Initially Transcribed Sequences Strongly Affect the Extent of Abortive Initiation by RNA Polymerase II*

(Received for publication, August 7, 1998, and in revised form, January 12, 1999)

Richard G. Keene and Donal S. Luse‡

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

We investigated transcript initiation and early elongation by RNA polymerase II using templates mismatched between −9 and +3 (bubble templates). Highly purified RNA polymerase II alone was able to initiate transcription specifically on these templates in the presence of dinucleotide primers. The length distribution of abortively initiated RNAs was similar for purified RNA polymerase II on bubble templates and polymerase II on double-stranded templates in HeLa nuclear extracts. Increasing the U content in the initial portion of the transcript caused similar increases in abortive initiation for transcription of bubble templates by pure polymerase and double-stranded templates in extracts. Thus, the level of abortive initiation by RNA polymerase II is at least partly determined by interactions of the polymerase with the transcript and/or the template, independent of the general transcription factors. Substitution of 5-bromo-UTP for UTP reduced abortive initiation on bubble templates, consistent with the idea that transcription complex stability during early elongation depends on the strength of the initial RNA-DNA hybrid. Interestingly, transcription of bubble templates in HeLa extracts gave very high levels of abortive initiation, suggesting that inability to reanneal the initially melted template segment inhibits transcript elongation in the presence of the initiation factors.

Both Escherichia coli RNA polymerase (reviewed in Refs. 1 and 2) and RNA polymerase II (3, 4) pass through a stage of abortive initiation between the formation of the initial phosphodiester bond and the establishment of a stable ternary transcription complex. The RNA chains released during abortive initiation by either RNA polymerase are generally from 2 to 10 nt (see, for example, Refs. 4–6), although aborted transcripts as long as 15 nt have been reported (7). The partitioning of initiations between abortive and productive pathways varies among promoters (e.g. see Refs. 7–10), but the mechanistic basis for this difference is not well understood. A number of recent results have emphasized the importance of the strength of the RNA-DNA hybrid in maintaining stability for transcription complexes during RNA chain elongation (Refs. 11–16; reviewed in Ref. 17). Thus, it is possible that the transcript template hybrid is a major factor in determining the relative levels of abortive initiation among promoters. However, the study of Rice et al. (18), which involved RNAse digestion of RNA polymerase II ternary complexes, indicated that the RNA-DNA hybrid within the transcription complex is very short (≤3 base pairs) and not a major factor in transcription complex stability. The RNA polymerase II initiation process on linear DNA templates requires, at minimum, TBP, TFII B, TFIIF, TFII E, and TFII H in addition to RNA polymerase II (reviewed in Ref. 19). The interaction of these general transcription factors with the template and/or the transcript could also strongly influence the abortive initiation process, and indeed, direct effects of TFII E and TFII H on the initiation/elongation transition have been reported (20–22).

In order to further analyze the role of template sequence in abortive initiation, we have extended the observation that templates mismatched in the vicinity of the transcription start site (“bubble templates”) support specific transcript initiation by RNA polymerase II with the addition of only TBP and TFII B (Ref. 23; see also Refs. 24–26). We found that priming RNA synthesis with an appropriate dinucleotide allows transcript initiation on bubble templates by RNA polymerase II alone. We have therefore been able to examine abortive and productive transcript initiation in the complete absence of the general transcription factors. These studies have allowed us to conclude that the level of abortive initiation by RNA polymerase II is at least partly determined by interactions of the polymerase with the transcript and/or the template, independent of transcription factors. Our results are consistent with a model in which the strength of the RNA-DNA hybrid is an important component of transcription complex stability during the transition from initiation to early elongation.

EXPERIMENTAL PROCEDURES

RNA Polymerase II Purification—Calf thymus was obtained frozen from ANTEC (Tyler, TX). RNA polymerase IIA was purified through the DE-52 chromatography step essentially as described by Hodo & Blatti (27), followed by affinity purification on an anti-RNA polymerase II C-terminal domain antibody column as described by Thompson et al. (28). Briefly, 1.5 kg of thymus was ground and clarified in a 4-liter volume; nucleic acid and associated proteins were precipitated by the addition of polyethyleneimine to 0.05%. RNA polymerase II was extracted from the polyethyleneimine pellet, precipitated with ammonium sulfate, and resuspended in sufficient buffer D (50 mM Tris, pH 7.9, 25% glycerol, 2 mM EDTA, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride) to reduce the ammonium sulfate concentration to 150 mM prior to batch adsorption at 20 mg of protein/ml on DE52 (Whatman) for 1.5 h. The DE52 was washed with Buffer D containing 150 mM ammonium sulfate until the absorbance declined to base line and then step-eluted with Buffer D containing 0.5 mM ammonium sulfate. Peak fractions containing polymerase II activity were pooled, concentrated by ammonium sulfate precipitation, and adsorbed to 5WG16 antibody immobilized to Sepharose (28). The antibody column was eluted four times with 1 volume of 40% glycerol, 0.5 mM ammonium sulfate (28); the last elution, done for 12 h at 4 °C, provided all of the RNA polymerase II used in this study. Polymerase pools were concentrated in Centriprep 30 concentrators (Amicon) and dialyzed to 50 mM

* This work was supported by National Institutes of Health Grant GM29487. 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 1754 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Molecular Biology, NC20, Lerner Research Institute, Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland, OH 44195. Tel.: 216-445-7688; Fax: 216-444-0512; E-mail: luse@cesmtp.ccf.org.

