THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 270, No. 5, Issue of February 3, pp. 2290–2297, 1995
Printed in U.S.A.

RNA Polymerase II Ternary Complexes May Become Arrested after Transcribing to within 10 Bases of the End of Linear Templates*

(Received for publication, September 19, 1994)

Michael G. Izban‡, Irakli Samkurashvili, and Donal S. Luse§

From the Department of Molecular Genetics, Biochemistry, and Microbiology, University of Cincinnati, College of Medicine, Cincinnati, Ohio 45267-0524

In the presence of elongation factor SII, arrested RNA polymerase II ternary complexes cleave 7–17 nucleotides from the 3′-ends of their nascent RNAs. It has been shown that transcription of linear templates generates apparent run-off RNAs, which are nevertheless truncated upon incubation with SII. By using high resolution gels, we demonstrate that transcription of blunt or 3′-overhung templates with RNA polymerase II generates two populations of ternary complexes. The first class pauses 5–10 bases prior to the end of the template strand. These complexes respond to SII by cleaving approximately 9–17 nucleotide RNAs from their 3′-ends and therefore may be termed arrested. A second class of complexes, which fail to respond to SII, transcribe to within 3 bases of the end of the template strand. These complexes appear to have run off the template since they have released their nascent RNAs. Run-off transcription occurs on all types of templates, but it is the predominant reaction on DNAs with 5′-overhung ends. Thus, RNA polymerase II ternary complexes that retain 5–10 bases of contact with the template strand downstream of the catalytic site become arrested. Further reduction of downstream template contacts can lead to termination. We also show that the addition of Sarkosyl to the elongation reactions significantly changes the pattern of transcriptional arrest near the end of linear templates.

It has been appreciated for some time that a fraction of the transcribing RNA polymerase II ternary complexes may stop transcription after traversing certain sequences, thereby entering a state called arrest (see, for instance, Refs. 1–3). The sites that provoke arrest typically encode runs of U within the transcript; other sequence features are usually present as well (1, 2). Arrested polymerases retain their RNAs in ternary complex but cannot continue transcription. Elongation factor SII has been shown to facilitate the passage of polymerase II through arrest sites (4–6). Several laboratories have shown that in the presence of SII, arrested ternary complexes acquire ribonuclease activity, cleaving their nascent transcripts from the 3′-end (7–13). After cleavage, transcription continues from the newly created 3′-OH group (9, 10). Thus, the transcript cleavage activity rescues the transcription complex from an inactive state and allows another opportunity for passage through the arrest site. We have demonstrated that the initial SII-facilitated cleavage by arrested complexes usually occurs in a 7–17-nt[1] increment (10). Elongation-competent ternary complexes that have halted transcription simply because the next NTP needed for elongation is missing (a condition we refer to as stalling) also cleave their nascent RNA in the presence of SII. However, transcript cleavage in stalled complexes generates predominantly dinucleotides (pNpNs) with trace levels of NMPs (11). The large difference in SII-facilitated cleavage increment between elongation-competent stalled complexes and elongation-incompetent arrested complexes strongly suggests that arrest is accompanied by an internal reorganization of the ternary complex.

It has been recently reported that a significant portion of promoter-initiated RNA polymerase II ternary complexes that have transcribed to the end of linear templates do not actually run off the DNA but remain active, as judged by their ability to truncate nascent transcripts in the presence of elongation factor SII (9, 12, 14). The initial increment of transcript cleavage was not identified in these studies, so it was not possible to determine whether the active complexes were stalled or arrested. We show here that RNA polymerase II transcription complexes that approach the end of the template can have two fates. If transcription continues to within a few bases of the end of the template strand, the ternary complex apparently dissociates. However, on some templates there is strong tendency for the transcription complexes to accumulate 5–10 nt upstream of the end of the template strand. The majority of these complexes will cleave their nascent RNAs when SII is added, and the increment of cleavage (usually 10–15 nt) clearly indicates that these complexes are arrested. Thus, the state of transcriptional arrest may be caused not only by distinctive sequences within the transcript and template but apparently also by the destabilization or loss of downstream template contacts.

MATERIALS AND METHODS

Reagents—Ultrapure (fast protein liquid chromatography purified) unlabeled NTPs and dATP were purchased from Pharmacia Biotech Inc. [α-^{32}P]CTP (800 Ci/mmol) was obtained from DuPont NEN. Bio-Gels A-1.5m and A-5m were purchased from Bio-Rad. The restriction enzymes *Pvu*II, *Hin*dIII, *Bgl*I, *Fsp*I, and *Eco*RI were purchased from Life Technologies, Inc.

Plasmids—The plasmids pML20 and pMB5 have been described elsewhere (8, 15). For the pMB5 template, *Eco*RI cleaves the template strand at +87 relative to the start site of transcription.

Elongation Factor SII—Recombinant human SII (rSII) was purified as described by Yoo *et al.* (16) and was either a gift from R. Landick (Washington University, St. Louis, Dept. of Biology) or purified in our laboratory. The concentration of rSII was determined by the Coomassie binding assay (Bio-Rad) using bovine serum albumin as a standard. The

* This work was supported by National Institutes of Health Grant GM 29487. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Dept. of Obstetrics and Gynecology and Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, TX 77555.

§ To whom correspondence should be addressed: Dept. of Molecular Biology, NC20, Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland, OH 44195. Tel.: 216-445-7688; Fax: 216-444-0512.

[1] The abbreviations used are: nt, nucleotide(s); rSII, recombinant human SII.

concentration used in each experiment is given in the figure legends.

Assembly and Purification of Stalled and Run-off Ternary Complexes—Ternary complexes stalled early in elongation and after "run-off" transcription on linear templates were generated essentially as previously described in protocols for circular templates (17). Templates were linearized by digesting with 10 units of restriction enzyme per μg of DNA for 2 h at 37 °C. Preinitiation complexes were formed by incubating 19 μg/ml of linear DNA with HeLa cell nuclear extract for 30 min at 25 °C prior to purification by Bio-Gel A-1.5m gel filtration. Elongation complexes stalled after the incorporation of a U residue at +20 (U20 complexes) were generated by incubating preinitiation complexes with 2 mM ApC, 10 μM dATP, 20 μM GTP, 20 μM UTP, 1 μM [α-^{32}P]CTP at 25 °C for 5 min followed by another 5-min incubation after the addition of CTP to 10 μM. The stalled ternary complexes were further purified by brief exposure to 1% Sarkosyl (5 min at 25 °C) followed by another round of Bio-Gel A-1.5m gel filtration. Sarkosyl-rinsed U20 complexes were chased to the end of the template by supplying 7.2 mM $MgCl_2$ and either 1 mM NTPs for 2–5 min at 37 °C to generate 5′-end-labeled complexes or 1 mM ATP, UTP, and GTP and 20 μM [α-^{32}P]CTP at 37 °C for 30 min to generate uniformly labeled complexes. In some cases, complexes were chased with 1 mM NTPs and 0.3% Sarkosyl for 2 min at 37 °C. U35 complexes, which we used to produce RNA ladders as size markers, were generated on pML20 templates in a two step process. U20 complexes were produced and incubated with 20 μM ATP to allow elongation through the triple A stop at positions +21, +22, and +23. The ATP was removed from the reaction mixture by gel filtration, and the A23 complexes were chased to the next A stop by supplying 20 μM CTP, GTP, and UTP. The resultant U35 complexes (typically 200 μl) were then used to generate the RNA ladders, either by supplying all four NTPs to 1 mM and incubating at 30 or 37 °C or by adding CTP, UTP, and GTP to 1 mM and ATP to 50 μM and incubating at 37 °C. Aliquots were removed every 30 or 15 s, respectively. The reactions were stopped by adding EDTA to a 3-fold excess over the Mg^{2+} concentration. The RNAs were purified, concentrated by ethanol precipitation, and resolved by polyacrylamide gel electrophoresis as previously described (10, 11).

