past +18/+19 which is exaggerated by the formation of a weak hybrid at the 3' end of the transcript. A transcript length of 18 to 19 nt marks the point at which the RNA exit channel should be stimulated by high levels of the NTP incorporated in the polymerase II ternary complex can translocate, or slip, upstream by 2 bases, reanneal to the DNA template, and support continued transcript elongation. Although there have been many reports of transcript slippage during the early stages of transcription, we would emphasize that the phenomenon we report here is mechanistically unique. For E. coli RNA polymerase, slippage has been observed beginning at +4 to +6, at the end of a stretch of either U or A residues in the transcript, leading to the introduction of long runs of U or A in the RNA (10, 13, 14, 25). This prevented the synthesis of longer, correctly templated RNAs at the pyrB1 promoter (25). A second class of slippage events includes cases in which 3-nt (2) or 4-nt (35) initiation transcripts slipped upstream by 3 bases, re-paired with the template, and supported continued transcription. None of the examples just cited correspond to the process we have documented with RNA polymerase II. The generation of long A or U polymers (10, 25) is presumed to occur in a stuttering mode that requires only a single-base upstream slippage at the 3' end of the transcript and thus does not entail large-scale disruption of the RNA-DNA hybrid as an initiating event. Unlike some of the stuttering reactions, slippage on pML20-40 does not remove RNA polymerase II from the normal mode of productive transcription. The 22-, 24-, and 26-nt RNAs produced by slippage on pML20-40 in the absence of ATP were as likely to be elongated in subsequent chase reactions as were the RNAs which had not slipped (data not shown in this case [see also Fig. 7]). In one case, the stuttering reaction was found to be stimulated by high levels of the NTP incorporated in the polymer (25), whereas in our case, slippage is most pronounced at low NTP concentrations (Table 1). The examples with upstream slippage of short oligonucleotides (2, 35) more closely resemble the reaction we report here. However, the longest RNA which was translocated by slippage was 4 nt, and neither of these events could be observed with all NTPs present. Slippage in these cases required that the polymerase be halted at the appropriate point early in transcription by leaving some NTPs out of the reaction mixture. In contrast, slippage in our case is easily observed even with NTPs at 815 mM (Fig. 7C). With CpU priming, the majority of initiations undergo at least one round of slippage at these high NTP concentrations (Fig. 7C).

Slippage by RNA polymerase II during promoter-directed initiation has been reported in only one other instance of which we are aware. Cowie et al. (3) showed that the sequences of a significant proportion of the 5' ends of polyomavirus early transcripts generated either in vivo or in vitro could not be explained without invoking upstream slippage and re-pairing during the synthesis of the initial three bonds. As in our case, this reaction apparently proceeds at physiological NTP concentrations. However, as with the E. coli RNA polymerase examples cited above, the transcript which undergoes slippage is much shorter than the 7-mer which can be the precursor to slippage in our case. In our view, the ability to reposition a 7-bp RNA-DNA hybrid by 2 bases is quite unanticipated in light of the detailed structure of the yeast RNA polymerase II ternary complex bearing a 15-nt transcript (8). In that structure, the 3-bp sequence at the 3' end of the RNA-DNA hybrid is in a tight heteroduplex and is in close contact with the RNA polymerase. It is difficult to imagine how this intimate association could be disrupted to allow displacement upstream. However, it is important to note that the yeast ternary complex contains a 15-nt RNA and thus has passed through the promoter escape transition, while complexes paused in our experiments at +7 would not have escaped the promoter (reference 5 and references therein). Our laboratory had attempted to induce slippage at promoter-distal locations by introducing the sequence (AC)10 into the template strand downstream of one of our AdML promoter-based constructs. We transcribed this sequence with very limiting ATP and CTP levels, and we also caused RNA polymerase to resume transcription within this sequence after upstream translocation from an arrest site. In no case did we obtain any evidence of transcript slippage, in experiments where 1% slippage should have been detected (32). All of these results, taken together, suggest that prior to escape, the clamp domain may not be tightly closed over the RNA-DNA hybrid. This more open configuration would allow the flexibility needed for slippage (8). This idea can be tested by moving the repetitive element in our template further downstream from transcription start, with the expectation that at some point before +15 this element will no longer cause upstream slippage.