The abbreviations used are: nt, nucleotide(s); TBP, TATA box-binding protein; TFII, general factor for transcript initiation by RNA polymerase II; AdMLP, adenovirus major late promoter; AdUMut, variant of the AdMLP bubble template; Br-UTP, 5-bromo-UTP.

This paper is available on line at http://www.jbc.org

11526
Tris-HCl, pH 7.9, 25% glycerol, 150 mM ammonium sulfate, 0.2 mM EDTA, and 1 mM dithiothreitol prior to storage at −80 °C. The RNA polymerase II was >95% pure and essentially all in the IIA form as assayed by SDS-PAGE. The polymerase was free of detectable RNase activity.

**Templates for in Vitro Transcription**—To form bubble templates, 86-nt oligos (from Operon Technologies; see Fig. 1) were mixed at 2 mg/ml total DNA concentration with a 10% molar excess of the non-template DNA strand in 25 mM Tris, pH 7.9, 8 mM MgCl2, 50 mM KCl. Templates were reannealed by heating to 99 °C for 3 min followed by cooling, first to 90 °C for 10 min and then to room temperature at 0.2 °C/min. Reannealed templates were resuspended on 12% non-denaturing polyacrylamide gels and eluted from gel slices by diffusion in 0.5 M ammonium acetate, 0.5% SDS, and 2 mM EDTA followed by ethanol precipitation. To remove acrylamide and other impurities, the templates were gel-filtered through a Sephadex G-50 spin column and ethanol-precipitated prior to use.

**In Vitro Transcription with Pure RNA Polymerase II**—Transcription conditions were similar to those of Pan & Greenblatt (23). Bubble duplex templates (0.22 pmol) were incubated with 0.18 pmol of highly purified RNA polymerase II in a 10-μl reaction volume in transcription buffer (10 mM Tris-HCl, pH 7.9, 10% glycerol, 8 mM MgCl2, 75 mM KCl, and 5 mM β-mercaptoethanol) supplemented with 200 μM μg/ml acetylated bovine serum albumin. After 30-min preincubations at 37 °C, transcription was initiated by adding 2.5 mM of a nucleotide mix. Final nucleotide concentrations in the reactions were 1 mM of dinucleotide primer and, unless otherwise indicated, 10 μM each of CTP and UTP; in some reactions, dATP was also present at 10 μM. In most experiments, both CTP and UTP were 32P-labeled (NEN; specific activity in the reaction 10,000 Ci/mmol). For sarkosyl and heparin challenge experiments, 0.5 μl of water, 1.3% sarkosyl, or 2.6 mg/ml heparin was added to the 10-μl preincubation reactions 30 s prior to the addition of the initiating or chase nucleotides, as indicated in the figure legends. Chase reactions were performed by adding 1.38 μl of 5 mM GTP, UTP, and CTP in transcription buffer to a 12.5-μl reaction at the indicated times. 5-Bromo-UTP and the sodium salts of sarkosyl and heparin were obtained from Sigma. Nuclease digestions were performed on transcription reactions after the addition of a-amanitin to 2 μg/ml. RNase H (U.S. Biochemical Corp.; 10 units/ml) or RNase A (Sigma; 10 μg/ml) was then added and incubated for 10 min at 37 °C.

Transcription reactions were terminated by the addition of 85% phenol/CHCl3, and digested with calf intestinal alkaline phosphatase to reduce background levels as described (4). Reactions containing RNase A were treated with proteinase K (Life Technologies, Inc.) at 1 mg/ml for 10 min prior to phenol/CHCl3 extraction. RNAs were resuspended in 7.5 μl of water, 1.0% sarkosyl, and 100 μM EDTA and electrophoresed in 28% polyacrylamide gel (25:3 acrylamide: bisacrylamide). The gels were silver stained (27) and scanned in a PhosphorImager (Molecular Dynamics, Inc., Sunnyvale, CA) and analyzed using ImageQuant software. After drawing a line from the transcript at C15 through the short transcripts, the peak areas were calculated. The intensity of each band was corrected by dividing the number of radioactive CTP and UTP residues present in that particular transcript, which allowed us to calculate the relative abundance of all of the transcripts in a gel lane. Blockage at a particular position (formally analogous to termination efficiency) was computed by taking the number of transcripts at that position and dividing by the sum of those transcripts and all longer transcripts.

**RESULTS**

Our laboratory has studied abortive initiation by RNA polymerase II at a variety of promoters using HeLa nuclear extracts as the source of the transcriptional machinery. This work revealed that the efficiency of clearance can vary significantly among promoters (4, 10). In order to further explore the molecular basis of this effect, we took advantage of the fact that the general transcription factor requirements for initiation by RNA polymerase II are considerably simplified with templates that are mismatched (23–26) or easily denatured (25, 30) at the initiation site. We began our experiments with a variant of the AdMLP in which bases −9 to +3 are mispaired (Fig. 1). This bubble duplex template was based on the work of Pan and Greenblatt (23), who showed that an AdMLP with a −9 to +3 unpaired region supported strong in vitro promoter activity when assayed with RNA polymerase II, TFII B, and TFIIIB. We found that RNA polymerase II alone would initiate transcription accurately on this template, provided that a dinucleotide primer is provided (see below). Thus, we have been able to study the effects of template sequence on RNA polymerase II promoter clearance in the absence of general transcription factors.