SII-facilitated Truncation Reactions—SII-facilitated truncation reactions were performed as previously described (10, 11). Briefly, chased complexes were subjected to another step of gel filtration to remove the Mg^{2+} and unincorporated NTPs. Pooled void volume fractions (190 μl) were divided into 30-μl aliquots and supplemented with components as described in the figure legends. Time course reactions were performed in appropriately sized pooled reactions, 30-μl portions were removed at the times indicated in the figure legends, and the reaction was stopped by chelating the Mg^{2+} with EDTA. Samples were treated with proteinase K (150 μg/ml) for 1 h prior to sequential phenol:$CHCl_3$ and chloroform extractions. The RNAs were concentrated by ethanol precipitation or by vacuum lyophilization when we wished to characterize the short RNAs generated during the SII-facilitated transcript truncation. RNAs were resolved on denaturing (7.8 M urea) polyacrylamide gels run in TBE (89 mM Tris-HCl, 89 mM borate, 2 mM EDTA, pH 8.3) as indicated in the figure legends. After electrophoresis, the gels were exposed to preflashed Kodak X-AR film at −70 °C with a Lightning Plus intensifying screen and to PhosphorImager screens. Data were collected and quantitated from the phosphor screens using the Molecular Dynamics PhosphorImager (Sunnyvale, CA) and ImageQuant software.

RESULTS

To determine the nature of transcript cleavage in complexes that have reached the end of a linear template, we used the pML20 plasmid, which contains the Ad 2 ML promoter, linearized with one of the following restriction endonucleases: *Hin*dIII (which cuts the template strand at position +72 relative to the transcriptional start site, leaving a 5′-overhung end), *Bgl*I (+220, 3′-overhung end), *Fsp*I (+213, blunt end), and *Pvu*II (+163, blunt end). We will refer to the sets of ternary complexes that had transcribed to the end of these templates as pML20-*Hin*dIII, pML20-*Bgl*I, pML20-*Fsp*I, and pML20-*Pvu*II complexes, respectively. The strategy we employed to generate ternary complexes on linear templates was essentially identical to that used on circular templates (17). Briefly, preinitiation complexes were formed on the linear templates by incubation in HeLa cell nuclear extracts; the complexes were purified by gel filtration on Bio-Gel A-1.5m. Incubation in the presence of an ApC dinucleotide primer and appropriate concentrations of GTP, UTP, and [α-^{32}P]CTP leads to the formation of complexes stalled after the incorporation of a U residue at position +20 (U20 complex), prior to a triple A stop. Stalled U20 complexes are very stable and can be highly purified by transient exposure to the detergent Sarkosyl followed by another round of gel filtration. We have referred to this procedure as Sarkosyl rinsing (17). The second gel filtration step removes not only the Sarkosyl but also most DNA binding proteins, which disassociate from the DNA in the detergent, as well as the NTPs and Mg^{2+}. Sarkosyl rinsing also removes the known transcript elongation factors, leaving ternary complexes that apparently contain only RNA polymerase II itself. Sarkosyl-rinsed complexes were chased to the end of the linear templates by supplying Mg^{2+} and either nonlabeled NTPs to 1 mM, which generated 5′-end-labeled transcripts, or 1 mM ATP, GTP, and UTP and 20 μM [α-^{32}P]CTP, which generated uniformly labeled transcripts. The latter method was used when we sought to characterize the SII-facilitated cleavage products directly. In either case, the RNA polymerases chased efficiently to near the end of the template (data not shown). It is important to note that after run-off transcription, the ternary complexes were subsequently subjected to another round of gel filtration to remove Mg^{2+} and unincorporated NTPs before the addition of the SII factor.

Transcription and SII-facilitated Transcript Truncation on Linear Templates—U20 complexes (Fig. 1, *lane 1*) were chased as described above to generate uniformly labeled pML20-*Bgl*I complexes (*lane 3*). While most of the RNAs were roughly the size expected for elongation to the *Bgl*I site, we also observed some RNAs much larger than the expected 220-nt run-off. The longer RNAs were generated even when gel-purified DNA fragments were used as templates in transcription reactions. Furthermore, their production was α-amanitin sensitive (data not shown). These RNAs may have been generated either by end-to-end transcription of the linear DNA (the maximum length of RNA was approximately equal to that of the DNA fragments used in the reactions) or by RNA polymerase II initiations at cryptic promoters within pUC 18 (see, for example, Ref. 18). Regardless of their origins, these RNAs were produced only in trace amounts (about 4% of the total, assuming an average length of 800 nt for the aberrant transcripts and a 25% C content for both RNAs). The presence of this minor population of complexes does not compromise the results presented below.

pML20-*Bgl*I complexes were supplied with SII and Mg^{2+} and incubated for 5 (*lane 5*) or 30 (*lane 6*) min. Consistent with previous studies (9, 12, 14), complexes that had apparently transcribed to the end of this linear template nevertheless remained responsive to SII. The results of the 30-min SII-facilitated reaction (*lane 6*) demonstrated that RNA polymerase is capable of removing hundreds of bases via the transcript truncation reaction. The production of shortened transcripts in the presence of SII (*lanes 5* and *6*) was accompanied by a reduction in length of the minor population of longer RNAs, as expected since these RNAs are associated within RNA polymerase II ternary complexes. Also as expected, in the absence of SII, very little transcript truncation was observed (*lane 4*). The broad band of RNAs associated with pML20-*Bgl*I complexes (*lane 3*) suggested to us that transcription to the end of the linear template actually generated a heterogenous population of ternary complexes. Therefore, we set out to map the distribution of the 3′-ends of the RNAs in these complexes.

High Resolution Mapping of RNA Associated with Run-off Ternary Complexes—Using high resolution sequencing gels and appropriate RNA markers, we precisely mapped the distribution of RNAs generated by transcription to the end of *Hin*dIII-, *Bgl*I-, *Fsp*I-, and *Pvu*II-linearized templates (see Fig.

