MOLECULAR AND CELLULAR BIOLOGY, Nov. 1991, p. 5398–5409
0270-7306/91/115398-12$02.00/0

Vol. 11, No. 11

A Heat-Labile Factor Promotes Premature 3′ End Formation in Exon 1 of the Murine Adenosine Deaminase Gene in a Cell-Free Transcription System

JEFFREY W. INNIS[1]†* AND RODNEY E. KELLEMS[1,2]

Institute for Molecular Genetics[1] and Verna and Marrs McLean Department of Biochemistry,[2] Baylor College of Medicine, Houston, Texas 77030

Received 7 May 1991/Accepted 30 July 1991

An elongation block to RNA polymerase II transcription in exon 1 is a major regulatory step in expression of the murine adenosine deaminase (ADA) gene. Previous work in the laboratory identified abundant short transcripts with 3′ termini in exon 1 in steady-state RNA from injected oocytes. Using a cell-free system to investigate the mechanism of premature 3′ end formation, we found that polymerase II generates prominent ADA transcripts approximately 96 to 100 nucleotides in length which are similar to the major short transcripts found in steady-state RNA from oocytes injected with ADA templates. We have determined that these transcripts are the processed products of 108- to 112-nucleotide precursors. Precursor formation is (i) favored in reactions using circular templates, (ii) not the result of a posttranscriptional processing event, (iii) sensitive to low concentrations of Sarkosyl, and (iv) dependent on a factor(s) which is inactivated in crude extracts at 47°C for 15 min. The cell-free system will allow further characterization of the template and factor requirements involved in the control of premature 3′ end formation by RNA polymerase II.

Transcriptional regulation of gene expression involves multiple steps related to initiation, elongation, and termination. Production of a complete primary transcript after initiation depends on the ability of RNA polymerase to traverse the transcription unit. Through modification of the elongating polymerase complex or the secondary structure of the RNA, and via signals in the DNA or RNA, control is achieved over whether RNA polymerase synthesizes a complete mRNA. Insight into the basic mechanisms of this level of control of gene expression has come primarily from research into prokaryotic transcription termination and antitermination. It is now well established that synthesis of a primary transcript in eukaryotes also is regulated by postinitiation transcriptional signals.

Experimental evidence in both eukaryotic and prokaryotic systems suggests that regulation of elongation occurs at multiple steps often very early after initiation. For example, a preinitiated transcription complex may pause at specific sites near the 5′ end of a gene and wait for additional instructions before continuing on with transcription. This type of process is known to play a major role in the transcription of some *Drosophila* genes (37, 38). For the *Drosophila hsp70* gene, the polymerase incorporates approximately 25 nucleotides (nt) and then pauses. Upon heat shock, the polymerase is released and further elongation of the nascent RNA chain ensues. Thus, completion of the mRNA is achieved by release of a block to the previously initiated polymerase complex. In phage lambda, transcription of the late genes involves a pause by RNA polymerase at +16 relative to the transcription initiation site, whereupon Q protein interacts with the transcription complex together with NusA at the *qut* site (11, 12), allowing transcription to proceed through downstream terminators. Thus, modification of the transcription complex early after initiation may be a rate-limiting step(s) in the formation of a processive or a termination-prone RNA polymerase. The former state allows the polymerase to read through potential termination sites or antiterminate. Antitermination by N protein of phage lambda is achieved by a similar mechanism and has been well studied (24). In human immunodeficiency virus type 1 (HIV-1), Tat protein, in conjunction with other cellular proteins, modifies the polymerase complex to be highly efficient at elongation (7, 18, 22, 23, 27, 28, 42, 48, 50). The steps in this process of antitermination are unknown.

Additional mechanisms to control postinitiation transcriptional elongation include attenuation and physical obstruction to polymerase progression (33, 39, 41, 52). Attenuation in many biosynthetic operons of bacteria is a well-understood mechanism of postinitiation transcription control and is based on the modification of RNA secondary structure. Antitermination in prokaryotes via this mode relies on ribosomes or, in a few cases, nonribosomal proteins to disrupt the ability of the transcribed RNA to form a specific secondary structure conducive to termination. Physical obstruction has been described in prokaryotes to inhibit RNA polymerase progression at specific sites. Thus, transcription termination and antitermination appear to control the expression of numerous genes by diverse mechanisms. The events described here occur within a short distance of the transcription initiation site and involve signals in the DNA, RNA, or RNA polymerase complex together with specific *trans*-acting factors.

Eukaryotic genes which have been shown to be transcriptionally regulated at the level of postinitiation transcriptional elongation include the c-*myc* (3, 4), L-*myc* (19), c-*myb* (2, 49), c-*mos* (29), c-*fos* (10, 21, 40), adenosine deaminase (ADA) (5, 6, 25, 34), epidermal growth factor receptor (13), histone H3.3 (44), *Drosophila hsp70*, *hsp26*, β1-tubulin, glyceraldehyde-3-phosphate dehydrogenase, polyubiquitin, and copia (37, 38), HIV-1 (7, 18, 22, 23, 27, 28, 42, 48, 50), simian virus 40 (15, 16), polyomavirus (45), and minute virus

* Corresponding author.
† Present address: Department of Human Genetics, University of Michigan Medical School, Medical Science II M4708, Ann Arbor, MI 48109-0618.

of mice (1, 35) genes and the adenovirus major late promoter (MLP) transcription unit (14, 26, 51). For a recent review of this literature, see Spencer and Groudine (46). For many of these genes, data from nuclear run-on assays suggested that transcription elongation beyond the first exon is governed by metabolic, tissue-specific, or developmental signals. Through the use of oocyte injection experiments, virus-infected cells, and in vitro transcription assays, some of the *cis* elements and *trans*-acting factors involved in postinitiation transcriptional elongation control in these genes have been elucidated, but the details of these events in eukaryotes are not well understood.

We have been studying the occurrence of conditional elongation of transcription through exon 1 of the murine ADA gene. In cells or tissues with high levels of ADA mRNA and activity, nascent transcripts detected by nuclear run-on assays are located throughout the gene. In tissues with low ADA mRNA and activity, nascent transcripts are clustered in exon 1 (6, 25). The same is true for the human ADA gene (5). Thus, increased expression of ADA mRNA is concurrent with transcriptional elongation beyond the first exon. Short RNA molecules with 3′ ends which terminate in the region of elongation block have not been detected in cultured cells. However, as with the c-*myc* gene, abundant short transcripts from the murine and human ADA genes accumulate in injected oocytes and the 3′ ends of these short transcripts map within exon 1 and correspond to the region of elongation block seen in nuclear run-on experiments (5, 34). These data suggest that the ADA and c-*myc* genes regulate transcription by conditional premature 3′ end formation within exon 1. For the murine ADA gene, the 3′ terminus of the major short transcript is at +96 relative to the transcription initiation site in RNA from injected oocytes (34). Posttranscriptional cleavage of longer ADA transcripts did not account for the generation of short transcripts, suggesting the possibility of cotranscriptional cleavage or termination as the mechanism for their formation (33a). Deletion of sequences downstream of +96 from +104 to +168 severely reduced the accumulation of the 96-nt short transcript in oocytes; however, slightly longer (+50 to 200 nt) RNA molecules became more abundant. Therefore, sequences downstream of the 3′ terminus of the major short transcript were necessary for accumulation of the 96-nt RNA. Moreover, transcription complexes that fail to generate 3′ ends near +96 appear to travel only a short distance downstream and generate transcripts with other premature 3′ termini. Thus, there appears to be conservation of the general mechanism involved in premature 3′ end formation in oocytes between the mouse and human ADA genes, as has been seen for the murine and human c-*myc* genes (3, 4).

Oocyte injection experiments with murine ADA templates demonstrated accumulation of abundant short transcripts with a discrete 3′ end. To learn more about this, we have used a cell-free transcription system to reproduce the process of premature 3′ end formation in exon 1 and have investigated the mechanism by which these short transcripts are generated. We show that the short transcripts are generated under standard reaction conditions with circular templates in a HeLa nuclear extract at sites similar to those seen in injected oocytes. Using this assay, we have delineated the conditions which augment or inhibit premature 3′ end formation and present evidence that a heat-labile factor(s) is necessary for the recognition of these signals in the first exon of the murine ADA gene. The assay also provides a convenient method for the purification of the factor(s) which promotes short transcript generation. The results support the hypothesis that efficient transcription through ADA exon 1 involves regulation of a transcription arrest signal. Further identification of the relevant *cis*-acting sequences and factors which influence this event in the ADA gene can be achieved by using the cell-free system, and this system also should be helpful in the identification of key cellular factors involved which promote complete mRNA synthesis.

MATERIALS AND METHODS

Cells and nuclear extract. HeLa S3 cells in log-phase suspension culture (250 ml) were obtained from the laboratory of Sophia Tsai and expanded in RPMI 1640 with 5% fetal calf serum. After 3 days, 4 to 6 liters of cells (5×10^5 to 7×10^5 cells per ml) was collected by centrifugation at room temperature in a Sorvall HG-4 rotor at 1,800 rpm for 10 min. Cells were washed with phosphate-buffered saline once at 25°C and once at 0°C. Extract preparation was performed according to Dignam et al. (9). Aliquots were frozen in a dry ice-ethanol slush and placed at −70°C for storage. The protein concentration of these extracts was usually between 4 and 6 mg/ml.