**Pure RNA Polymerase II Directs Initiation Events on Bubble Duplex Templates in the Absence of General Transcription Factors**—As a control for the bubble duplex reactions, we prepared preinitiation complexes on conventional double-stranded DNAs bearing the AdMLP. Templates were incubated in HeLa extracts and purified by gel filtration to remove most of the contaminating NTPs. Initiation on these templates may be primed with either the ApC or CpA dinucleotides, which pair with the +1/+2 and −1/+1 positions on the template strand (see Fig. 1 and Refs. 3, 4, and 31). To simplify discussion, we will identify transcripts produced with either dinucleotide primer by their 3'-ends, using +1 as the normal start site. The addition of limiting levels of labeled UTP and CTP to ApC- or CpA-primed preinitiation complexes on double-stranded templates allowed initiation and transcription elongation until +15, where a G residue is required (C15 RNAs; lanes 11–12 and 29–30 of Fig. 2). Some complexes continued elongation to +20, where UTP is required, presumably because of the presence of trace levels of GTP in the other nucleotides. All transcripts longer than 3 nt were made by RNA polymerase II, since their synthesis was abolished by 2 μg/ml a-amanitin (lanes 13 and 31). As expected from earlier work with preinitiation complexes assembled in crude nuclear extracts (4), some ApCpU and CpApG synthesis was aminatin-resistant.

**Transcript initiation by RNA polymerase II on double-stranded templates normally requires either ATP or dATP as an energy source (3, 32, 33) and is sensitive to the detergent sarkosyl at 0.05% (34) or the polyanion heparin (see, for example, Ref. 35). There was no a-amanitin-sensitive transcription of the double-stranded AdMLP template when dATP was absent from the reaction (Fig. 2, lanes 14 and 32) or when 0.05% sarkosyl or 100 μg/ml heparin was added before the NTPs (lanes 15, 16, 33, and 34). All of the RNAs longer than about eight bases chased efficiently when GTP and excess CTP and UTP were added to the original reactions (lane 18; data not shown for Cpa priming), while the shorter RNAs appeared to be abortively initiated. These results agree with our earlier studies on this system (4, 10). The polymerase II ternary elongation complex is not expected to be sensitive to sarkosyl and heparin, and we saw no effect of these reagents on the elonga-

---

2 I. Samkurashvili and D. S. Luse, unpublished observations.
Considerably more RNA was made by pure RNA polymerase II than by ternary complexes. Elongation resumed for complexes with the ApC dinucleotide. Elongation of complexes stalled downstream of U7 (lanes 19 to 20; data not shown for CpA priming). Transcription of the single-stranded templates was essentially complete in 5 min (compare lanes 11 and 12, and compare lanes 29 and 30). When pure RNA polymerase II alone was incubated with the AdMLP bubble template, the dinucleotide primer ApC, and limiting levels of labeled UTP and CTP, synthesis of amanitin-sensitive short RNAs was observed (Fig. 2, lanes 1–3). Similar results were obtained with pure RNA polymerase II and the CpA primer on the bubble templates (lanes 23–25). No RNA was synthesized in the pure polymerase reactions when ATP, CTP, and UTP were used as substrates without a dinucleotide primer (data not shown). Most of the RNAs made in the pure polymerase reactions resulted from pausing at or upstream of +15. Some RNAs of >15 nt were also obtained. Possible origins for these RNA will be discussed below.

Essentially no RNA was made in the pure polymerase reactions when either 0.05% sarkosyl or 100 μg/ml heparin was added before the NTPs (lanes 5, 6, 27, and 28). This is consistent with a true initiation event. As expected, transcription by pure polymerase II on the bubble templates did not require ATP or dATP (compare lanes 2 and 4 or lanes 24 and 26). Considerably more RNA was made by pure RNA polymerase II on the bubble templates if the 5-min reactions were extended to 30 min (compare lanes 1 and 2, and compare lanes 23 and 24), in contrast to the case of transcription of the double-stranded templates in extracts. Chase of dinucleotide-primed pure polymerase II reactions with GTP, CTP, and UTP resulted in production of the expected U20 product (Fig. 2, lane 8; data not shown for CpA primer). Elongation resumed for complexes with RNAs as short as 9 nt, but in no case were the pure RNA polymerase II ternary complexes fully active for continued RNA synthesis. Typically, 40–75% of dinucleotide-primed C15 complexes made by pure polymerase II on the bubble templates could resume transcription upon chase. The failure of complexes in the pure polymerase system to chase quantitatively was not simply the result of the longer reaction times used on the bubble templates, since 30-min initial reactions with double-stranded templates yielded C15 complexes that were fully active in a subsequent chase (lane 22). In contrast to initiation, transcript elongation on the bubble templates by pure polymerase II was not sensitive to sarkosyl or heparin (lanes 9 and 10; data not shown for CpA primer).

To determine whether the short transcripts that could not be chased in the pure polymerase reactions were released from ternary complex, we performed gel filtration with Sephacryl S-200 on ApC-primed reactions. Ternary complexes should elute in the void volume, whereas released transcripts should appear in the included fractions in this experiment. As seen in the left part of Fig. 3A, RNAs 7 nt and shorter appeared exclusively in the later eluting included fractions, indicating that these RNAs were released from the transcription complex. Active ternary complexes were almost entirely confined to the void volume. In the experiment in Fig. 3A, 75% of the C15 complexes that were excluded from the column could chase (lanes 4 and 5). Presumably, the failure of some transcripts to chase in lane 5 resulted from ongoing transcript release by complexes halted at C15 prior to the pooling of fractions and the addition of chase NTPs. A very small proportion of active complexes trailed into the initial included fractions (lanes 6 and 7). The void volume fraction contained only RNAs 7 nt or longer, and at least some of the RNAs of each length could chase. In contrast, essentially all of the RNAs in the included fractions failed to chase. These results are very similar to those obtained upon gel filtration of dinucleotide-primed RNAs made on double-stranded templates by RNA polymerase II in extracts, except that in the latter case all of the complexes running in the void volume were active in transcript elongation upon subsequent chase (3, 4).