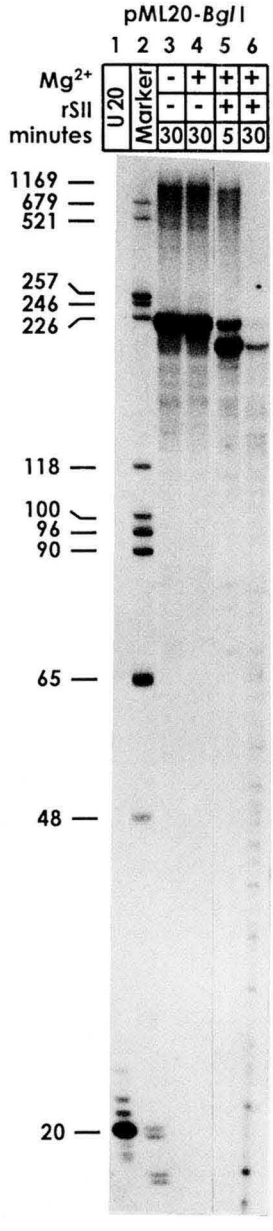

FIG. 1. **Transcription to near the end of *Bgl*I-digested pML20 template and subsequent SII-facilitated truncation.** Sarkosyl-rinsed U20 complexes generated on *Bgl*I-digested pML20 template (*lane 1*) were run to the end of the linear template by supplying Mg^{2+} and 1 mM ATP, UTP, and GTP and 20 μM [α-^{32}P]CTP and incubating for 30 min at 37 °C. After gel filtration, the resultant run-off complexes (*lane 3*) were incubated at 37 °C after the addition of 7.8 mM Mg^{2+} and 10 μg/ml recombinant SII as indicated. The RNAs were purified and then concentrated by lyophilization. One-third of the reactions were resolved on a 10% (acrylamide/bisacrylamide (29:1)) sequencing gel stopped after the bromphenol blue dye marker had run 37 cm. The DNA markers (*lane 2*) have been previously described (17). DNA markers of sizes 48–1169 nt are indicated at the *left* as is the RNA associated with U20 complex (*lane 1*).

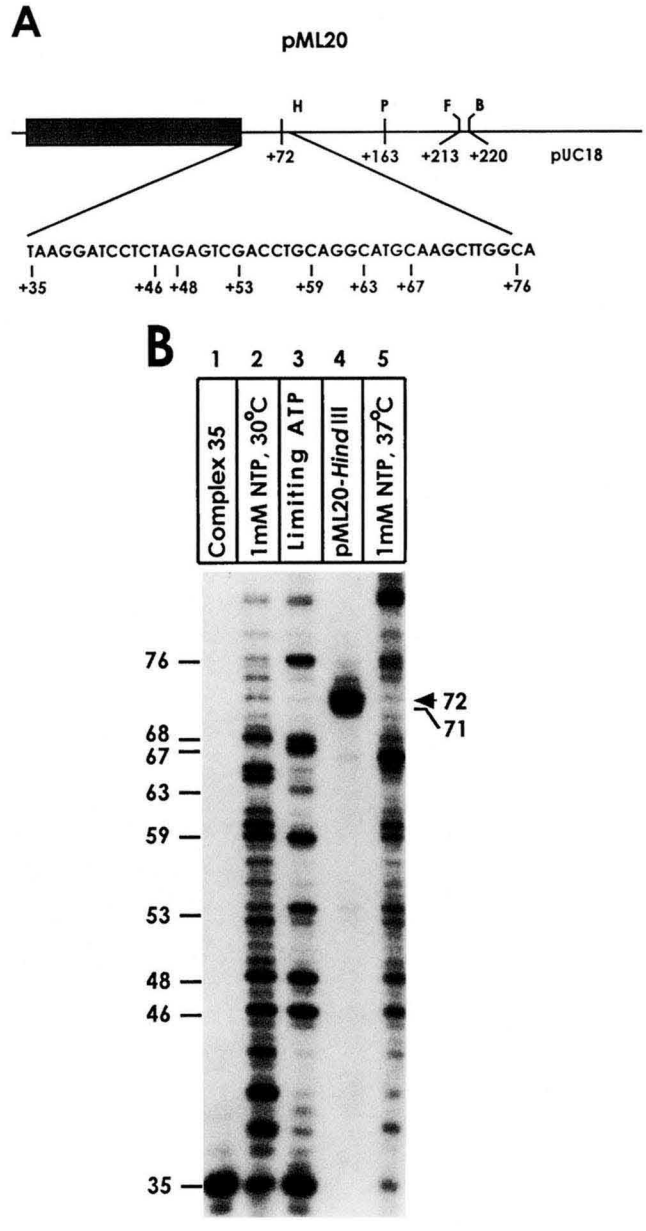

FIG. 2. **The pML20 template and high resolution mapping of RNAs generated by run-off transcription.** *A*, a portion of the pUC 18-based pML20 template is shown. The *stippled box* represents DNA derived from the adenovirus 2 genome containing the major late promoter from −174 to +35 relative to the start site of transcription. Restriction sites within the vector downstream of the promoter are shown (*H*, *Hin*dIII; *P*, *Pvu*II; *F*, *Fsp*I; *B*, *Bgl*I). The sites of template strand cleavage relative to the start site of transcription are indicated. The non-template strand from +35 to +77 with sites of prolonged RNA polymerase II residence time during A-limitation transcription reactions is also shown. *B*, *lane 4* contains RNAs generated by transcription to the end of the *Hin*dIII-linearized template. The length of the predominant run-off transcripts are indicated at the *right*. The *solid arrowhead* indicates the end of the template strand. Uniform RNA (*lanes 2* and *5*) and A limitation (*lane 3*) ladders were generated on circular pML20 template as described under "Materials and Methods." The RNAs derived from RNA polymerase that had paused prior to sites of A incorporation and U35 complex are indicated at the left. The purified transcripts were resolved on an 8% (acrylamide/bisacrylamide (19:1)) sequencing gel stopped after the bromphenol blue had reached 35 cm.

2*A*). In some instances, DNA sequencing ladders were also used to assist in the analysis. Shown in Fig. 2*B* are the results we obtained with pML20-*Hin*dIII complexes. RNA markers were generated with circular pML20 templates transcribed for various times in the presence of saturating levels of NTPs (see "Materials and Methods"). By pooling RNAs purified from several of these reactions, we were able to generate RNA populations with bands at nearly every base and with identical sequence to the test RNAs (Fig. 2*B*, *lanes 2* and *5*). To aid in assigning lengths within these relatively uniform RNA ladders, we also generated a second set of RNA markers in which we mixed elongation reactions performed for various times with one NTP limiting. An ATP-limitation ladder is shown in *lane 3* of Fig. 2*B*; transcripts associated with RNA polymerases that have paused prior to sites of A incorporation are more abun-