In vitro transcription reactions. All solutions were prepared by using distilled, deionized H_2O treated with diethyl pyrocarbonate and autoclaving for 45 min. Supercoiled (circular) or linear plasmid DNA molecules containing the ADA promoter, first exon, and part of the first intron were used in transcription reactions. Linearized or circular templates (300 ng) were incubated at 30°C for 15 min in the presence of 50 to 60 μg of protein (HeLa nuclear extract), buffers, salts, and nucleoside triphosphates (NTPs) in a final volume of 20 μl. The final concentrations were 10 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.9), 50 mM KCl, 4 mM $MgCl_2$, 8 to 11% glycerol, and 3 mM dithiothreitol. Transcription was started by the addition of NTPs to 200 μM ATP, CTP, and UTP and 15 μM GTP with 20 to 40 μCi of [α-^{32}P]GTP (3,000 Ci/mmol). Experiments utilizing labeled CTP or UTP with 200 μM unlabeled GTP gave identical results. No preincubation was necessary with circular templates, but in some experiments 5 min elapsed at 30°C prior to NTP addition. The reactions were stopped with 200 μl of 50 mM Tris (pH 7.4)–1% sodium dodecyl sulfate (SDS)–5 mM EDTA. The samples were extracted with 200 μl of phenol-chloroform (1:1, vol/vol), and the RNA was precipitated with 2.5 volumes of ethanol in the presence of 7.5 μg of yeast RNA and 330 mM sodium acetate. RNA was redissolved in 5 or 6 μl of a solution containing 90% formamide, 0.5× Tris-borate-EDTA, and dyes and separated by electrophoresis in 6 or 8% polyacrylamide–7 M urea gels. Gels were exposed to Kodak X-Omat AR film with or without intensifying screens. Sarkosyl (*N*-lauroyl sarcosinate; Sigma) was dissolved in diethyl pyrocarbonate-treated water. Aliquots were added to transcription reactions as described in the figure legends.

To prepare heat-treated nuclear extracts, crude nuclear extracts were incubated at 47°C for 15 min as described by Nakajima et al. (32). A fraction of crude extract enriched for transcription factor TFIID, a kind gift from M. Sawadogo and M. Szentirmay, was purified by P11 and DE52 chromatography of crude nuclear extract.

Plasmid DNA. Plasmids used in transcription reactions were subcloned pieces of genomic cosmid clones and have been described elsewhere (17). The vector DNA is pTZ19R (Pharmacia), and ADA sequences from −154 to +206 relative to the transcription initiation site in the largest subclone

(0.4) used in this study were cloned into the polylinker region. All plasmid DNA was purified from *Escherichia coli* HB101 by two rounds of CsCl density equilibrium centrifugation. All linear templates generated by restriction enzyme digestion were checked for complete cleavage by agarose gel electrophoresis and ethidium bromide staining prior to use in transcription assays. The adenovirus/ADA template was constructed by placing the adenovirus MLP from −259 to +33 in front of the murine ADA sequence from −4 to +160.

Nuclease protection analysis. RNA from transcription reactions was prepared for nuclease protection analysis or primer extension by termination of the transcription reaction with 10 volumes of a solution containing 50 mM Tris, (pH 7.4), 1% SDS, and 5 mM EDTA; proteinase K was then added, and the solution was incubated for 20 min at room temperature. Nucleic acids were purified by organic extraction and ethanol precipitation. Samples were treated with DNase I (RQ1 DNase I; Promega) for 15 min at 37°C, and RNA was recovered after organic extractions and ethanol precipitation in the presence of yeast carrier RNA. A uniformly radiolabeled RNA probe complementary to ADA mRNA was generated by T7 RNA polymerase according to the manufacturer's instructions, using [α-^{32}P]UTP (see Fig. 1C). The probe is longer than depicted in the figure because of inclusion of polylinker sequence. RNA was mixed with probe (300,000 cpm) and ethanol precipitated. The pellet was dissolved in 15 μl of 80% formamide–0.2 M NaCl–40 mM piperazine-*N*,*N*′-bis(ethanesulfonic acid) (PIPES; pH 6.4)–1 mM EDTA and denatured at 90°C for 5 min. The mixture was incubated at 47°C for 18 h. Then 350 μl of a solution containing 10 mM Tris (pH 7.5), 300 mM NaCl, 5 mM EDTA, 10 μg of RNase A per ml, and 1,900 U of RNase T_1 (Bethesda Research Laboratories) per ml was added, and the mixture was incubated at 23°C for 90 min. After digestion, 10 μl of 20% SDS and 20 μg of proteinase K were added, and the mixture was incubated at 37°C for 15 min. The RNA was extracted with phenol-chloroform and ethanol precipitated. DNA markers were used in the experiments reported. No significant difference in migration between these markers and defined RNA markers was detectable in our gel system.

Primer extension. RNA purified from transcription reactions as described above for nuclease protection was mixed with a synthetic ^{32}P-labeled oligonucleotide primer (10^5 cpm; 1 ng) complementary to the murine ADA mRNA from +13 to +37 (P_a) or +113 to +137 (P_b). Hybridization was carried out at 68°C for 4 h in a 10-μl volume of 10 mM Tris-HCl (pH 7.5)–0.3 M NaCl–1 mM EDTA. Then 90 μl of reverse transcription buffer containing 55 mM Tris-HCl (pH 7.5), 8.8 mM $MgCl_2$, 11 mM dithiothreitol, 33 mM NaCl, 400 μg of actinomycin D per ml, 550 μM dNTPs, and 8 to 10 U of avian myeloblastosis virus reverse transcriptase (Life Sciences, Inc.) was added, and the mixture was incubated at 42°C for 1 h. Extended products were precipitated with ethanol, resuspended in 95% formamide–dyes, heated at 100°C for 3 min, and separated by electrophoresis in 8% polyacrylamide–7 M urea gels.

Isolation of labeled RNA from in vitro reactions. Scaled-up in vitro reactions were performed with designated templates, and RNA products were separated by gel electrophoresis in 6% polyacrylamide–7 M urea. RNA was eluted from gel slices (using an autoradiogram as a guide) in a solution of 0.5 M ammonium acetate, 0.1% SDS, and 0.1 mM EDTA by vigorous shaking for 2 h at room temperature. Carrier yeast RNA (2.5 μg) was added, and the solution was extracted twice with phenol-chloroform and once with chloroform. The RNA was precipitated by the addition of 2.5 volumes of ethanol and then incubated at −70°C for 1 h. The precipitate was collected by centrifugation, and the RNA was dissolved in 10 to 20 μl of 10 mM Tris-HCl (pH 7.5)–1 mM EDTA. Aliquots of this preparation were used as described in the legend to Fig. 4.

RESULTS

Circular ADA templates generate short transcripts predominantly 96 to 100 nt in length. To study the transcriptional control of the murine ADA gene, we used an in vitro assay consisting of nuclear extracts from HeLa cells and cloned ADA gene fragments (17). These templates contained the ADA promoter, first exon, and part of the first intron. All ADA transcripts initiated accurately at the start site used in vivo and were transcribed by RNA polymerase II. The ADA promoter was utilized efficiently, and RNA products were radiolabeled by including [α-^{32}P]GTP in the transcription reactions. We observed that, in addition to the runoff product, a large number of transcripts from linear templates were less than full length (Fig. 1A, lane 1). The most abundant short transcripts appeared to be a broad band approximately 137 to 145 nt in length, hereafter referred to as site III transcripts. Other less than full-length transcripts were seen reproducibly (between site III and full length), but these were usually minor in comparison. Similarly, when supercoiled molecules containing these same ADA gene fragments were used to initiate transcription, identical-length short transcripts were generated (Fig. 1A, lane 2); however, the relative abundances of short and long transcripts from circular templates appeared to be different from those in the RNA products generated from linear templates (see below and Fig. 2). Since supercoiled molecules were immediately converted to relaxed circles in these extracts (data not shown), we will refer to these as circular templates. Short transcripts of predominantly 96 to 100 nt (hereafter referred to as site I transcripts) and 108 to 112 nt (hereafter referred to as site II transcripts) were easily observed in reactions with circular templates. Site I transcripts are similar in length to the major short transcripts (96 nt) which accumulate in steady-state RNA from injected oocytes (34). Generation of these short RNA molecules was independent of the radioactive NTP used in the reactions (data not shown). Additional short RNA molecules as well as much longer read-through transcripts from circular templates also were observed. Different extract preparations varied greatly with respect to the proportion of site I and II transcripts generated on circular templates. For example, Fig. 1B shows the RNA products from a reaction using the same circular template as in Fig. 1A but with an extract which generates a greater proportion of transcripts at sites I and II. It can be seen that site I and II transcripts are much more abundant than longer RNA molecules. The site I transcripts are the most abundant short transcripts in oocytes as well (34). We have identified other conditions, including dilution of the extract protein and elevation of the KCl concentration after initiation, which greatly increase the proportion of transcripts at sites I and II (data not shown). For these reasons, and because these transcripts are the most abundant in the pulse-chase analysis (see below), we have chosen to focus our attention on those transcripts at sites I, II, and III.