As noted above, RNAs longer than 15 nt were observed in
transcription reactions with pure RNA polymerase II in the absence of GTP. Some of these 15-nt RNAs may have resulted from GTP contamination in the other nucleotides, but not all of them can be explained in this way. In particular, many of these RNAs did not comigrate with transcripts produced by read-through of the G-stop on the double-stranded templates. All of the 15-nt RNAs must be polymerase II products, since their synthesis was amanitin-sensitive. While we have not further investigated the origins of these anomalous transcripts, we suspect that they resulted from the addition of NTPs to transcripts released from the template along with the RNA polymerase in binary complexes. This supposition seems reasonable, since Johnson and Chamberlin (36) demonstrated that binary complexes of polymerase II and longer RNAs could cleave these RNAs in the presence of TFIIS and then add a limited number of bases to the newly generated 3'-ends.

Increased Abortive Initiation by RNA Polymerase II Occurs during Transcription of a Variant AdMLP Template with a More U-rich Initially Transcribed Region—We showed that a mouse b-globin promoter supported much higher levels of abortive initiation and correspondingly lower levels of productive transcription when compared with the AdMLP (10). In a series of subsequent experiments, we found3 that simply making the first six transcribed bases of the AdMLP identical to the same six bases of the globin promoter resulted in a template that supported high levels of abortive initiation, approaching those seen with the intact globin promoter. Only two base changes, at positions +4 and +6, were needed to match the AdMLP and globin promoter over this initially transcribed region. In order to explore the role of template sequence alone in this effect, we synthesized a variant of the AdMLP bubble template, called AdUMut, incorporating the two base changes just discussed (see Fig. 1). The results of transcribing this template with either pure RNA polymerase II or HeLa nuclear extract are shown in Fig. 4. The AdMLP bubble template was transcribed as a control.

As in the case of the AdMLP template, transcription of the AdUMut template with either dinucleotide primer, pure RNA polymerase II, and labeled UTP and CTP resulted in elongation to the G stop at C15, with some production of longer RNAs. The striking difference between the templates was the much higher level of abortive initiation obtained on AdUMut with either primer (in Fig. 4, compare lanes 1–3 with lanes 7–9, and compare lanes 13–15 with lanes 19–21). To quantify this difference, we computed the fraction of transcription complexes that could continue RNA synthesis past a certain template location, a parameter we call blockage. Blockage is defined as the ratio of the number of transcripts stopped at a particular location to the number of transcripts at that position plus all transcripts that read through to downstream locations. These values, expressed as percentages, are given in Table I for either ApC- or CpA-primed reactions with pure polymerase II. We observed

---

3 J. Kitzmiller and D. S. Luse, unpublished observations.
from 1.4- to 7-fold increases in abortive initiation on the AdU-Mut template relative to the AdMLP template within the first 8 nt of transcribed sequence. Interestingly, the maximum increase in abortive initiation with either primer occurs, not at a particular template location but at a transcript length of 8 nt (Table I).

Since the large majority of the short (<9-nt) RNAs made by pure RNA polymerase II on the AdUMut template could not be chased, it seemed likely that they had been released from ternary complex, as was the case with comparably sized RNAs made on the AdMLP bubble templates. We again used gel filtration on Sephacryl S-200 to test this. After the initial 30-min incubation of pure polymerase II and the AdUMut bubble template, none of the transcripts stalled at U3 to U8 and 47% of the transcription complexes stalled at C15 could be chased (Fig. 3A, lane 18). Only a very small proportion of active ternary complexes survived the gel filtration on the AdUMut template as compared with the AdMLP template (compare lanes 19 and 20 with lanes 4 and 5). All of the AdUMut transcripts 8 nt or shorter were found exclusively in the later eluting excluded fractions, indicating that they had been released from ternary complex.

We also transcribed the AdMLP and AdUMut bubble templates in HeLa nuclear extracts. For these studies, we incubated the bubble duplexes in HeLa extracts and then purified preinitiation complexes by gel filtration. On a given template, we observed even greater levels of abortive initiation with extracts than with pure polymerase II reactions under identical conditions (in Fig. 4 compare lanes 5 and 2, lanes 11 and 8, lanes 17 and 14, and lanes 23 and 20, respectively; see also Table I). The AdUMut template supported greater levels of abortive initiation than the AdMLP template when transcribed by polymerase II in extracts, with either primer (compare lanes 5 and 11, and compare lanes 17 and 23), in agreement with the effect seen with pure RNA polymerase II. The short transcripts synthesized in extracts were not simply due to transcriptional stalling, since the addition of chase NTPs did not cause these complexes to resume elongation (compare lanes 6, 12, 18, and 24 with unchased reactions). The amount of transcript produced from the bubble templates in extracts did not increase between 5 and 30 min (data not shown). Thus, the relatively slow RNA synthesis by pure polymerase II that we observe on bubble templates in comparison with transcription of double-stranded templates (see Fig. 2) is not an intrinsic property of RNA polymerase II transcribing a bubble template. In the extract reactions, small amounts of ~11–15-nt RNAs were made even in the presence of 2 μg/ml α-amanitin (lanes 4, 10, 16, and 22); we presume that these are transcripts by RNA polymerases I and/or III from the nuclear extracts.