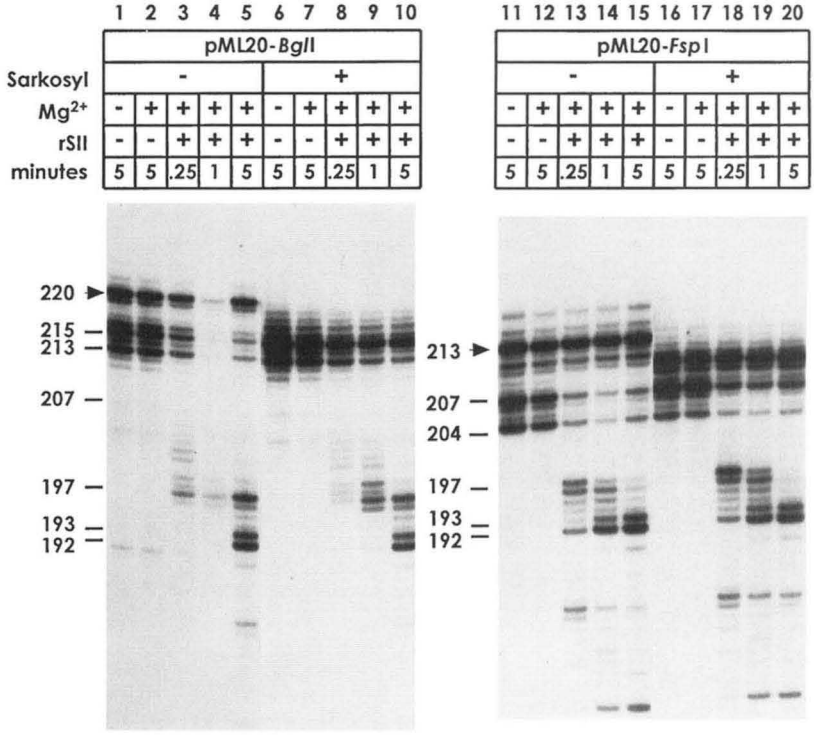

FIG. 3. **SII-facilitated transcript cleavage by complexes arrested near the end of linear templates.** Transcription to the end of *Bgl*I- and *Fsp*I-digested templates was performed in the presence (*lanes 6–10* and *16–20*) or absence (*lanes 1–5* and *10–15*) of 0.3% Sarkosyl. After the removal of NTPs and Sarkosyl by gel filtration, pML20-*Bgl*I (*lanes 1* and *6*) and pML20-*Fsp*I (*lanes 11* and *16*) complexes were supplied with 7.8 mM Mg^{2+} and 1.6 μg/ml rSII as indicated. After incubation at 37 °C for the times shown, the purified RNAs were resolved on a single 5% (acrylamide/bisacrylamide (19:1)) sequencing gel (29 × 80 cm) run until the xylene cyanol had reached 75 cm. Transcript sizes are indicated at the *left* of each template set. The *solid arrowhead* indicates the end of the template strand. The RNAs in *lane 4* were partially lost during purification.

dant in this case. The sizes of these RNAs are given in Fig. 2*B* and are also indicated in Fig. 2*A*. The data in Fig. 2*B* clearly show that transcripts from the *Hin*dIII-linearized template were 71 or 72 nt (*lane 4*), which means that transcription proceeded to within 1 base of the end of the template strand in this case. The sequence of the end of the *Hin*dIII-cut template strand, with the last two transcribed bases underlined, is 3′-CGTTCGA-5′. The RNAs generated by transcription to the end of the template strand (+72) are indicated in Fig. 2 and subsequent figures by a *solid arrowhead*. The distributions of RNAs generated by transcription on the various linear templates are summarized in Fig. 6.

We also tested a second linear template that contained a 4-nt 5′-overhang, namely the pMB5 plasmid (15) cut with *Eco*RI. In this case as well, transcription proceeded to within 2 nucleotides of the end of the template (data not shown). It is important to note that neither the pML20-*Hin*dIII nor the pMB5-*Eco*RI complexes were capable of SII-facilitated transcript truncation (data not shown). This is in contrast to complexes generated on 3′-overhung (Fig. 1) and blunt-ended templates (see below). Since polymerases appeared to be inactive after transcription to the end of the *Hin*dIII-cut template, we tested for the possibility that the ternary complexes were no longer intact. Using gel filtration with Bio-Gel A-5m, we determined that the 71- and 72-nt RNAs generated on pML20-*Hin*dIII complexes were not retained in ternary complexes since they eluted later than the DNA template (data not shown). Thus, it appears that transcription to within a few nucleotides of the end of the template strand causes destabilization of the ternary complex. If this is a general feature of ternary complexes, irrespective of the type of template end, one would predict that the SII-responsive pML20-*Bgl*I complexes must have halted transcription prior to the end of the template strand.

Transcription to the End of a Variety of Linear Templates and Subsequent SII-facilitated Cleavage—High resolution mapping of pML20-*Bgl*I transcripts showed that RNAs produced by transcription to the end of this linear template ranged from 212 to 220 nt in length (Fig. 3, *lane 1*). Complexes that had transcribed to within 3 nt of the end of the template strand (218–220-nt RNAs) did not respond to SII treatment, whereas pML20-*Bgl*I complexes paused 4–8 nt from the end of the template strand (213–216-nt RNAs) underwent transcript truncation in the presence of SII (*lanes 3–5*; note that the majority of the sample run in *lane 4* was lost). Transcription to the end (+213) of the blunt-ended *Fsp*I-digested template generated a slightly different pattern of transcription products (*lane 11*). However, we again observed that complexes that had transcribed to within 3–4 nt of the end of the template strand failed to cleave their transcripts when supplied with SII, whereas complexes that had halted transcription 5–10 nt from the template end underwent SII-facilitated transcript cleavage (*lanes 13–15*). Increasing the concentration of SII and/or the incubation time did not significantly alter these results (data not shown, but see Fig. 5). Also, increasing the duration of the initial transcription reaction gave only a slight increase in the number of complexes that completed transcription to the end of the template strand (data not shown). Thus, it is not likely that the observed heterogeneity of complexes reflects insufficient chase time.

It is important to note that the polymerases that stop upstream of the template end are not simply impeded by other DNA binding proteins. Most proteins that bind DNA nonspecifically are removed by the Sarkosyl rinse before the chase step. Furthermore, when we added an additional treatment in which the Sarkosyl-rinsed early elongation complexes were treated with 100 μg/ml heparin and repurified by gel filtration, there was no change in the proportion of polymerases paused upstream of the template end or the locations at which pausing took place (data not shown). The distribution of pML20-*Bgl*I complexes shifted toward complexes halted 5–8 nt prior to the end of the template strand in chase reactions containing limiting CTP concentrations (data not shown). Thus, it is not surprising that nearly all uniformly labeled pML20-*Bgl*I complexes remained SII responsive (Fig. 1, *lanes 5* and *6*). Run-off RNAs generated with pML20-*Bgl*I and pML20-*Fsp*I complexes were too large to separate from the template by gel filtration, which precluded a direct test for the release of RNAs from these transcription complexes. However, based on the data obtained