To characterize the 5′ and 3′ ends of ADA transcripts, we performed primer extension with RNA generated in transcription reactions from linear or circular templates (Fig. 1C). The 5′ ends of RNA generated from the two templates were identical and agreed with the transcription initiation

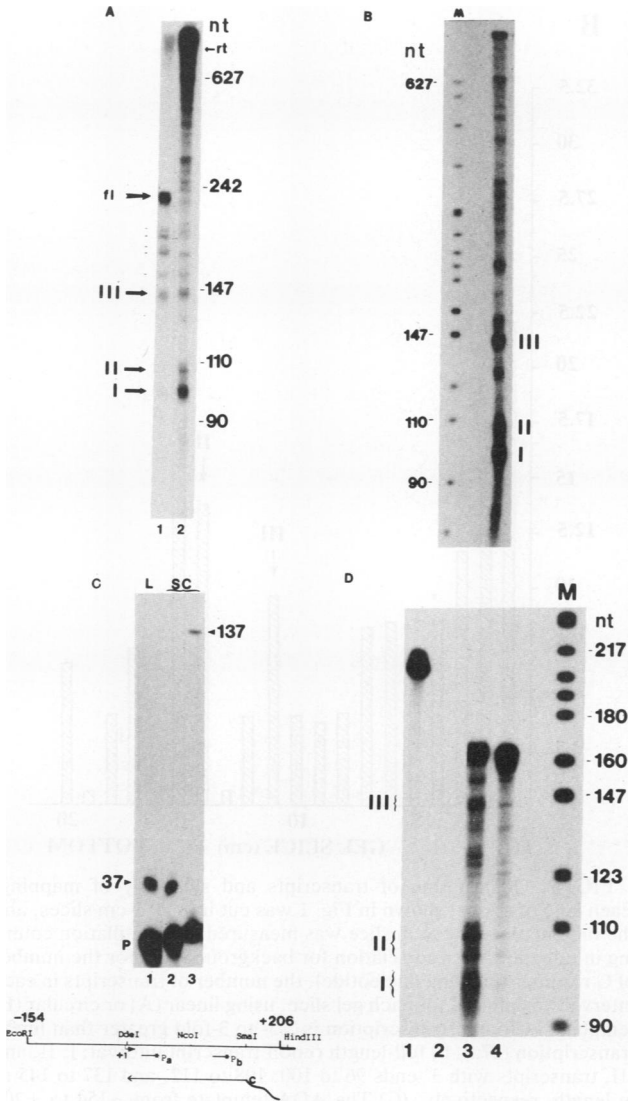

FIG. 1. Characterization of short transcripts from the ADA promoter. (A) A genomic clone (0.4) comprising ADA sequences from −154 to +206 in plasmid vector pTZ19R was used to direct production of ADA transcripts in vitro on circular or linear templates as described in Materials and Methods. Lanes: 1, labeled RNA products of a reaction in which 300 ng of *Hin*dIII-linearized plasmid DNA was incubated with extract and NTPs, including [α-^{32}P]GTP (20 μCi), at 30°C for 15 min; 2, labeled reaction products from transcription of the circular template. The products were separated by polyacrylamide gel electrophoresis. fl, full-length runoff transcript from the linear template; rt, read-through transcripts generated on the circular templates. Sites I, II, and III indicate the major reproducible products in the reaction. Small dashes denote other consistently observed transcripts. (B) Different extract preparations vary with respect to the ratio of short versus read-through transcripts. Circular 0.4 DNA was transcribed in an extract preparation different from that used for panel A. Otherwise, identical conditions were used. M, marker DNA. The positions of site I, II, and III RNA molecules are indicated. (C) 5′ end characterization by primer extension analysis was carried out on RNA from reactions using linear templates (L) or circular 0.4 templates (SC) and a primer (P_a) complementary to ADA RNA from +13 to +37 (lanes 1 and 2) or a primer further downstream (P_b) on RNA from circular templates. Extended products of 37 and 137 bp are indicated. Probes used in mapping experiments are depicted at the bottom. (D) RNase mapping. Probe C was generated by T7 polymerase transcription and measures 205 bp. It is complementary to the ADA mRNA from +3 to +168 (shown) but is longer at the 5′ end because of polylinker sequences in the vector (curved line) downstream from the T7 promoter. Probe was annealed with RNA from transcription reactions using circular or linear templates and digested with RNase A and RNase T_1. Protected products were separated in 6% polyacrylamide–7 M urea. Major short transcripts are shown at the right and correspond to those observed at sites I, II, and III seen in panels A and B. Lanes: 1, probe; 2, control extract RNA and carrier RNA; 3, circular template; 4, linear template; M, DNA markers.

sites seen in murine tissue ADA mRNA and ADA RNA from injected oocytes (6, 34). Use of a primer further downstream (P_b, +113 to +137) showed that a single initiation site was used and that short RNA molecules were not the result of initiation at different positions. To demonstrate that the short RNA molecules observed in runoff reactions were transcribed from ADA sequences and to verify their presumed 3′ ends, we performed RNase mapping (Fig. 1D). Short ADA transcripts with 3′ termini at sites I, II, and III as well as the full-length protected product at 165 nt were the major transcripts observed. These data agree with the in vitro labeling data. Similar results were observed with S1 mapping experiments (data not shown). No RNA initiating upstream in plasmid sequences or in the ADA 5′ flanking region was detected in protection experiments using longer probes. Therefore, transcripts of different sizes are the result of 3′ end heterogeneity.

Examination of the apparent distribution of RNA transcripts suggests that a greater proportion of the total transcripts from circular templates accumulated at sites I and II compared with results for linear templates. Most extract preparations generated short and long transcripts on circular templates somewhere between the two examples shown in Fig. 1A and B. To quantitate the proportion of transcripts of various lengths, we divided each lane of Fig. 1A into 1-cm slices and determined the radioactivity present in each slice. After correction for background and normalization for the number of G residues in the transcripts for each slice interval, we obtained a value that is proportional to the number of transcripts in each gel slice. These results are shown in Fig. 2A and B. It can be seen that a greater proportion of the total transcripts from circular templates than from linear templates accumulated at sites I and II (30% versus 13%). This finding implies that template topology may affect the distribution of these short RNA transcripts. We have performed other experiments to explore the parameters which affect the distribution of short and full-length transcripts. Dilution of the extract used in Fig. 1A increases greatly the proportion of short RNA products at site I or II, producing a pattern similar to that shown in Fig. 1B. This finding implies that a balance of factors may be critical to the synthesis of a complete mRNA and for premature 3′ end formation by the transcription apparatus (data not shown). A summary of the results is shown in Fig. 2C. The 3′ ends of the site I transcripts map near the end of exon 1 approximately 23 to 27 bp in front of the splice site (+122), and site III is positioned approximately 18 to 23 bp into intron 1. Other minor downstream 3′ ends are included, and the width of the bands reflects the relative abundance of observed transcripts with a given 3′ end. In conclusion, nuclear extracts from HeLa cells generate short ADA transcripts which are similar to the short ADA transcripts generated in oocytes, and circular templates favor the formation of transcripts with 3′ ends at sites I and II.