**Fig. 3. Short transcripts produced on bubble templates are released from ternary complex.** The indicated templates were transcribed with pure RNA polymerase II, ApC and 10 μM 32P-labeled UTP and CTP. Reactions were applied to Sephacryl S-200 columns; pools of five consecutive drops were made beginning at drop 13 (void) and successive fractions (included) were collected. In the left two panels, RNAs are shown from an aliquot of the 30-min input reaction (lanes 2 and 17), as well as RNAs from reactions performed in the presence of 2 μg/ml α-amanitin for 30 min (lanes 1 and 16) or chased with 0.5 mM CTP, GTP, and UTP for 5 min (lanes 3 and 18) after a 25-min initial pulse. Numbers give the position and base at the 3'-end of indicated RNA, counting from the A at +1. A, for each fraction, RNAs are shown from an untreated aliquot (lanes marked with a minus sign) and an aliquot chased with 0.5 mM CTP, GTP, and UTP (lanes marked with a plus sign). B, for each fraction, RNAs are shown from an untreated aliquot (lanes marked with a minus sign) and aliquots treated with RNase H or RNase A. An aliquot of the void fraction was chased with 0.5 mM CTP, GTP, and UTP (lane 2).
merases and either bubble or oligo(dC)-tailed templates, it was reported that the transcripts formed extended RNA-DNA hybrids (37–40). To determine whether the short transcripts released by polymerase II as abortive initiation products were present as RNA-DNA hybrids (RNAse H sensitive and RNAse A resistant) or were released from ternary complex as normally blocked at C15 were also present in included fractions (Fig. 3). In all cases, both CTP and UTP were present at 10 μM and were 32P-labeled.

The percentages of transcription complexes at template positions +4 and +10 that cannot continue transcription are given in the table for the indicated combinations of bubble templates and dinucleotide primers. This blockage value is the number of complexes with n-mer RNAs plus the number of complexes with RNAs longer than n. The values are the averages of six independently processed samples; the S.D. (σ, – 1) is indicated. In all cases, both CTP and UTP were present at 10 μM and were 32P-labeled.

**TABLE I**

Percentages of transcription complexes that fail to continue RNA synthesis between positions +4 and +10 during transcription of AdMLP and AdUMut bubble templates by pure RNA polymerase II

| RNA 3'-end | AdMLP, ApC | AdMLP, ApC | AdMLP, ApC | AdMLP, ApC | AdMLP, CpA | AdMLP, CpA | AdMLP, CpA |
|------------|------------|------------|------------|------------|------------|------------|------------|
| C4/U4      | 33.1 ± 2.4 | 46.2 ± 2.2 | 1.40       | 23.3 ± 3.7 | 22.8 ± 6.2 | 0.98       |
| U5         | 28.4 ± 4.0 | 43.6 ± 2.6 | 1.54       | 9.6 ± 1.6  | 18.4 ± 3.7 | 1.92       |
| C4/U4      | 11.7 ± 4.0 | 27.5 ± 1.8 | 2.35       | 6.0 ± 2.1  | 19.0 ± 2.0 | 3.17       |
| U7         | 13.8 ± 2.0 | 52.9 ± 3.0 | 3.83       | 3.6 ± 0.7  | 20.4 ± 2.5 | 5.67       |
| U8         | 6.2 ± 0.9  | 40.4 ± 5.4 | 6.51       | 3.3 ± 0.4  | 6.9 ± 0.9  | 2.09       |
| C9         | 2.5 ± 1.1  | 5.5 ± 1.8  | 2.20       | 3.7 ± 2.1  | 3.5 ± 2.4  | 0.93       |
| C10        | 2.1 ± 0.7  | 2.5 ± 1.9  | 1.19       | 2.8 ± 0.5  | 3.9 ± 0.7  | 1.40       |

In particular, the decreased stability of C15 complexes is not the result of formation of continuous RNA-DNA hybrids.

The Incorporation of a Nucleotide Analog That Should Increase the Stability of the RNA-DNA Hybrid Reduces Abortive Initiation—We compared initiation by pure RNA polymerase II in the presence of either UTP or 5-bromo-UTP (Br-UTP), a UTP analog that strengthens the RNA-DNA hybrid. For these experiments, labeled CTP was present at 0.5 μM, and no nonlabeled CTP or labeled UTP was added. The latter condition avoided competition between labeled UTP and Br-UTP for incorporation into RNA (see Fig. 5). When the AdMLP bubble template was transcribed with UTP and limiting labeled CTP for 30 min, the expected transcripts stalled at C15 were produced, as well as prominent RNAs corresponding to stops before C incorporation at U5, U8, and U14 (Fig. 5, lane 2). A fraction of transcription complexes stalled at C15 and U14 chased after the addition of GTP, CTP, and UTP (lane 3) or GTP, CTP, and Br-UTP (GU/C chase, lane 4). Incorporation of Br-UTP decreased electrophoretic mobility relative to UTP-containing RNAs (compare, for example, the 20-mer in lanes 3 and 4). Performing initial transcriptions on the AdMLP with Br-UTP resulted in a significant reduction in abortive...
initiation (compare lanes 2 and 6 of Fig. 5), most prominently at U5. Blockage was almost eliminated at U5 for the particular reaction shown in Fig. 5, and it was reduced an average of 3.7-fold over eight experiments. Blockage after position U8 was reduced by an average of 1.7 fold over eight experiments. For reactions with Br-UTP on the AdUMut bubble template, we used the same conditions as for the AdMLP template except that the CpA dinucleotide primer was substituted. Transcription for 30 min with UTP produced transcripts stalled at C15 and U14 as well as transcripts stopped at U3 and from U5 to U8 (Fig. 5, lane 12). Incorporation of Br-UTP during the initial transcription reaction reduced abortive initiation, as for the AdMLP template (compare lanes 11 and 15). Blockage was eliminated after U3 (lane 15) and reduced from 1.6- to 2.2-fold (the average of eight experiments) from U5 through C9.