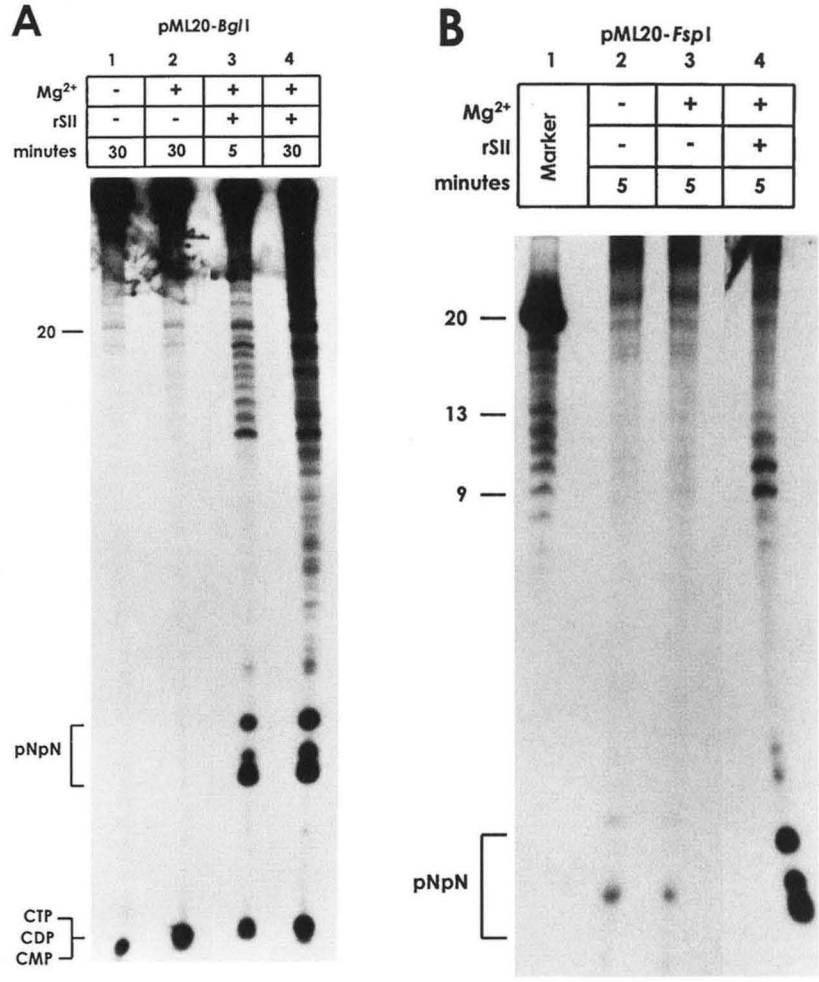

FIG. 4. **The cleavage products generated from SII-facilitated transcript truncation by pML20-*Bgl*I complexes.** *A*, one-third of the reactions described in Fig. 1 were resolved on 16 × 23-cm sequencing gels (28%, acrylamide/bisacrylamide (25:3)). The bromphenol blue dye marker was run 12.5 cm. U20 complex RNA and radiolabeled pCpC were run as standards. The locations of notable RNAs are shown at the *left*. *B*, uniformly labeled pML20-*Fsp*I complexes (*lane 2*) were incubated for 5 min at 37 °C after the addition of 7.8 mM Mg^{2+} and 1.6 μg/ml rSII as indicated. The purified RNAs were fractionated as described in *panel A*. Labeled RNA from U20 complex was used as size markers (*lane 1*), and transcript lengths are shown at the *left* as are the location of the pNpN cleavage products.

with pML20-*Hin*dIII complexes, it seems reasonable to suppose that the transcripts that extended to within a few bases of the ends of the template strand on the *Bgl*I- and *Fsp*I-cut templates were not retained in ternary complex.

In all of the run-off experiments, we also observed a minor population of RNAs that were 1–5 nt longer than would be expected based on the length of the template strand (Fig. 2*B*, *lane 4*; Fig. 3, *lanes 1–5* and *11–15*). We are not certain of the origin of these RNAs, but in this context two points are worth noting. First, other *in vitro* transcription systems have also yielded run-off RNAs that were longer than the distance between the initiation site and the end of the template strand (19, 20). Also, Johnson and Chamberlin (21) have recently observed that binary complexes of RNA polymerase II and RNA are capable of adding a small number of bases to the 3′-end of the RNA.

The rate of transcript elongation is essentially identical during transcription on linear templates in the presence or absence of 0.3% Sarkosyl (Ref. 4 and data not shown). However, a low level of Sarkosyl affects the ability of the known elongation factors to function and inhibits initiation by RNA polymerase II (4, 6, 22, 23), presumably by disrupting transcription factor-polymerase interactions. The results in Fig. 3, *lanes 6* and *16*, show that the distribution of transcription complexes at the ends of *Bgl*I- and *Fsp*I-cleaved templates was altered when run-off transcription was performed in the presence of 0.3% Sarkosyl (*lanes 6* and *16*). The detergent appeared to inhibit continued transcription by complexes that would normally transcribe to near the end of the template (compare pML20-*Bgl*I complexes in *lanes 1* and *6* and pML20-*Fsp*I complexes in *lanes 11* and *16*). The inability of ternary complexes to transcribe to the end of linear templates in the presence of Sarkosyl appears to be a general phenomenon, as the vast majority of pML20-*Hin*dIII complexes halted transcription 7 nt prior to the end of the template strand under these conditions (data not shown). Sarkosyl had little effect on the accumulation or distribution of complexes that remained SII responsive (*lanes* 8–10 and 18–20). However, the ability of normally responsive complexes to perform the truncation reaction upon SII addition was slightly reduced after exposure to Sarkosyl.

As indicated above, SII-facilitated transcript cleavage by bona fide arrested complexes occurs in a large (7–17 nt) cleavage increment, detectable by monitoring transcript shortening at very early times after SII addition. The initial cleavage increment with pML20-*Bgl*I and pML20-*Fsp*I complexes also appeared to be large, suggesting the interesting possibility that these complexes were functionally arrested. After the 0.25-min incubation with SII, RNAs of 196–204 nt are evident with a concomitant decrease in original pML20-*Bgl*I complex transcripts (Fig. 3, *lane 3*). Similarly, 192–197-nt RNAs accumulate in the pML20-*Fsp*I reaction (*lane 13*). The preference for particular sites for SII-facilitated cleavage by pML20-*Bgl*I (*lane 8*) and pML20-*Fsp*I (*lane 18*) complexes was not altered by generating these complexes in the presence of Sarkosyl.