Site I transcripts are the processed products of site II

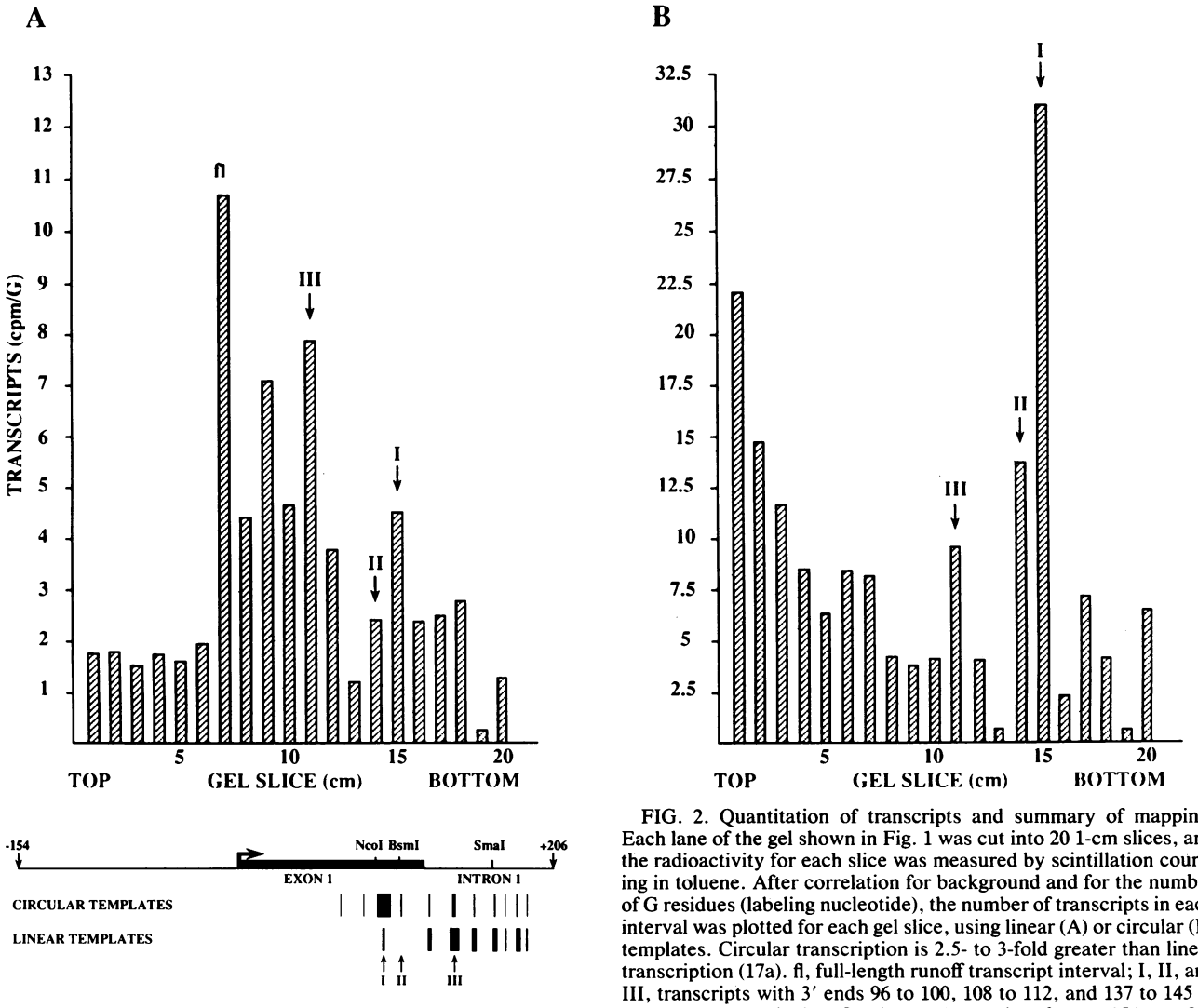

FIG. 2. Quantitation of transcripts and summary of mapping. Each lane of the gel shown in Fig. 1 was cut into 20 1-cm slices, and the radioactivity for each slice was measured by scintillation counting in toluene. After correlation for background and for the number of G residues (labeling nucleotide), the number of transcripts in each interval was plotted for each gel slice, using linear (A) or circular (B) templates. Circular transcription is 2.5- to 3-fold greater than linear transcription (17a). fl, full-length runoff transcript interval; I, II, and III, transcripts with 3′ ends 96 to 100, 108 to 112, and 137 to 145 nt in length, respectively. (C) The ADA template from −154 to +206 including the promoter, exon 1, and part of intron 1 is shown. Bars underneath depict the approximate position of the 3′ ends of short ADA transcripts which accumulate in reactions with circular or linear templates derived from the data in Fig. 1. The width of the band indicates the relative abundance of transcripts at these positions for each template. The transcripts with 3′ ends at sites I, II, and III are marked by arrows.

precursors. Using the transcription assay, we wanted to determine whether the site I transcripts are primary transcription products or the result of degradation or processing of a longer precursor. To address this question, we performed a pulse-chase experiment (Fig. 3). A transcription reaction using circular templates and NTPs, including 1 μM [α-^{32}P]GTP (3,000 Ci/mmol), was initiated. After 1 min, nonradioactive GTP was added to 1 mM and the reaction was allowed to proceed. Two minutes later, α-amanitin was added to 1 μg/ml and an aliquot was removed. RNA was immediately purified from the initial aliquot and compared with RNA in aliquots removed from the reaction mixture at various times after α-amanitin addition. Site II transcripts (approximately 108 to 112 nt in length) were the predominant transcript early in the reaction. As the reaction was continued, there was a gradual decrease in the amount of these transcripts and a concomitant increase in the abundance of the site I RNA molecules while the total amount of both remained constant. This result strongly suggested that the site I molecules resulted from degradation or specific cleavage of the 3′ end of site II transcripts. Site III transcripts were not degraded. Pulse-chase reactions performed without α-amanitin gave identical results (data not shown).

To test whether longer ADA transcripts could be substrates for specific cleavage, we uniformly radiolabeled RNA in transcription reactions using linear or circular templates with [α-^{32}P]GTP. We subsequently purified the labeled ADA RNA molecules of the size of the full-length runoff product and site III RNA by elution after gel electrophoresis and reintroduced these transcripts into extracts under standard reaction conditions (Fig. 4A). The full-length transcripts were not cleaved to site III, II, or I over 15 min. The site III RNA was degraded to a number of slightly shorter intermediates without specific cleavage to site I or II. Thus, longer RNA molecules are not substrates for posttranscriptional processing to site II or I when added to the extract after preparation. Therefore, site II precursors likely are formed by termination or cotranscriptional processing.

We have attempted to extend both populations of tran-

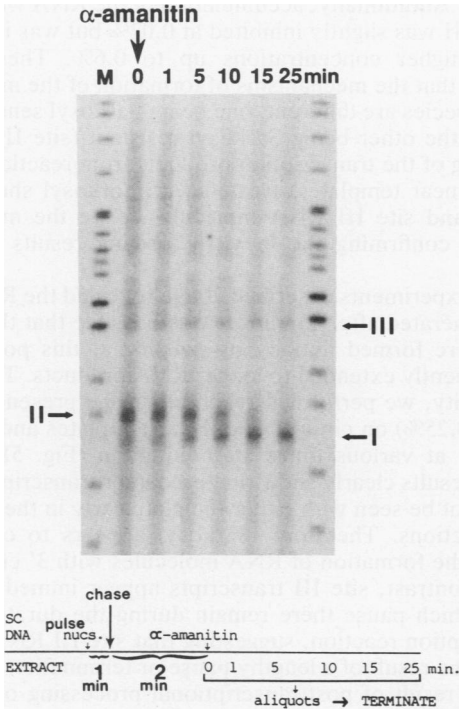

FIG. 3. Pulse-chase analysis of short transcript generation. Transcription was initiated with 0.4 circular (SC) templates in the presence of [α-^{32}P]GTP (1 μM; 3,000 Ci/mmol); 1 min later, a 1,000-fold excess of unlabeled GTP was added. The reaction was continued for 2 min at 30°C, whereupon α-amanitin was added to 1 μg/ml. An aliquot was taken at that time (lane 0) and at subsequent times indicated above the lanes. The reactions were terminated immediately after removal of aliquots, and RNA was purified and separated by polyacrylamide gel electrophoresis. Trichloroacetic acid-precipitable counts per minute were similar for all aliquots (data not shown).

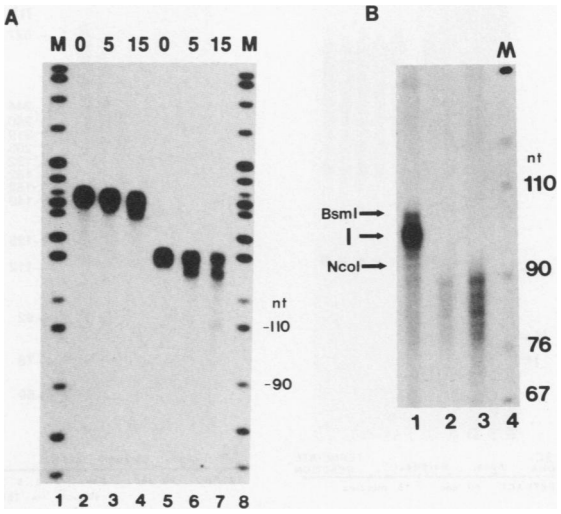

FIG. 4. Evidence that RNA molecules with 3′ ends at site III or greater in length are not precursors for site I, II or III RNA transcripts. (A) In vitro transcription reactions were carried out, and RNA molecules corresponding to the full-length 0.3 *Hin*dIII-generated runoff RNA (203 bp) (lanes 2 to 4) and the site III RNA (lanes 5 to 7) were cut and eluted from the gel. The purified RNA was incubated with extract under transcription reaction conditions without DNA in the presence of NTPs. Aliquots were removed at 0, 5, and 15 min, and the RNA was purified and separated by gel electrophoresis. (B) The transcripts with 3′ ends at site I are stable degraded intermediates of runoff transcripts with 3′ ends at +103. 0.4 DNA template was cleaved with *Bsm*I at +103 (see Fig. 2) or *Nco*I at +92. Lanes: 1, labeled reaction products generated from the *Bsm*I-cleaved template; 2 and 3, labeled RNA products from the *Nco*I-cleaved template with 150 and 300 ng, respectively, of DNA template per reaction; 4, marker DNA. Site I is indicated, as are the run-off transcripts at +103 and +92.

scripts into longer RNA molecules by adding fresh NTPs to the reaction 15 min after initiation. Repletion of NTPs does not result in extension of these transcripts (data not shown). We also have used the procedure of Culotta et al. (8) to sediment transcription complexes in reaction mixtures away from free RNA and found that site I transcripts are greatly enriched in the soluble (nonsedimented) fraction. The site II transcripts also are present in the soluble fraction but are much less abundant because of processing (data not shown). These results imply that both groups of transcripts are unattached to sedimentable complexes or are released from the template and would not be substrates for extension into longer RNA molecules by polymerase II. We have not found any other conditions in which preformed site I or site II transcripts are elongated (see Fig. 5D). We cannot rule out a cotranscriptional RNA cleavage event as an explanation for the generation of the precursors. However, the precursor RNA transcripts also may be the result of termination and are subject to degradation. Since α-amanitin arrests polymerase progression, the results suggest that degradation or processing of the precursor is uncoupled from transcriptional elongation. Different extract preparations vary with respect to the rate of cleavage of the site II RNA molecules (data not shown).

We performed additional experiments to investigate how the RNAs with 3′ ends at site I were formed in the in vitro reactions. Linear templates with 3′ ends at +103, where a *Bsm*I restriction site is located (Fig. 2), generate runoff RNA molecules which are processed to the 3′ ends at site I, whereas runoff RNA molecules with 3′ ends at +92 (*Nco*I site) are degraded rapidly (Fig. 4B). These results suggest that sequences in the RNA downstream of +103 or at the 3′ end of the precursors at site II are not necessary in order for the processing/cleavage reaction to occur. Generation of the 3′ ends at site I may involve exonucleolytic degradation of the 3′ end of the site II RNA to a stable structure.