If abortive initiation is reduced when Br-UTP is substituted for UTP, one might expect more full-length (U14 and C15) transcripts to accumulate in the presence of the U analog. For the experiment shown in Fig. 5, we computed the amount of RNA at each transcript length, beginning at C5 for the AdMLP reactions and U5 for the AdUMut reactions and then summed these numbers. For the AdMLP template with UTP as substrate, 10% of the total RNA C5 and longer was U14 or C15, but when Br-UTP was substituted for UTP 22% of the total RNA was U14 or C15. The effect was even stronger with the AdUMut template. Transcription with UTP gave 6% of the total RNA as U14/C15, while 31% of the total was U14/C15 when Br-UTP was substituted.

DISCUSSION

In order to further explore the role of the sequence of the initially transcribed region in initiation efficiency by RNA polymerase II, we have taken advantage of the simplification provided by the use of bubble templates. In pilot experiments that led to the present study, we reproduced the earlier result (23) that transcription of AdMLP bubble templates by RNA polymerase II with normal NTP substrates requires, at minimum, TBP and TFIIB (data not shown). We were somewhat surprised to discover that priming RNA synthesis with a promoter-complementary dinucleotide was sufficient to remove any initiation factor requirement (Fig. 2). The start of transcription on bubble templates by pure RNA polymerase II in the presence of dinucleotide primers represents a reasonable model of normal transcript initiation.

The ability to start transcription and escape abortive initiation on the bubble templates without any additional proteins shows that RNA polymerase II has no absolute factor requirement for either process. One might argue that this point is already well established, since templates with 3'-dC extensions
support transcription by polymerase II (39). However, such templates do not provide a close structural analog to an initiation intermediate, as the bubble templates do. Furthermore, in most cases the large majority of transcripts on these dC-tailed templates cannot be extended beyond about 15 nt (38). The most important consequence of having a factor-independent initiation reaction was the ability to compare transcription efficiencies on the same initially transcribed sequence in the presence and absence of factors. The change of two bases, at +4 and +6, between the AdUMut and AdMLP bubble templates resulted in a greatly increased level of abortive initiation by pure RNA polymerase II on AdUMut. The effect is strongest for the fifth through the eighth bases added, with an average increase in abortive initiation of 3–4-fold during this stage of transcription (Table I). This result is very similar to the 5-fold increase in abortive initiation that we reported in an earlier comparison, using double-stranded DNAs and nuclear extracts, of AdMLP and the AdUMut-like mouse β-globin promoter (10). The fact that we can essentially duplicate the nuclear extract/double-stranded template result with bubble templates and pure RNA polymerase II suggests that the efficiency with which polymerase escapes abortive initiation is to a considerable extent determined by the transcript-template sequence alone.

What is the mechanistic basis for the increase in abortive initiation seen with AdUMut versus the AdMLP template? In approaching this question it is useful to briefly review recent findings on the extent of the RNA-DNA hybrid within the transcription complex, an area that has been somewhat controversial. Rice, Kane, and Chamberlin (18) studied this issue by treating RNA polymerase II ternary complexes with ribonuclease and then testing for retention of the 3′ portion of transcripts within the complexes. They concluded that RNAs as short as 3 nt could be retained in active ternary complexes and extended upon NTP addition (18). While this finding argues against any extensive RNA-DNA hybrid, several other recent studies have reached the opposite conclusion. Nuelder et al. (12) incorporated a cross-linkable U analog at various positions within transcripts made by defined E. coli ternary complexes. This analog cross-linked to the template A residue with which it was presumably base-paired, and only to that residue, as long as the U analog was present between 2 and 8 nt upstream of the 3′-end of the RNA. When the analog was present further upstream, only cross-links to the RNA polymerase were observed. Incorporation of other nucleotide analogs that would strengthen the RNA-DNA hybrid reduced the tendency for transcriptional arrest when these analogs were placed within eight bases of the 3′-end, but no effect was seen upon incorporation further upstream. Nuelder et al. (12) concluded that the RNA-DNA hybrid within their transcription complexes is 8 nt long. Essentially the same conclusion was reached by Kashlev and colleagues in two very recent reports. In the first of these (15), transcription complexes were assembled with 30 base bubble templates mismatched over the central nine bases. Short RNAs were hybridized to one strand of the bubble region and these hybrids were challenged with E. coli RNA polymerase. Functional transcription complexes were obtained by this approach with RNAs of 6 nt or longer. Six-nucleotide RNAs gave rather unstable complexes; stability was increased with 7-nucleotide RNAs and was maximal with 8-nucleotide RNAs. Significantly, 8-nucleotide RNAs that could hybridize only over the 3′ six bases gave rise to unstable complexes. These results were also taken to indicate that the RNA-DNA hybrid in the transcription complex is 8 base pairs long. Komissarova and Kashlev (16) reached similar conclusions from an entirely different method, one which uses normal transcription complexes. They found (16) that RNase digestion of ternary E. coli polymerase complexes did not result in cleavages closer than 14–16 bases from the 3′-end of the nascent RNA. Transcripts could be truncated down to 8–10 nt, but no further, with the E. coli GreB factor or pyrophosphorylase. Most significantly, Komissarova and Kashlev (16) were able to detect apparent RNase cleavage to within three bases of the 3′-end of the nascent RNA in active complexes, but this resulted from the failure of a denaturing agent to immediately inactivate the RNase at the end of the experiment. When this problem was eliminated, no RNase cleavages closer than 14–16 bases from the 3′-end were detected.