To determine directly the SII-facilitated cleavage increment, we uniformly labeled pML20-*Bgl*I and pML20-*Fsp*I complexes and resolved the SII-facilitated cleavage products on 28% sequencing gels (Fig. 4, *A* and *B*). As expected, gel filtration of pML20-*Bgl*I complexes removed all but a low level of the labeling nucleotide (Fig. 4*A*, *lane 1*). The 5-min SII-facilitated cleav-

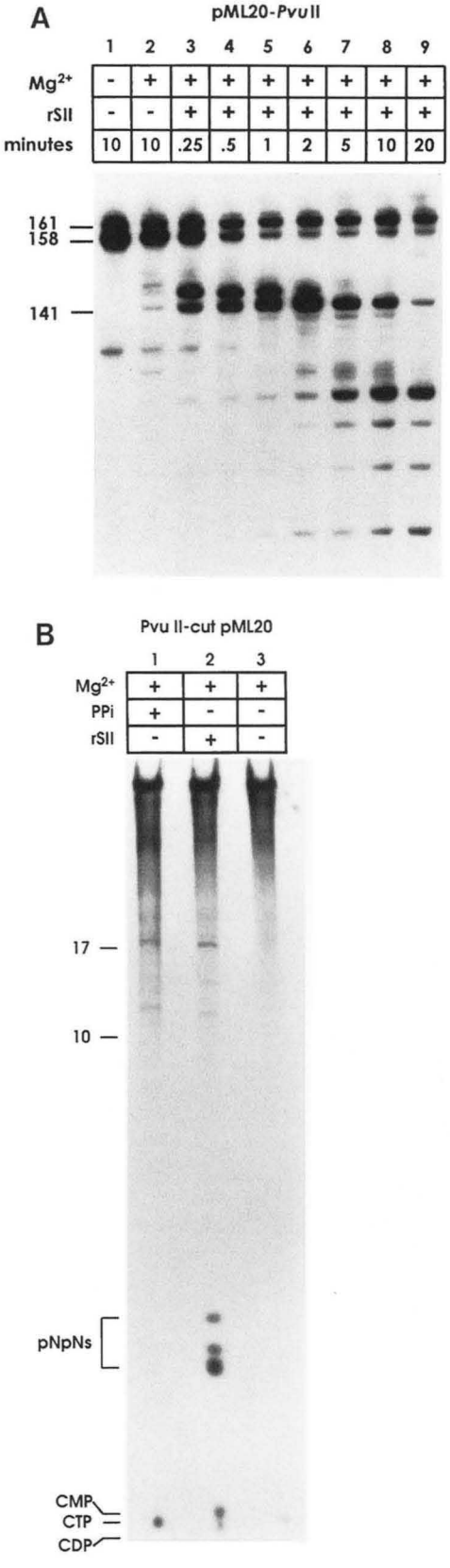

FIG. 5. **SII-facilitated transcript truncation by ternary complexes arrested at the end of *Pvu*II-digested template.** *A*, pML20-*Pvu*II complexes (*lane 1*) were supplemented with 7.8 mM Mg^{2+} and 1.5 μg/ml rSII as shown. After incubation at 37 °C for the times indicated, the RNAs were purified and resolved on a 10% (acrylamide/bisacrylamide (19:1)) sequencing gel. Transcript sizes are shown at the *left*. *B*, uniformly labeled pML20-*Pvu*II ternary complexes were supplied with 8 mM Mg^{2+} (*lanes 1–3*), 3.2 μg/ml rSII (*lane 2*) and 2 mM PPi (*lane 1*) as indicated. After a 15-min incubation at 37 °C, the RNAs were purified and resolved by electrophoresis as described in the Fig. 4 legend. Marker positions are indicated on the *left* (markers not shown).

age reaction, which generated complexes containing primarily 192 and 193 nt RNAs (see Fig. 1, *lane 5*; additional data not shown), clearly liberated both large (approximately 10–17 nt) and dinucleotide cleavage products. A time course of SII-facilitated transcript truncation demonstrated that the maximum level of large cleavage products was achieved within minutes, whereas the pNpNs products continued to accumulate over time (data not shown). This is consistent with the model that arrested complexes regain elongation competence after the initial cleavage. The increase in the dinucleotide products was obvious after the 30-min incubation (*lane 4*). We also observed an increase in the level of radiolabeled mononucleotide (presumably CMP) in reactions supplied with SII (*lanes 3* and *4*). However, a similar level of this product was present in the control reaction receiving Mg^{2+} but no SII (*lane 2*). Essentially no other cleavage products were observed in the control reaction. Therefore, the accumulation of this level of radiolabeled mononucleotide was not a property of SII-facilitated cleavage. SII-facilitated cleavage products generated by uniformly labeled pML20-*Fsp*I complexes are displayed in Fig. 4*B*, *lane 4*. The 5-min SII-facilitated cleavage reaction liberated large (approximately 9–11 nt) fragments and pNpNs. The sizes of the large cleavage products produced by both pML20-*Bgl*I complexes (10–17 nt) and pML20-*Fsp*I complexes (9–11 nt) were consistent with the initial increment of transcript shortening observed in Fig. 3.

We observed similar properties with run-off transcription complexes generated on *Pvu*II-digested pML20 templates. A time course of SII-facilitated cleavage using 5′-end-labeled pML20-*Pvu*II complexes is shown in Fig. 5*A*. The major RNA species generated during run-off transcription were 158 and 161 nt in length (*lane 1*); a complete run-off RNA on this template would be 163 nt long. The time course (*lanes 3–9*) demonstrated that ternary complexes that halted transcription 5 nt prior to the end of the template strand appeared to cleave their nascent transcripts in a large increment, whereas complexes transcribing to within 2–3 nt of the end of the template did not respond to SII treatment. It is important to note that both the pattern of run-off RNAs and the initial cleavage increment were different on the two blunt-ended templates (compare Figs. 3 and 5). For example, *Fsp*I-cleaved templates showed an initial cleavage increment of 9–12 nt, while *Pvu*II-cut templates gave a 12–17-nt initial increment. These data strongly suggest that nucleic acid sequence influences the exact site(s) of both transcription cessation and subsequent SII-facilitated cleavage.

We have recently demonstrated that in the presence of pyrophosphate (PPi), complexes arrested within bona fide intrinsic arrest sites cleave their nascent RNAs at essentially the same sites used by SII-facilitated cleavage (24). Interestingly, complexes arrested at the ends of linear templates also truncate their nascent transcripts in the presence of PPi at sites similar to those utilized during SII-facilitated truncation (Fig. 5*B* and additional data not shown). To illustrate, we show the products released during SII-facilitated and PPi-mediated transcript truncation by pML20-*Pvu*II complexes. Uniformly labeled pML20-*Pvu*II complexes (Fig. 5*B*) were incubated for 15 min with either SII (*lane 2*) or PPi (*lane 1*), and the RNAs were fractionated on a 28% sequencing gel. As expected, in the presence of SII, large RNAs were liberated; we presume that the dinucleotide products were liberated after the cleavages that generated the long RNAs, based on additional time course data not shown here. In the presence of 2 mM PPi, a similar

distribution of large RNAs was observed (*lane 1*). The large cleavage products produced in the presence of PPi should have 5′-triphosphate termini and thus should have slightly different mobilities than the 5′-monophosphate fragments generated by SII-facilitated cleavage (see Ref. 24; compare *lanes 1* and *2* in Fig 5*B*). Also as expected, subsequent PPi-mediated transcript truncation via conventional pyrophosphorolysis generated NTPs (Fig. 5*B*, *lane 1*, and additional time course data not shown). These results suggest that the large PPi-mediated cleavage increment is not a unique feature of ternary complexes that become arrested within bona fide arrest sites. Rather, pyrophosphorolysis at locations far upstream of the nascent RNA 3′-end appears to be a general feature of complexes that have adopted an "arrested" configuration.