Sarkosyl inhibits generation of the site II RNA precursors as well as processing to site I. We reasoned that protein-protein and protein-nucleic acid interactions were essential for premature 3′ end formation in exon 1. Other groups (14, 48, 51) have demonstrated marked effects on the elongation or termination properties of RNA polymerase II with various quantities of Sarkosyl in vitro. We wanted to investigate whether Sarkosyl would affect the generation of ADA transcripts with prematurely formed 3′ ends. A transcription reaction with a circular (Fig. 5A) template was initiated and allowed to proceed for 1 min. Sarkosyl was then added at various concentrations, and the reaction was allowed to proceed for 15 min. Production of all short transcripts between 96 and 112 nt was abolished in the presence of 0.05% Sarkosyl. Reactions with linear templates were similarly affected (data not shown). Further evaluation showed that inhibition began to be apparent in crude extracts at 0.02% Sarkosyl and was complete at 0.025% (data not

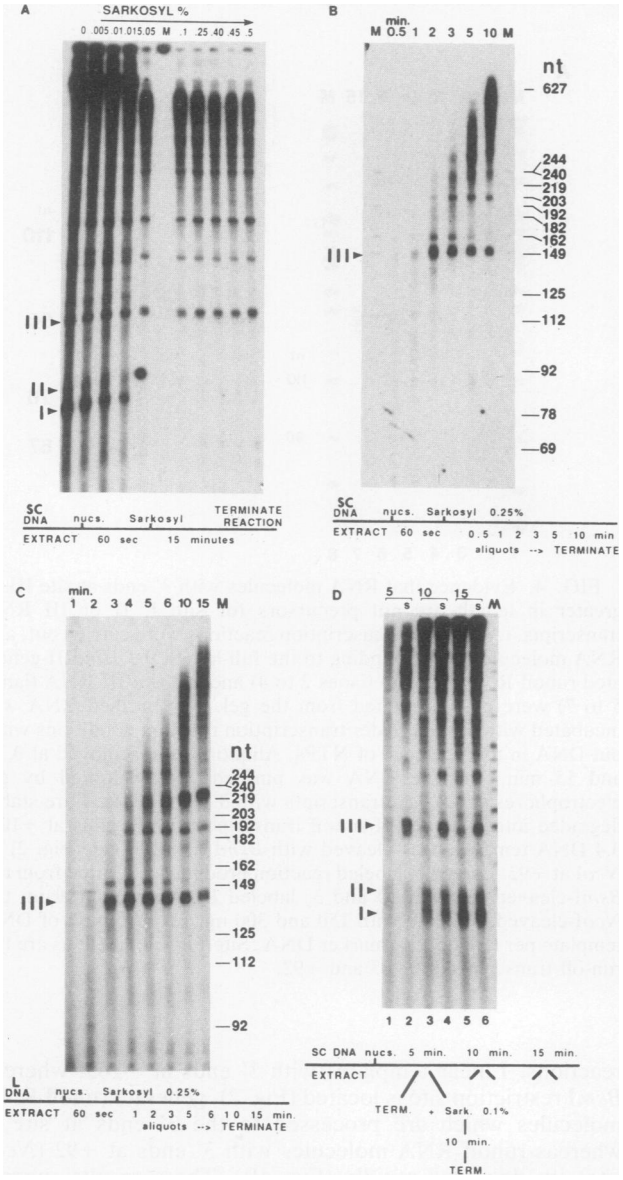

FIG. 5. Inhibition by Sarkosyl of premature 3′ end formation at site II but not at site III. (A) A circular template (0.4) was used to direct transcription. Sixty seconds after NTPs were added to initiate the reaction, Sarkosyl was added to various concentrations from 0 to 0.5%, and the reactions were allowed to proceed for 15 min at 30°C. RNA was purified and separated by gel electrophoresis. The RNA molecules with 3′ ends at sites I, II, and III are shown. Lanes M in all panels, markers. The reaction schemes are depicted below the autoradiograms. SC, supercoiled circular; L, linear. An artifact of exposure adjacent to lane 5 is not site II RNA. Sarkosyl completely inhibits site II formation at 0.05%. (B) Circular DNA was used to direct transcription for 1 min. Sarkosyl was added to 0.25%, and aliquots were taken at later times from 0.5 to 10 min, as shown above the lanes. The prominent site III RNA molecules are indicated on the left. (C) Linear DNA (0.4 *Hin*dIII) was mixed with extract and NTPs. One minute after initiation, Sarkosyl was added to 0.25%; aliquots were taken from this reaction at 1, 2, 3, 4, 5, 6, 10, and 15 min. RNA was purified and separated by electrophoresis. The arrowhead indicates the site III transcripts. (D) A transcription reaction was initiated with a circular template, and after 5 min the reaction in an aliquot was terminated. An aliquot of identical volume was made up to 0.1% Sarkosyl and allowed to continue transcription for 10 min at 30°C, whereupon the reaction was terminated. Aliquots were treated the same at 10- and 15-min intervals. The RNA products were separated by gel electrophoresis. Lanes S, Sarkosyl-treated aliquots.

shown). Additionally, accumulation of the RNA with 3′ ends at site III was slightly inhibited at 0.02% but was insensitive to all higher concentrations up to 0.6%. These results suggest that the mechanisms of formation of the major short RNA species are different, one being Sarkosyl sensitive (site II) and the other being Sarkosyl resistant (site III). RNase mapping of the transcription products from reactions (circular or linear templates) treated with Sarkosyl showed full-length and site III RNA molecules to be the major transcripts, confirming the in vitro labeling results (data not shown).

The experiments described above revealed the RNA products generated after 15 min. It was possible that the precursors were formed initially by pausing at this position but subsequently extended to longer RNA products. To test this possibility, we performed reactions in the presence of Sarkosyl (0.25%) on circular and linear templates and removed aliquots at various times after initiation (Fig. 5B and C). These results clearly show that precursor transcripts (site II) could not be seen with either template early in the course of the reactions. Therefore, Sarkosyl appears to completely inhibit the formation of RNA molecules with 3′ ends at site II. In contrast, site III transcripts appear immediately and most which pause there remain during the duration of the transcription reaction, suggesting that site III RNA production is the result of a lengthy pause or termination event and not the result of posttranscriptional processing of a longer transcript.

Since Sarkosyl was noted to abolish production of the precursor RNA transcripts, we tested whether it inhibited processing to site I as well. A transcription reaction was initiated, and two aliquots were removed at 5, 10, and 15 min. Transcription in one aliquot was stopped, and the other was made up to 0.25% Sarkosyl and allowed to incubate for another 10 min at 30°C before the reaction was terminated. RNA was purified from all six aliquots and separated by gel electrophoresis (Fig. 5D). We compared the abundance of RNA with 3′ ends at site I and site II before and after incubation with Sarkosyl. No change in the abundance of these transcripts was observed during the course of each incubation. Thus, Sarkosyl inhibits both the production of the precursors and the subsequent cleavage or processing of this RNA to products. Accumulation of RNA with 3′ ends at site III increased with incubation in Sarkosyl and results from elongation of shorter transcripts beyond site II during the 10-min incubation.

A heat-labile factor is necessary for premature 3′ end formation. In the course of work performed to study ADA promoter function, we found that transcription was abolished by heating the crude extract at 47°C for 15 min, as had been reported for the adenovirus MLP (32). Transcription was restored on both the adenovirus MLP and the ADA promoter by addition of a fraction enriched for TFIID (Fig. 6A). When transcription was restored in heat-treated extracts with this fraction, transcription arrest at site II was not observed and the accumulation of site III RNA also was reduced. Primer extension analysis of RNA from transcription reactions with heat-treated extracts revealed no ADA products (Fig. 6B), while analysis of RNA from reactions using fraction TFIID-reconstituted extracts showed that ADA transcripts had the identical transcription initiation site seen in RNA produced from crude extracts. Thus, loss of the ability to generate precursor transcripts with 3′ ends at site II is not the result of a change in transcription initiation site. To test whether an inhibitor of site II transcript generation was present in fraction TFIID, we performed reactions using