Our results are consistent with a model in which the stable transcript elongation complex contains an 8-base pair RNA-DNA hybrid, with RNA upstream of the hybrid interacting in some way with the RNA polymerase. First, we obtained much more abortive initiation with the AdUMut template relative to the AdMLP DNA. With the ApC primer, the greatest difference was obtained for 8-nucleotide RNAs (Table I). This is the RNA length at which the AdUMut transcript (ACUUUUUU) has the longest continuous run of U residues at its 3′-end. In comparison, the corresponding 8-nucleotide AdMLP transcript is ACUCUCUU. Since the U3a hybrid is unusually weak (reviewed in Ref. 17), one would expect that abortive initiation would occur more frequently on AdUMut than on AdMLP and that the greatest difference would be seen at +8. The decrease in abortive initiation and increased yield of 15-nt RNA that we observed when Br-UTP was substituted for UTP is also consistent with the importance of RNA-DNA hybrid strength in determining the stability of the early transcription complex.

All of our results cannot be explained simply by invoking differences in hybrid strength. A striking example is the relative level of abortive initiation on the AdUMut and AdMLP templates when CpA was used as a primer instead of ApC. In this case (Table I), the greatest increase in abortive initiation with the AdUMut template compared with AdMLP was seen at U7, although the longest U-run occurs for the U8 transcript. This finding can be explained if RNA immediately upstream of the 8-base pair hybrid region enters an RNA binding channel on the RNA polymerase (see Refs. 16 and 41). The putative RNA-RNA polymerase interaction would stabilize the transcription complex and partially compensate for differences in hybrid stability at the 3′-end of the RNA. It is also important to note that while we favor an explanation of our data based on differential hybrid strength, other models are also possible. For example, if U-rich RNAs interact less well with the RNA polymerase than do other RNAs, then the difference in abortive initiation between the AdUMut and AdMLP templates may be explained without invoking an RNA-DNA hybrid longer than a few base pairs (18). Such an explanation would also require that the substitution of Br-U for U results in somewhat higher transcript affinity.

While the bubble template system has proven useful in investigating the role of transcript sequence in the initiation process, our results have also revealed some limitations to this approach. One of these is the slower rate of transcript accumulation on the bubble templates relative to fully double-stranded templates. We have performed template challenge experiments and also some preliminary assays with matrix-attached bubble templates. In all of these tests (data not shown), template commitment with the bubble templates and pure RNA polymerase II was found to be very rapid (about 2 min). Thus, the slow accumulation of transcripts on the bubble templates must result from either slow initiation by RNA polymerase II or a slow isomerization of the template-bound polymerase into an initiation-competent state. Bubble templates transcribed in nuclear extracts showed relatively rapid RNA synthesis (Fig. 4),
which indicates that slow initiation is not an inherent property of bubble templates. This result raises the possibility that while transcription factors are not absolutely required for initiation by RNA polymerase II, such factors may serve to stimulate the process of initiation.

Another limitation of the bubble templates is the relative instability of transcription complexes, such as those bearing C\textsuperscript{15} RNAs, which have escaped the abortive initiation stage. On bubble templates, we typically see one-third to one-half of the C\textsuperscript{15} complexes fail to chase after a 30-min transcription reaction, while analogous complexes on double-stranded templates retain full transcriptional competence. The recent results of Holstege et al. (42) suggest a possible reason for this instability. These workers demonstrated that during transcription initiation by RNA polymerase II at the AdMLP, the initial denatured region of \(-9\) to \(-2\) extends continuously downstream until 10 bonds are made, at which point the upstream portion of the bubble abruptly reanneals, leaving a denatured region extending downstream from about +3. Such upstream reclosure cannot occur on our bubble templates, which would suggest that complexes stalled at locations just downstream of +10 might be destabilized by an inappropriately melted upstream region. We do not think that the simple presence of a long stretch of free template strand upstream of the polymerase is the instability signal, because the addition of DNA oligonucleotides complementary to the otherwise unpaired segment of the template strand upstream of \(-2\) did not increase the stability of C\textsuperscript{15} complexes (data not shown). The strain imposed by the inappropriately long transcription bubble in complexes with >11 nt nascent RNAs could cause the ejection of the polymerase as a binary transcript/polymerase complex, thereby leading to the limited, template-independent synthesis of RNAs >15 nt as discussed above.

The inability to reclose the upstream region could also explain the increased level of abortive initiation on bubble templates seen in extract transcription, relative to the pure RNA polymerase II reaction. This effect was particularly striking for the AdMLP template, on which almost no promoter clearance was achieved in extract transcription (Fig. 4). We can imagine that blocking reannealing with a permanently unpaired upstream region might lock initiation factors in place and prevent transcription from continuing. It is tempting to suggest that TFIIH is involved in this effect. Kumar et al. (43) recently showed that the lack of TFIIH in reactions performed with highly purified polymerase II transcription factors leads to a failure of transcription to progress beyond about +15 to +20. Thus, we can speculate that failure of the bubble to reclose might trap TFIIH in a configuration within the complex such that it actually blocks promoter clearance by RNA polymerase II.