DISCUSSION

We have investigated in detail the fates of RNA polymerase II ternary complexes that reach the ends of various linear templates. We find that such complexes partition into two groups: those that transcribe to within 3–4 bases of the end of the template strand and those that stop transcription 5–10 bases prior to the end of the template strand. The former group does not respond to SII. These polymerases appear to have terminated transcription, as one might have expected. However, polymerases that stop 5–10 bases upstream of the end of the template strand remain responsive to SII. Remarkably, these SII-responsive complexes closely resemble complexes arrested during transcription of circular templates, based on a number of criteria. First, the initial SII-facilitated cleavage for complexes halted near the end of the template strand removes 9–17 nt fragments; subsequent cleavages by these complexes occur primarily in the dinucleotide increments characteristic of elongation-competent complexes. Both of these properties exactly parallel those of complexes that have become arrested while transcribing circular DNA templates (10). Furthermore, arrest on circular templates may be relieved by pyrophosphorolytic transcript cleavage, which like SII-facilitated truncation generates 7–17 nt fragments from arrested complexes (24). Pyrophosphate treatment also causes transcript cleavage by polymerases halted 5–10 bases upstream of the template end; in this case as well, cleavage occurs in a large (10–17 nt) increment (Fig. 5*B* and data not shown). Finally, transcription at suboptimal CTP concentrations on the *Bgl*I and *Pvu*II templates leads to the majority of ternary complexes arresting 5–10 nt upstream of the end of the template strand, just before the incorporation of C residues into the transcript (Fig. 1 and data not shown; see also Fig. 6). Similar effects of NTP limitations on arrest within intrinsic arrest sites are well documented (2, 8, 25, 26). Thus, SII-responsive ternary complexes paused near the ends of templates have all the characteristics previously described for complexes arrested at particular DNA sequences during transcription of circular templates.

Explanations of transcriptional arrest have focused on the influence of particular template sequences on the arrest process. Arrest generally correlates with the transcription of an A-rich segment of the template strand, such that arrested complexes contain transcripts ending in 3–5 consecutive U residues (a structure we will refer to as the U tail). Arrest sites also contain other features that influence the loss of elongation competence, since not all A-rich segments of the template strand cause arrest (2, 10). Some arrest sites contain sequences immediately upstream of the A-rich region that encode self-complementary sections in the transcript (27, 28). Another class of arrest sites appear to contain DNA segments at or near the point of arrest, which are capable of adopting a bent configuration (1, 2, 7, 10, 29). The best studied of the latter class of sites contains two runs of A residues on the template strand,

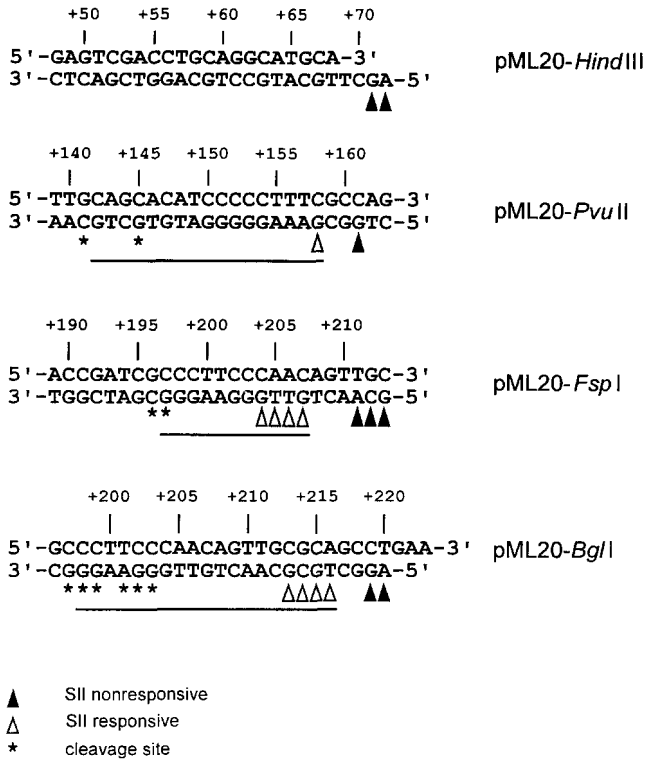

FIG. 6. **Summary of RNA patterns generated by run-off transcription on linear DNA templates and after very early times during subsequent SII-facilitated transcript truncation.** The last 25–30 base pairs of the linear templates used in this study are shown; the *numbers* above the sequences indicate the distance downstream of the transcription start site. *Arrowheads* indicate the major locations of 3′-ends generated in the initial transcription reaction. *Solid arrowheads* indicate SII-nonresponsive complexes; *open arrowheads* indicate SII-responsive complexes. The apparent location of new 3′-ends generated by transcript cleavage are indicated by the *asterisks*. The maximum sizes of the segments removed from the transcripts are indicated by the *bars* below the sequences.

spaced about 10 base pairs apart (2, 7, 10). Arrest occurs within the first A run, but the integrity of the downstream A run is required for arrest (2). These observations led to the proposal (2) that a bent template configuration is one of the signals for transcriptional arrest. This view was further refined by Reines, (7) who suggested that transcription through intrinsic arrest sites in the bent configuration may actually impede elongation, thereby increasing the likelihood of complexes to adopt the arrested configuration. This model is consistent with the observation that extending the residence time of RNA polymerase at potential sites of arrest increases the number of complexes that become blocked in elongation (2, 8, 25, 26).

In an initial exploration into the effects of residence time on transcriptional arrest of complexes containing U tails, we have shown that 3, 4, or 5 consecutive A residues in the template strand provide essentially no barrier to RNA polymerase II in the presence of excess NTPs. However, if the polymerase is forced to stall after the synthesis of, for example, a 5-nt U tail, about half of the ternary complexes become arrested (10). These and other observations led us to propose that some ternary complexes in which the transcript ends in a U tail may exist in equilibrium between the active and arrested states (10). These studies reinforced the notion that other template and/or RNA components in addition to transient pausing must be involved in "fixing" the arrested configuration (2, 25).

The necessity for transcript cleavage in the relief of arrest led us to speculate that arrest reflects the loss of contact between the 3′-end of the transcript and the catalytic center of

the polymerase (8). After arrest, the active site must resume transcription from a new 3′-end, which may be as far as 17 nt upstream of the original site of polymerization. Furthermore, pyrophosphorolytic transcript cleavage by arrested complexes also occurs in a large increment (7–17 nt, Fig. 5*B*; see also Ref. 24). Thus, arrest may be viewed as a major conformational change within the RNA polymerase, one consequence of which is the upstream displacement of the catalytic center of the enzyme.