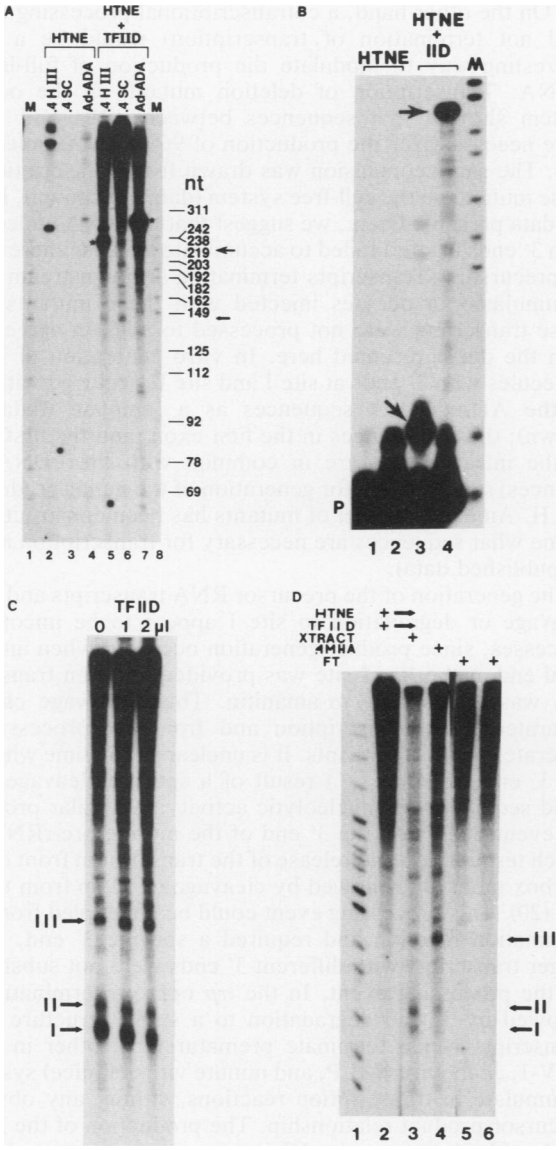

FIG. 6. Evidence that premature 3′ end formation at site II and downstream sites is dependent on a factor(s) which is inactivated at 47°C for 15 min. (A) Linear (0.4 *Hin*dIII) or circular (0.4 SC) ADA templates or a linearized plasmid (Ad-ADA) carrying the adenovirus MLP driving transcription of ADA sequences from −4 to +160 were incubated under transcription reaction conditions with HeLa nuclear extracts that were pretreated at 47°C for 15 min (HTNE) in lanes 2 to 4, respectively. The reactions were repeated with the addition of 1 μl of a fraction of HeLa nuclear extracts enriched for TFIID (lanes 5 to 7). The large arrow points to the position of +144 in the ADA sequences for Ad-ADA in lane 7. The small arrow points to the accumulation of transcripts with 3′ ends at site III in reactions with linear and circular ADA templates in lanes 5 and 6. The three bands in lane 2 are reproducible but are not products of ADA transcription. Asterisks denote the runoff transcripts from the ADA template in lane 5 and the Ad-ADA template in lane 7. Lanes M in panels A and B contain markers. (B) Primer extension analysis of RNA from reactions using circular 0.4 ADA plasmid DNA and heat-treated nuclear extract (HTNE) alone or with fraction TFIID. Primers used: lanes 1 and 3, P_a; lanes 2 and 4, P_b. P denotes the position of both primers (25 bp). P_b shows radioactivity above 25 bp which is not an extended product. The arrows point to the extended products. Lane 3 shows an extended product of approximately 37 bp; lane 4 shows an extended product of 137 bp. (C) Premature 3′ end formation at site II is not inhibited by a component(s) in fraction TFIID. Crude extract (not heat treated) was used in reactions with circular 0.4 DNA as in Fig. 1A with the addition of 0, 1, or 2 μl of fraction TFIID. Transcripts were radiolabeled with [α-^{32}P]GTP. (D) The heat-labile activity which promotes premature 3′ end formation binds to and can be eluted from a heparin agarose column. An extract was partially purified on heparin agarose as described in Materials and Methods. Transcription reactions with circular 0.4 DNA were initiated in heat-treated nuclear extract (HTNE) supplemented with 1 μl of fraction TFIID as in panel A. Crude extract (XTRACT, 1 μl), a protein fraction eluted from the heparin agarose column with a 0.4 KCl step gradient (.4MHA, 1 μl), or protein-containing flowthrough fractions (FT, 2 μl each) were added. Transcripts with 3′ ends at sites I, II, and III are indicated. Lane 2 shows RNA products from a reaction containing heat-treated nuclear extract, TFIID, and DNA. Lane 1, marker DNA.

crude extracts in the presence of one or two times (1 or 2 μl, respectively) the amount of fraction TFIID used to reconstitute transcription (Fig. 6C). These results show that lack of premature 3′ end formation at site II is not the result of an inhibitor present in the fraction which restores transcription. Therefore, a heat-labile activity is required for the transcription machinery to generate transcripts with 3′ ends formed prematurely at site II.

We sought to determine whether this reaction scheme could be used as an assay to purify the factor. A small amount of crude extract and an extract fraction eluted from a heparin agarose column were both capable of restoring premature 3′ end formation (Fig. 6D, lanes 3 and 4). Flowthrough fractions from this column had no activity (lanes 5 and 6). Thus, a component(s) of the extract that promotes premature 3′ end formation is heat labile and binds to and can be eluted from a heparin agarose column. Further purification of this factor is under way.

DISCUSSION

We have developed a cell-free assay to explore transcription arrest in exon 1 of the murine ADA gene which was originally identified by nuclear run-on and oocyte injection experiments. We have identified short ADA transcripts which accumulate in the in vitro assay under standard reaction conditions which are similar to the short transcripts that accumulate in steady-state RNA from oocytes injected with murine ADA templates (34). Thus, the transcriptional mechanisms within oocytes and nuclear extracts from mammalian cells both recognize signals present in the template or its RNA product which direct the production of incomplete, nonpolyadenylated ADA transcripts with 3′ ends at discrete sites. Using the assay, we have provided evidence to support the conclusion that the RNA transcripts with 3′ ends at site I are the stable processed products of 108- to 112-nt RNA precursors. Precursor formation may be the consequence of transcription arrest or cotranscriptional processing, since specific posttranscriptional cleavage of longer ADA transcripts was not observed. In addition, formation of the discrete ends at 108 to 112 nt (site II) is favored over downstream sites in reactions initiated with circular templates and is disrupted by low concentrations of Sarkosyl. Further evidence that protein-protein or protein-nucleic acid interactions are involved was provided by the observation that mild heat treatment of the extract inactivates a component(s) necessary for generating premature 3′ ends in exon 1. This activity was restored by adding back an extract fraction. When disrupted by Sarkosyl, complexes generate tran-

scripts which terminate further downstream. In heat-treated extracts, premature 3′ end formation at all sites is inhibited. Thus, the transcription complex appears to arrest transcription particularly at site II in HeLa nuclear extracts; however, production of full-length mRNA may involve bypassing multiple sites in this region.

We have shown that template topology affects not only the transcription initiation rate of the murine ADA gene but also the proportion of transcripts which terminate at site II versus downstream sites. In oocytes injected with covalently closed supercoiled ADA templates, transcripts with 3′ ends at site I were the predominant short RNA molecules. Other minor sites existed further downstream; however, the data suggest that a covalently closed template favors generation of the short product in oocytes and in nuclear extracts. How does template topology affect premature 3′ end formation? While it has been observed before that template supercoiling increases transcription initiation of some genes in vitro compared with results for linear templates (30), the precise reasons for this difference are incompletely understood. The tension available in a supercoiled molecule may be used by the transcription machinery to help convert the closed template to an open complex. In HeLa nuclear extracts, supercoiled DNA is rapidly converted to multiple covalently closed topoisomers (data not shown), most of which are completely relaxed. The effect observed here may be related to the use of covalently closed templates and not supercoiling. Topoisomerase activity in this process is possible but is not readily obvious simply by examination of the sequence in the region near site II (43, 47). In experiments not shown here, no cleavage sites were detected in a labeled linear double-stranded DNA fragment encompassing this region when the fragment was incubated in HeLa nuclear extracts under transcription reaction conditions. This result does not rule out involvement of topoisomerase activity in the occurrence of premature 3′ end formation. Generation of a prematurely terminated transcript involves multiple steps from pausing of the polymerase to the release of the transcript from the template or transcription complex. In prokaryotes, Rho factor is required for release of the RNA from the template at Rho-dependent terminators but by itself cannot cause termination by RNA polymerase at specific sites. The mechanism of transcript release is unknown. From these considerations, if premature 3′ end formation in ADA can be attributed to premature termination, further work will be needed to elucidate the step or steps in termination which are facilitated by using a covalently closed template for transcription. A possible effect of supercoiling may be addressed by using a relaxed, covalently closed template or a positively supercoiled template to start the reaction. The advantage of starting reactions with supercoiled or circular templates may be applicable to the study of 3′ end formation or transcription termination in other genes.

The major stable short RNA transcripts at site I are cleavage products of 108- to 112-nt precursors. These precursor transcripts may be the result of premature termination of transcription; however, we cannot rule out the possibility that a cotranscriptional processing event is responsible. Longer RNA molecules are not cleaved specifically to site II when added back to extract; thus, posttranscriptional processing is not likely the mechanism for site II transcript formation. In addition, it is difficult to imagine how cotranscriptional RNA processing would be affected by template topology, unless RNA molecules were hybridized more often to their linear template and not to the circular template and therefore not available to the processing activity. On the other hand, a cotranscriptional processing event (and not termination of transcription) would be a very interesting way to modulate the production of full-length mRNA. Transcription of deletion mutants in the oocyte system showed that sequences between +104 and +168 were necessary for the production of 96-nt RNA molecules (34). The same conclusion was drawn from transcription of these mutants in the cell-free system (data not shown). From the data presented here, we suggest that the RNA molecules with 3′ ends at site I failed to accumulate because there were no precursors. Transcripts terminating at downstream sites accumulated in oocytes injected with these mutants, but those transcripts were not processed to +96, in agreement with the data presented here. In vitro generation of RNA molecules with 3′ ends at site I and site II occurred with use of the ADA cDNA sequences as a template (data not shown); thus, sequences in the first exon (and the first 3 bp of the intron which are in common with the cDNA sequences) are sufficient for generation of the major product at site II. An extensive set of mutants has been constructed to define what sequences are necessary for transcription arrest (unpublished data).