In summary, the bubble template system has allowed us to demonstrate a major role for the sequence of the initially transcribed region in the extent of abortive initiation by RNA polymerase II. This approach should also be useful in the future in analyzing the contributions of subsets of the RNA polymerase II general transcription factors to the polymerase II initiation process.

Acknowledgments—We thank Danny Reinberg, Robert Landick, and Zachary Burton for generous gifts of transcription factors and factor expression constructs and Donna Driscoll, Richard Gronostajski, and Richard Fadgett for advice during the course of these experiments. We also thank Natalia Komissarova and Mikhail Kashlev for communicating results prior to publication.

REFERENCES

1. McClure, W. R. (1985) Annu. Rev. Biochem. 54, 171–204
2. von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. (1984) Annu. Rev. Biochem. 53, 389–446
3. Luse, D. S., and Jacob, G. A. (1987) J. Biol. Chem. 262, 14990–14997
4. Jacob, G. A., Luse, S. W., and Luse, D. S. (1991) J. Biol. Chem. 266, 22537–22544
5. Carrozzini, A. J., and Gralla, J. D. (1985) J. Mol. Biol. 183, 165–177
6. Straney, D. C., and Crothers, D. M. (1987) J. Mol. Biol. 193, 267–278
7. Hsu, L. M., Vo, N. V., and Chamberlin, M. J. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 11588–11592
8. Krummel, B., and Chamberlin, M. J. (1989) Biochemistry 28, 7829–7842
9. Ellinger, T., Behnke, D., Bujard, H., and Gralla, J. D. (1994) J. Mol. Biol. 239, 455–465
10. Jacob, G. A., Kitzmiller, J. A., and Luse, D. S. (1994) J. Biol. Chem. 269, 3655–3663
11. Reeder, T. C., and Hawley, D. K. (1996) Cell 87, 767–777
12. Nudler, E., Mustaa, A., Lukhanin, E., and Goldfarb, A. (1997) Cell 89, 33–41
13. Komissarova, N., and Kashlev, M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 1755–1760
14. Komissarova, N., and Kashlev, M. (1997) J. Biol. Chem. 272, 15329–15338
15. Sidorenkov, I., Komissarova, N., and Kashlev, M. (1998) Mol. Cell 2, 55–64
16. Komissarova, N., and Kashlev, M. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 14499–14504
17. Landick, R. (1997) Cell 88, 741–744
18. Rice, G. A., Kane, C. M., and Chamberlin, M. J. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 4245–4249
19. Orphanides, G., Lagrange, T., and Reinberg, D. (1996) Genes Dev. 10, 2657–2683
20. Goodrich, J. A., and Tjian, R. (1994) Cell 77, 145–156
21. Dvir, A., Conaway, R. C., and Conaway, J. W. (1996) J. Biol. Chem. 271, 23352–23356
22. Dvir, A., Conaway, R. C., and Conaway, J. W. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 9006–9010
23. Pan, G., and Greenblatt, J. (1994) J. Biol. Chem. 269, 30101–30104
24. Tantin, D., and Carey, M. (1994) J. Biol. Chem. 269, 17397–17400
25. Holstege, F. C. P., Tantin, D., Carey, M., Van der Vliet, P. C., and Timmers, H. T. M. (1996) EMBO J. 15, 810–819
26. Holstege, F. C. P., Van der Vliet, P. C., and Timmers, H. T. M. (1996) EMBO J. 15, 1666–1677
27. Hodo, H. G., and Blatti, S. P. (1977) Biochemistry 16, 2335–2343
28. Thompson, N. E., Aronson, D. B., and Burgess, R. R. (1990) J. Biol. Chem. 265, 7069–7077
29. Lehn, M. G., and Luse, D. S. (1992) Genes Dev. 6, 1342–1356
30. Timmers, H. T. M. (1984) EMBO J. 13, 391–399
31. Luse, D. S., Kochel, T., Kuenzle, R. D., Coppola, J. A., and Cai, H. (1987) J. Biol. Chem. 262, 2725–2731
32. Bunick, D., Zandomeni, R., Ackerman, S., and Weinmann, R. (1982) Cell 29, 877–886
33. Sawadogo, M., and Roeder, R. G. (1984) J. Biol. Chem. 259, 5321–5326
34. Hawley, D. K., and Roeder, R. G. (1985) J. Biol. Chem. 260, 8163–8172
35. Reinberg, D., and Roeder, R. G. (1987) J. Biol. Chem. 262, 3310–3321
36. Johnson, T. L., and Chamberlin, M. J. (1994) Cell 77, 217–224
37. Dedrick, R. L., and Chamberlin, M. J. (1985) Biochemistry 24, 2245–2253
38. Sluder, A. E., Price, D. H., and Greenleaf, A. L. (1988) J. Biol. Chem. 263, 9917–9925
39. Kadesch, T., and Chamberlin, M. J. (1982) J. Biol. Chem. 257, 5286–5295
40. Duarte, S. S., Hart, C. R., and von Hippe1, P. H. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 9539–9543
41. Nudler, E., Gusarow, L., Aretissova, E., Kozlov, M., and Goldfarb, A. (1998) Science 281, 424–428
42. Holstege, F. C. P., Fiedler, U., and Timmers, H. M. (1997) EMBO J. 16, 7468–7480
43. Kumar, K. P., Akoulitchev, S., and Reinberg, D. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 9767–9772
44. Cowie, A., Jat, P., and Kamen, R. (1982) J. Mol. Biol. 159, 225–255
45. Biswas, T. K. (1997) Arch. Biochem. Biophys. 340, 250–256