It is not immediately clear how the apparent arrest of polymerases near the ends of linear templates fits into previous observations on arrest. In particular, the 3′-ends of the RNAs within complexes that arrest near the ends of templates are not necessarily U-rich (see Fig. 6). Some recent findings suggest that arrest can occur at locations other than canonical arrest sites. Reines and Mote (14) have shown that physically blocking the path of the elongating RNA polymerase with a protein can apparently cause arrest. A new model for the ternary complex explains translocation by the alternative locking and sliding of upstream and downstream DNA binding domains (30); attached to these domains are the active site and domains for RNA binding. In such a model, blocking translocation might generate strain within the polymerase, which could be relieved by an upstream movement of part of the complex. This relaxation could carry the active site upstream and thereby produce arrest. Along these lines, we may explain our current results by emphasizing the importance to the RNA polymerase of appropriate contacts with the template downstream of the site of chain elongation. As the polymerase approaches within 10 bases of the end of a linear template, it should experience some loss of downstream template contacts, since footprinting studies indicate that stalled transcription complexes protect at least 15 base pairs of template downstream of the site of chain elongation from attack by DNase I (31, 32). Thus, arrest at the end of linear templates may arise because partial or complete loss of downstream DNA contacts prevents translocation, which in turn causes a conformational change in the RNA polymerase that forces the 3′-end of the transcript away from the active site. This raises the possibility that loss or destabilization of downstream DNA contacts during transcription through intrinsic arrest sites, perhaps while they are in the bent configuration, might favor the arrested configuration. Exonuclease III footprinting studies of an arrested complex within the histone H3.3 intrinsic arrest site have demonstrated that reestablishing the elongation competency of these complexes via transcript truncation is accompanied by a slight forward extension of the downstream exonuclease III footprint boundary (33).

The simple linkage of the loss of downstream contacts with arrest does not explain the second pathway that we observed at the end of templates, in which polymerases synthesize RNA to the end of the template strand, with apparently catastrophic consequences, instead of halting 5–10 bases upstream. However, even at the strongest arrest sites, polymerases partition between arrest and elongation competence when excess NTPs are present (1, 4, 8, 17, 25, 26). Thus, it is not unexpected that some polymerases do not arrest 5–10 bases upstream of the end of the template strand but instead continue transcription.

We have demonstrated that the detergent Sarkosyl affects the point at which SII-nonresponsive complexes stop transcribing near the ends of linear templates (Fig. 3). The sites at which SII-responsive complexes arrest and the extent to which arrest occurs on a particular template are largely unaffected by Sarkosyl (Fig. 3). It is important to emphasize that Sarkosyl has little effect on either the rate of transcript elongation or the arrest frequency at intrinsic arrest sites (Refs. 4, 6, 17, and 25 and data not shown). Thus, in the absence of appropriate polymerase-template interactions, Sarkosyl may destabilize the ternary complex. Interestingly, under certain circumstances, RNA polymerase II-RNA binary complexes can perform a limited, non-templated polymerization reaction that is sensitive to Sarkosyl (21).

In summary, we have shown that RNA polymerase II can fail to run off the end of linear templates by falling into a state of transcriptional arrest just before reaching the end of the template. This observation broadens the definition of the circumstances that will cause arrest. It reinforces the idea (see also Refs. 2, 10, 25) that elements other than the sequence of the nascent RNA 3′-end are important for continued transcription by the elongation complex. Achieving a detailed picture of the RNA polymerase II ternary complex and the ways in which it may fail during RNA chain elongation will be crucial to eventually understanding the control of gene expression during elongation.

Acknowledgment—We thank Robert Landick for the generous gift of purified recombinant human SII.

REFERENCES

1. Reines, D., Wells, D., Chamberlin, M. J., and Kane, C. M. (1987) *J. Mol. Biol.* **196,** 299–312
2. Kerppola, T. K., and Kane, C. M. (1990) *Biochemistry* **29,** 269–278
3. Resnekov, O., and Aloni, Y. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 12–16
4. Reines, D., Chamberlin, M. J., and Kane, C. M. (1989) *J. Biol. Chem.* **264,** 10799–10809
5. Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991) *Mol. Cell. Biol.* **11,** 1195–1206
6. Izban, M. G., and Luse, D. S. (1992) *J. Biol. Chem.* **267,** 13647–13655
7. Reines, D. (1992) *J. Biol. Chem.* **267,** 3795–3800
8. Izban, M. G., and Luse, D. S. (1992) *Genes & Dev.* **6,** 1342–1356
9. Reines, D., Ghanouni, P., Li, Q. Q., and Mote, J., Jr., (1992) *J. Biol. Chem.* **267,** 15516–15522
10. Izban, M. G., and Luse, D. S. (1993) *J. Biol. Chem.* **268,** 12874–12885
11. Izban, M. G., and Luse, D. S. (1993) *J. Biol. Chem.* **268,** 12864–12873
12. Wang, D. G., and Hawley, D. K. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 843–847
13. Guo, H., and Price, D. H. (1993) *J. Biol. Chem.* **268,** 18762–18770
14. Reines, D., and Mote, J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 1917–1921
15. Jacob, G. A., Kitzmiller, J. A., and Luse, D. S. (1994) *J. Biol. Chem.* **269,** 3655–3663
16. Yoo, O., Yoon, H., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Agarwal, K. (1991) *Nucleic Acids Res.* **19,** 1073–1079
17. Izban, M. G., and Luse, D. S. (1991) *Genes & Dev.* **5,** 683–696
18. Coppola, J. A., Field, A. S., and Luse, D. S. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 1251–1255
19. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, R., and Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056
20. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15,** 8783–8798
21. Johnson, T. L., and Chamberlin, M. J. (1994) *Cell* **77,** 217–224
22. Hawley, D. K., and Roeder, R. G. (1985) *J. Biol. Chem.* **260,** 8163–8172
23. Kephart, D. D., Marshall, N. F., and Price, D. H. (1992) *Mol. Cell. Biol.* **12,** 2067–2077
24. Rudd, M. D., Izban, M. G., and Luse, D. S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 8057–8061
25. Wiest, D. K., Wang, D., and Hawley, D. K. (1992) *J. Biol. Chem.* **267,** 7733–7744
26. Wiest, D. K., and Hawley, D. K. (1990) *Mol. Cell. Biol.* **10,** 5782–5795
27. Kessler, M., Ben-Asher, E., and Aloni, Y. (1989) *J. Biol. Chem.* **264,** 9785–9790
28. Resnekov, O., Kessler, M., and Aloni, Y. (1989) *J. Biol. Chem.* **264,** 9953–9959
29. London, L., Keene, R. G., and Landick, R. (1991) *Mol. Cell. Biol.* **11,** 4599–4615
30. Chamberlin, M. J. (1994) *The Harvey Lectures*, pp. 1–21, Wiley-Liss, Inc., New York
31. Linn, S. C., and Luse, D. S. (1991) *Mol. Cell. Biol.* **11,** 1508–1522
32. Rice, G. A., Chamberlin, M. J., and Kane, C. M. (1993) *Nucleic Acids Res.* **21,** 113–118
33. Gu, W., Powell, W., Mote, J., Jr., and Reines, D. (1993) *J. Biol. Chem.* **268,** 25604–25616