The generation of the precursor RNA transcripts and their cleavage or degradation to site I appear to be uncoupled processes, since product generation occurred when an artificial end at the *Bsm*I site was provided or when transcription was inhibited by α-amanitin. Thus, cleavage can be separated from transcription and from the process that generates site II transcripts. It is unclear at this time whether the 3′ end trimming is a result of a specific cleavage or a rapid sequential exonucleolytic activity. A similar processing event occurs on the 3′ end of the murine pre-rRNA, in which termination and release of the transcript in front of the *Sal* box motif are followed by cleavage of 10 bp from the 3′ end (20). This processing event could be uncoupled from the termination reaction and required a specific 3′ end, since longer transcripts with different 3′ ends were not substrates for the processing event. In the *trp* operon, termination is followed by 3′ end degradation to a stable structure (31). Transcripts which terminate prematurely in other in vitro (HIV-1, adenovirus MLP, and minute virus of mice) systems accumulate in transcription reactions without any obvious precursor-product relationship. The production of the ADA RNA with 3′ ends at site III in our reactions more closely resembles those events. The generation of the site I RNA products appear to result from degradation of only the site II RNA precursors, since longer RNA molecules were not cleaved in the add-back experiments and the pulse-chase reaction. However, it is possible that degradation of the 3′ ends of terminated RNA molecules may be random and other secondary structure may protect the 3′ ends of longer RNA molecules for short in vitro reaction times. In this view, the transcripts with 3′ ends at site I represent relatively stable intermediates. What makes these 3′ ends stable? First, the secondary structure of the RNA may retard further degradation. The 3′ ends of transcripts generated from linear templates cleaved at the *Nco*I site are rapidly degraded, whereas those from templates cleaved at +103 (*Bsm*I site) accumulate at between 96 and 100 nt. Therefore, it is possible that the 96- to 100-nt transcripts are stable degradation products due to secondary structure which is not present in the transcripts ending at the *Nco*I site. Another possibility to account for the increased stability of site I transcripts in oocytes and nuclear extract transcription reactions is that they may be bound by a protein involved in transcription arrest which stabilizes the 3′ end of the RNA.

Precedent for protein bound to RNA and involved in termination or antitermination of transcription is abundant (7, 33, 52). Finally, secondary structures that this region of the ADA mRNA may assume may carry signals involved in antitermination of transcription.

The generation of the 108- to 112-nt precursors occurs in crude extracts in a substantial fraction of elongating complexes under standard reaction conditions. For ADA, Sarkosyl abolishes premature 3′ end formation at this site (as well as processing), which is opposite the expectation for the outcome of this experiment based on data from other in vitro studies regarding premature 3′ end formation by polymerase II. For the adenovirus MLP and HIV long terminal repeat, in vitro premature termination events are stimulated to occur with Sarkosyl (14, 48, 51). Generation of the ADA precursors was sensitive to concentrations of Sarkosyl known to inhibit initiation of transcription on the adenovirus MLP (14). These same concentrations also inhibit initiation on the ADA promoter in crude extracts (data not shown). The observation that initiation and premature 3′ end formation are inhibited by the same concentrations of Sarkosyl may be fortuitous. On the other hand, it seems possible that communication between the initiation complex and the signals downstream at the site of arrest may occur. Sarkosyl does not affect the generation of site III RNA except minimally at concentrations of between 0.025 and 0.03% (data not shown). From the pulse-chase experiment, no change in the length or abundance of site III RNA was observed, and from the time course reactions in the presence of Sarkosyl, the 144-nt product appears to be a major termination site in vitro. Complexes that do not produce site II transcripts appear to elongate further and arrest at site III. This view is supported by the results of experiments in which Sarkosyl was added to reactions 5, 10, and 15 min after initiation and by direct comparison of the RNA products from linear and circular templates in the absence of Sarkosyl. In addition, in oocytes and HeLa nuclear extracts RNA polymerase fails to generate site II transcripts if sequences immediately downstream from this position are deleted (data not shown). Instead, the polymerase reads through this region and may generate transcripts which terminate further downstream. From these results, it can be concluded that the mechanism of formation of site III RNA is different from that of the shorter pair. Wiest and Hawley (51) demonstrated that elongation through a termination site in the first intron of the adenovirus major late gene was severely impaired by Sarkosyl. Moreover, transcription complexes behaved differently at each of two terminator sequences placed in tandem; termination at the second site was greater, implying that transcription complexes which read through the first site were still capable of terminating. Their results and ours suggest that elongation complexes which read through one transcription site are capable of being arrested further downstream; however, the factors which affect these events at separate sites in ADA appear to differ. Alternatively, the latter observation may imply that transcription complexes elongated further down the template are different. One way this could be tested is to increase or decrease the distance between the promoter and sites II or III to determine whether the efficiency of premature 3′ end formation is affected.

We have shown that a heat-labile factor(s) is necessary for the generation of short transcripts and that a fraction enriched for TFIID restores transcription to the ADA promoter. Transcription initiated accurately in the reconstituted reaction and lack of premature 3′ end formation was not due

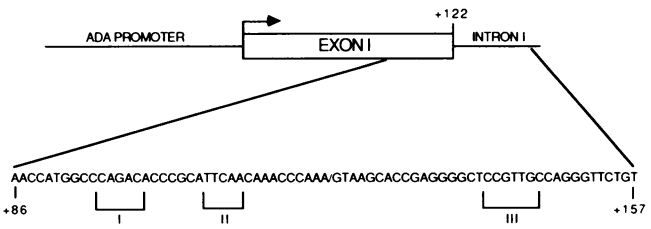

FIG. 7. Sequence of the ADA first exon and first intron surrounding sites I, II, and III. Exon 1 (122 bp) is shown as a block. The nucleotide sequence from +86 to +157 is shown. The slash denotes the exon-intron junction. The region of the first exon from +86 to +122 contains 15 C residues and 4 G residues. No sequence similarities have been observed between this region and those which promote premature 3′ end formation in other eukaryotic genes. Sites I, II, and III are outlined and refer to sequences from 96 to 100, 108 to 112, and 137 to 145 nt, respectively.

to an inhibitor in the enriched fraction. Activity can be restored to heat-treated extracts by the addition of a partially purified extract. We presume that the activity of this factor(s) to promote premature 3′ end formation is sensitive to low concentrations of Sarkosyl. The identity of this factor is unknown, and further purification is under way. We propose that the heat-labile factor(s) promotes transcription arrest within the first exon particularly at site II and at multiple sites within the 5′ end of the first intron by interacting with the DNA, RNA, or RNA polymerase and may be a target of cellular activities aimed at allowing elongation to proceed without arrest. Examination of the sequence in the region surrounding sites I and II (Fig. 7) shows a marked abundance of C residues (fifteen) relative to G residues (four) between +86 and +122, whereas RNA sequences upstream and downstream of this region have equal numbers of G and C residues in the RNA. Stretches of RNA in which C residues greatly outnumber G residues have been identified as Rho utilization sites in bacteria (36). It is tempting to speculate that the same features of the RNA are recognized by a similar protein in eukaryotes and lead to premature 3′ end formation under circumstances involving a regulated pause by the polymerase at site II. Future studies will focus on the sequence elements of the region which play a role in premature transcription arrest.

By studying the requirements for premature 3′ end formation in ADA in vitro, we can begin to propose experiments to address the questions relating to read-through of these sites and the subsequent formation of a complete mRNA. Elongation of the transcript by the polymerase past these sites may require a factor analogous to N protein of phage lambda (33) or the Tat protein of HIV-1 (7). Elucidation of the exact sequence elements and factors involved in the regulation of the ADA elongation block may be applicable to the study of transcriptional elongation control of other genes transcribed by RNA polymerase II.

ACKNOWLEDGMENTS

We thank Cynthia Lindsay for preparation of the manuscript. We thank Marilyn Szentirmay for fraction TFIID. We appreciate the helpful discussion and critical comments by Randy Mifflin, Muffin Alford, Shera Kash, John Winston, Zhi Chen, Jeffrey Chinsky, V. Ramamurthy, and Jim Lupski.

This work was supported by NIH grants GM30204 and AI25255, grant Q-893 from the Robert A. Welch Foundation, and a grant from Eli Lilly & Co. J.W.I. is a postdoctoral fellow supported by a National Research Service Award GM12921.