A particularly interesting and unanticipated aspect of the slippage reaction is its dependence on the site of transcript initiation. Initiation with the CpA primer (−1/+1) leads to almost no slippage, while initiation at +1 (ATP) or with ApC (+1/+2) supports slippage. Slippage occurs even more efficiently when transcription is primed with CpU (+2/+3). One might argue that CpA-initiated transcripts, which are 8 bases
long at the +7 stop, are long enough to be stabilized (for example, because the RNA-DNA hybrid is complete) and thus resist slippage, in contrast to the 7-nt ApC-initiated transcripts. However, recall that CpA-initiated transcripts made in the absence of GTP on the 6G template, which are only 6 nt long, support almost no slippage (Fig. 7B). These observations suggest that the newly initiated transcription complex is qualitatively different when initiation is primed with CpA from when it is primed with ApC or CpU. A mechanistic basis for this difference may lie in the relative positions of the 5′ ends of the transcripts and the transcription bubble. Holstege et al. have shown that the initial melted region for an AdML-based promoter very similar in sequence to the one used here is -9 to +2 (11). The leading edge of the transcription bubble advances in synchrony with transcription, but the upstream edge remains fixed, until the 10th bond is made. At this point, the upstream edge abruptly shifts to +3. Since the upstream edge of the bubble does not move during the synthesis of the transcripts which are precursors to slippage, it seems possible that different locations of the 5′ end of the RNA relative to the fixed upstream bubble edge could result in somewhat different structures for the various transcription complexes and thus lead to different stabilities of the RNA-DNA hybrids within these complexes. This idea is reinforced by the observation that the closure of the clamp structure of RNA polymerase, which must move inward to retain the nucleic acids during transcript elongation, depends on the interaction of so-called switch regions at the base of the clamp with the RNA-DNA hybrid (8, 21).

An idea put forward in earlier work on transcript slippage by E. coli RNA polymerase is that slippage should be inhibited once the transcript is long enough to reach an upstream RNA binding site (2, 35). Current models for RNA polymerase II would place the entry of the transcript into the exit channel at some point immediately upstream of the end of the RNA-DNA hybrid, i.e., at 10 to 11 nt from of the point of bond formation (8, 21). In the case of ApC-primed transcription of the pML20-40 plasmid, two rounds of slippage occur with relatively high efficiency while much lower levels of third- and fourth-round products can be observed (Fig. 4). The precursors for the first and second rounds of slippage are 7- and 9-mers, respectively (schematic in Fig. 3), while the much less prevalent subsequent rounds would begin with 11- and 13-mers. This would be consistent with the idea that transcripts longer than 9 nt could begin to fill the exit channel and thus be stabilized against slippage. However, it could also simply be a kinetic effect, since the polymerases in the reaction in Fig. 4 were free to continue transcription past the pause site at template position +7. In the reaction on the 6G template in Fig. 7A, where RNA polymerase was forced to stall after synthesizing the transcript CUCU, five rounds of slippage and reextension can be detected. The precursors to the last two rounds were 10 and 12 nt long, respectively. It should be possible to extend the reaction shown in Fig. 7A to much longer times, to determine if there is a maximum transcript length which can lead to slippage.

We cannot say that the path that RNA takes upstream of the RNA-DNA hybrid in complexes which have undergone multiple rounds of slippage is identical to that which would be taken in complexes where no slippage had occurred. However, we emphasize that the large majority of the slipped and reextended RNAs in the CpU-primed transcription experiment, even those which have gone through four or five rounds of slippage, appear to remain in active transcription complexes, since they chase with high efficiency (Fig. 7A).

In summary, we have shown that the process of early transcript elongation by RNA polymerase II is nonuniform. There are a number of extended pauses during the synthesis of the first 25 bases of the transcript, which probably reflects the structural transitions that the polymerase must make as it assumes the transcription elongation configuration. A particularly surprising consequence of the earliest of these pauses on the template we used is the ability of a 7-nt transcript to slip upstream, reanneal with the template, and be efficiently reextended by the RNA polymerase. Our studies cannot address the issue of whether slippage occurs in vivo from the AdML promoter, although easily detectable levels of slippage did take place in ATP-initiated reactions at near-physiological NTP levels (Fig. 7C). It is interesting to recall that slippage at the poliovirus early promoter was discovered by examination of viral RNA (3). Regardless of whether slippage plays an important role during initiation in vivo, the existence of this reaction provides useful insight into the mechanism by which RNA polymerase II passes from initiation into the stable elongation state.

ACKNOWLEDGMENT

This work was supported by grant GM 29487 from the National Institutes of Health.

REFERENCES


Vol. 22, 2002 TRANSCRIPT SLIPPAGE AFTER INITIATION BY RNA Pol II 39