REFERENCES

1. **Ben-Asher, E., and Y. Aloni.** 1984. Transcription of minute virus of mice, an autonomous parvovirus, may be regulated by attenuation. J. Virol. **52**:266–276.
2. **Bender, T. P., C. B. Thompson, and W. M. Kuehl.** 1987. Differential expression of c-*myb* mRNA in murine B lymphomas by a block to transcription elongation. Science **237**:1473–1476.
3. **Bentley, D. L., and M. Groudine.** 1986. A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. Nature **321**:702–706.
4. **Bentley, D. L., and M. Groudine.** 1988. Sequence requirements for premature termination of transcription in the human c-*myc* gene. Cell **53**:245–256.
5. **Chen, Z., M. L. Harless, D. A. Wright, and R. E. Kellems.** 1990. Identification and characterization of transcriptional arrest sites in exon 1 of the human adenosine deaminase gene. Mol. Cell. Biol. **10**:4555–4564.
6. **Chinsky, J. M., M.-C. Maa, V. Ramamurthy, and R. E. Kellems.** 1989. Adenosine deaminase gene expression: tissue-dependent regulation of transcriptional elongation. J. Biol. Chem. **264**: 14561–14565.
7. **Cullen, B. R.** 1990. The HIV-1 tat protein: an RNA sequence-specific processivity factor. Cell **63**:655–657.
8. **Culotta, V. C., R. J. Wides, and B. Sollner-Webb.** 1985. Eucaryotic transcription complexes are specifically associated in large sedimentable structures: rapid isolation of polymerase I, II, and III transcription factors. Mol. Cell. Biol. **5**:1582–1590.
9. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
10. **Fort, P., J. Rech, A. Vie, M. Piechaczyk, A. Bonnieu, P. Jeanteur, and J.-M. Blanchard.** 1987. Regulation of c-*fos* gene expression in hamster fibroblasts: initiation and elongation of transcription and mRNA degradation. Nucleic Acids Res. **15**: 5657–5667.
11. **Grayhack, E. J., and J. W. Roberts.** 1982. The phage λ Q gene product: activity of a transcription antiterminator *in vitro*. Cell **30**:637–648.
12. **Grayhack, E. J., X. Yang, L. Law, and J. W. Roberts.** 1985. Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. Cell **42**:259–269.
13. **Haley, J. D., and M. D. Waterfield.** 1991. Contributory effects of de novo transcription and premature transcript termination in the regulation of human epidermal growth factor receptor proto-oncogene RNA synthesis. J. Biol. Chem. **266**:1746–1753.
14. **Hawley, D. K., and R. G. Roeder.** 1985. Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. J. Biol. Chem. **260**:8163–8172.
15. **Hay, N., and Y. Aloni.** 1984. Attenuation in SV40 as a mechanism of transcription-termination by RNA polymerase B. Nucleic Acids Res. **30**:1401–1414.
16. **Hay, N., H. Skolnik-David, and Y. Aloni.** 1982. Attenuation in the control of SV40 gene expression. Cell **29**:183–193.
17. **Innis, J. W., D. J. Moore, S. F. Kash, V. Ramamurthy, M. Sawadogo, and R. E. Kellems.** 1991. The murine adenosine deaminase promoter requires an atypical TATA box which binds TFIID and transcriptional activity is stimulated by multiple upstream Sp1 binding sites. J. Biol. Chem., in press.
18. **Kao, S.-Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin.** 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by *tat* gene product. Nature (London) **330**:489–493.
19. **Krystal, G., M. Birrer, J. Way, M. Nau, E. Sausville, C. Thompson, J. Minna, and J. Battey.** 1988. Multiple mechanisms for transcriptional regulation of the *myc* gene family in small-cell lung cancer. Mol. Cell. Biol. **8**:3373–3381.
20. **Kuhn, A., and I. Grummt.** 1989. 3′-end formation of mouse pre-rRNA involves both transcription termination and a specific processing reaction. Genes Dev. **3**:224–231.
21. **Lamb, N. J. C., A. Fernandez, N. Tourkine, P. Jeanteur, and J.-M. Blanchard.** 1990. Demonstration in living cells of an intragenic negative regulatory element within the rodent c-fos gene. Cell **61**:485–496.
22. **Laspia, M. F., A. P. Rice, and M. B. Mathews.** 1989. HIV-1 tat protein increases transcriptional initiation and stabilizes elongation. Cell **59**:283–292.
23. **Laspia, M. F., A. P. Rice, and M. B. Mathews.** 1990. Synergy between HIV-1 tat and adenovirus E1A is principally due to stabilization of transcriptional elongation. Genes Dev. **4**:2397–2408.
24. **Lazinski, D., E. Grzadzielska, and A. Das.** 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59**:207–218.
25. **Maa, M.-C., J. M. Chinsky, V. Ramamurthy, B. D. Martin, and R. E. Kellems.** 1990. Identification of transcription stop sites at the 5′ and 3′ ends of the murine adenosine deaminase gene. J. Biol. Chem. **265**:12513–12519.
26. **Maderious, A., and S. Chen-Kiang.** 1984. Pausing and premature termination of human RNA polymerase II during transcription of adenovirus in vivo and in vitro. Proc. Natl. Acad. Sci. USA **81**:5931–5935.
27. **Marciniak, R. A., B. J. Calnan, A. D. Frankel, and P. A. Sharp.** 1990. HIV-1 tat protein *trans*-activates transcription *in vitro*. Cell **63**:791–802.
28. **Marciniak, R. A., M. A. Garcia-Blanco, and P. A. Sharp.** 1990. Identification and characterization of a HeLa nuclear protein that specifically binds to the *trans*-activation-response (TAR) element of human immunodeficiency virus. Proc. Natl. Acad. Sci. USA **87**:3624–3628.
29. **McGeady, M. L., T. G. Wood, J. V. Maizel, and G. F. Vande Woude.** 1986. Sequences upstream from the mouse c-*mos* oncogene may function as a transcription termination signal. DNA **5**:289–298.
30. **Mizutani, M., T. Ohta, H. Watanabe, H. Handa, and S. Hirose.** 1991. Negative supercoiling of DNA facilitates an interaction between transcription factor IID and the fibroin gene promoter. Proc. Natl. Acad. Sci. USA **88**:718–722.
31. **Mott, J. E., J. L. Galloway, and T. Platt.** 1985. Maturation of *Escherichia coli* tryptophan operon mRNA: evidence for 3′ exonucleolytic processing after rho-dependent termination. EMBO J. **4**:1887–1891.
32. **Nakajima, N., M. Horikoshi, and R. G. Roeder.** 1988. Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. Mol. Cell. Biol. **8**:4028–4040.
33. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.

33a. **Ramamurthy, V., and R. E. Kellems.** Unpublished data.

34. **Ramamurthy, V., M.-C. Maa, M. L. Harless, D. A. Wright, and R. E. Kellems.** 1990. Sequence requirements for transcriptional arrest in exon 1 of the murine adenosine deaminase gene. Mol. Cell. Biol. **10**:1484–1491.
35. **Resnekov, O., and Y. Aloni.** 1989. RNA polymerase II is capable of pausing and prematurely terminating transcription at a precise location *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. USA **86**:12–16.
36. **Richardson, J. P.** 1991. Preventing the synthesis of unused transcripts by Rho factor. Cell **64**:1047–1049.
37. **Rougvie, A. E., and J. T. Lis.** 1988. The RNA polymerase II molecule at the 5′ end of the uninduced hsp 70 gene of D. melanogaster is transcriptionally engaged. Cell **54**:795–804.
38. **Rougvie, A. E., and J. T. Lis.** 1990. Postinitiation transcriptional control in Drosophila melanogaster. Mol. Cell. Biol. **10**:6041–6045.
39. **Sancar, G. B., A. Sancar, J. W. Little, and W. D. Rupp.** 1982. The *uvr*B gene of Escherichia coli has both *lex*A-repressed and *lex*A-independent promoters. Cell **28**:523–530.
40. **Schneider-Schaulies, J., A. Schimpl, and E. Wecker.** 1987. Kinetics of cellular oncogene expression in mouse lymphocytes. II. Regulation of c-fos and c-myc gene expression. Eur. J. Immunol. **17**:713–718.
41. **Sellitti, M. A., P. A. Pavco, and D. A. Steege.** 1987. *lac* repressor blocks *in vivo* transcription of *lac* control region. Proc. Natl.

Acad. Sci. USA **84**:3199–3203.
42. **Sharp, P. A., and R. A. Marciniak.** 1989. HIV TAR: an RNA enhancer? Cell **59**:229–230.
43. **Shen, C. C., and C.-K. J. Shen.** 1990. Specificity and flexibility of the recognition of DNA helical structure by eukaryotic topoisomerase I. J. Mol. Biol. **212**:67–78.
44. **SivaRaman, L., D. Reines, and C. M. Kane.** 1990. Purified elongation factor SII is sufficient to promote read-through by purified RNA polymerase II at specific termination sites in the human histone H3.3 gene. J. Biol. Chem. **265**:14554–14560.
45. **Skarnes, W. C., D. C. Tessier, and N. H. Acheson.** 1988. RNA polymerases stall and/or prematurely terminate nearby both early and late promoters on polyomavirus DNA. J. Mol. Biol. **203**:153–171.
46. **Spencer, C. A., and M. Groudine.** 1990. Transcription elongation and eukaryotic gene regulation. Oncogene **5**:777–785.
47. **Spitzner, J. R., I. K. Chung, and M. T. Muller.** 1990. Eukaryotic topoisomerase II preferentially cleaves alternating purine-pyrimidine repeats. Nucleic Acids Res. **18**:1–11.
48. **Toohey, M. G., and K. A. Jones.** 1989. *In vitro* formation of short RNA polymerase II transcripts that terminate within the HIV-1 and HIV-2 promoter-proximal downstream regions. Genes Dev. **3**:265–282.
49. **Watson, R. J.** 1988. A transcriptional arrest mechanism involved in controlling constitutive levels of mouse c-myb mRNA. Oncogene **2**:267–272.
50. **Weeks, K. M., C. Ampe, S. C. Schultz, T. A. Steitz, and D. M. Crothers.** 1990. Fragments of the HIV-1 tat protein specifically bind TAR RNA. Science **249**:1281–1285.
51. **Wiest, D. K., and D. K. Hawley.** 1990. In vitro analysis of a transcription termination site for RNA polymerase II. Mol. Cell. Biol. **10**:5782–5795.
52. **Yanofsky, C.** 1988. Transcription attenuation. J. Biol. Chem. **263**:609–612.