Coupled Transcription-Polyadenylation in a Cell-free System*

(Randy C. Mifflin† and Rodney E. Kellemes‡)

From the Verna and Marrs McLean Department of Biochemistry, ‡Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030

To investigate the relationship between transcription and polyadenylation, an in vitro system has been developed in which endogenously transcribed pre-mRNAs containing functional polyadenylation sites are rapidly and accurately cleaved in a HeLa nuclear extract. Cleavage of endogenously transcribed substrates differed from that of exogenous substrates in that a proximal 3' terminus was not required, the reaction was more tolerant of increased Mg2+ levels, and endogenous substrates were cleaved more efficiently. A promoter dependence for this reaction was suggested by the observation that substrates transcribed by bacteriophage T7 RNA polymerase in the presence of nuclear extract were not cleaved. In addition, analysis of the bovine growth hormone poly(A) site indicated that it is highly efficient in vitro which agrees with previous in vivo data. The availability of an in vitro system in which transcription and polyadenylation are coupled should facilitate analysis of the relation between 3' end processing and RNA polymerase II transcription termination as well as the promoter requirements for polyadenylation.

An important process in the production of a functional RNA polymerase II transcript is the generation of a 3' terminus. For most eukaryotic non-histone mRNAs, this is accomplished by a two-step reaction involving cleavage 10-50 nucleotides downstream of a consensus hexanucleotide, AAUAAA, followed by consecutive addition of approximately 250 adenosine residues (reviewed in Wickens, 1990). Polyadenylation aids in the stabilization and translation of mRNAs (Bernstein et al., 1989; reviewed in Jackson and Standart, 1990). In vivo and in vitro studies have identified, in addition to the hexanucleotide, additional cis-acting elements required for efficient polyadenylation. The most common of these is a loosely conserved GU- or U-rich element within 50 bases 3' to the cleavage site (reviewed by Manley, 1988). Additional studies have suggested that sequences encompassing the cleavage site (Ryner et al., 1989) and the topology of the pre-mRNA substrate (Gimmi et al., 1989; Ahmed et al., 1991) further modulate the efficiency of this reaction. In vitro studies have allowed the initial characterization of the complexes which carry out these reactions and the subsequent purification of the factors comprising these complexes (Christofori and Keller, 1988; Takagaki et al., 1989, 1990; Gilmartin and Nevins, 1989).

The in vitro studies mentioned above are carried out by incubating previously synthesized RNAs containing poly(A) addition signals in nuclear extracts and monitoring complex formation or processing (Moore and Sharp, 1985; Humphrey et al., 1987; Skolnick-David et al., 1987). Although use of this assay has greatly contributed to our understanding of the cleavage/polyadenylation process, certain questions remain which cannot be addressed using this system. One such question is whether the process of cleavage/polyadenylation is coupled to pol II transcription in the intact cell. The two are not obligatorily coupled since exogenous substrates are cleaved in nuclear extracts, but in this case efficient cleavage is strictly dependent on a proximal 3' terminus, whereas in vivo this is not the case (Sperry and Berg, 1986). Polyadenylation of pol III-promoted transcripts can occur in vivo although this occurs at a reduced efficiency compared with similar pol II-promoted transcripts (Lewis and Manley, 1986).

Although the above mentioned studies demonstrate that transcription and polyadenylation need not be coupled under certain circumstances, evidence has been presented which suggests the two are coupled in vivo or communicate in some fashion. U1 small nuclear RNAs are transcribed by a pol II-like activity. When transcripts that contain a poly(A) addition signal are initiated from these promoters, the resultant RNAs are not polyadenylated (Hernandez and Weiner, 1986; Dahlberg and Schenborn, 1988; Neuman de Vegvar et al., 1986; Lobo and Hernandez, 1989). Similarly, pol I-promoted transcripts containing poly(A) addition signals are not polyadenylated in vivo (Smale and Tjian, 1985). It is unclear whether the two examples mentioned above result from transcription in nuclear compartments devoid of polyadenylation machinery or from communication between transcription and processing complexes. Finally, studies on the mechanism of pol II transcription termination indicate that transcription and polyadenylation communicate at some level since, in a number of cases, polyadenylation precedes and is required for termination to occur (Whitelaw and Proudfoot, 1986; Logan et al., 1987; Lanoix and Acheson, 1988; Connelly and Manley, 1988).

We were interested in developing an in vitro assay for RNA polymerase II transcription termination signals as a first step toward understanding this seemingly complex process. In order to mimic in vivo termination accurately in vitro, we felt it first necessary to establish conditions under which newly transcribed RNAs were accurately cleaved and polyadenylated in a relatively short time frame. Such an assay would not only facilitate studies of transcription termination but would also

*This research was supported by National Institutes of Health Grant GM30204. 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.

†Supported by National Institutes of Health Postdoctoral Fellowship 5 F32 GM 12253-03. Present address: Dept. Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, TX 77550.

‡To whom correspondence should be addressed. Tel.: 713-798-4572; Fax: 713-798-9428.

1The abbreviations used are: pol, polymerase; ADA, adenosine deaminase; BHG, bovine growth hormone; SVL, simian virus 40 late; bp, base pair(s); HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.
allow some of the questions raised above to be addressed in vitro. Using templates in which the mouse adenosine deaminase (ADA) promoter drives transcription of small fragments containing either the bovine growth hormone (BGH) or simian virus 40 late (SVL) polyadenylation signals in the Hela nuclear extract of Dignam et al. (1983), accurate cleavage was detected coincident with the appearance of full-length transcripts. Unlike the exogenous assay, cleavage efficiency was not dependent upon a proximal 3' terminus and was more tolerant of increased Mg2+ concentrations necessary for optimal transcription. Interestingly, similar templates driven by the phage T7 promoter in the presence of T7 RNA polymerase and nuclear extract generated transcripts that were not processed during this short time frame suggesting a promoter dependence for the rapid processing observed.

MATERIALS AND METHODS

DNA Constructs—The BGH poly(A) cassette used was obtained from the plasmid pDS1 BGH (pBGH in Gimmi et al., 1989). For the exogenous cleavage assay shown in Fig. 1, the 277-bp XhoI-BamHI BGH-specific fragment was inserted into the corresponding sites of pGEMI to create pBGHGEM. The BGH poly(A) substrate was generated by restricting pBGHGEM with EcoRI followed by transcription with SP6 polymerase. The SVL poly(A) cassette was obtained from pSVE (Spicer and Berg, 1986) as HindIII-EcoRI fragment (containing the SV40 BamH1-Blal fragment and pSP64 polylinker) which was inserted into the corresponding sites of pGEMI to create pSVLGM. The SVL poly(A) substrate was generated by transcription with T7 polymerase of either DraI- or HindIII-cut pSVLGM. The templates used for in vitro transcription evolved from the GaK expression cassette of pSP1BGH, except that the bacterial GaK gene was permutated to facilitate I in vitro analysis and the SV40 early promoter was replaced by the murine ADA promoter to support efficient initiation in 4 mM Mg2+(Innis et al., 1991).1 Specific [32P]CTP was included at a final specific activity of 2.5 × 106 cpm/μg of carrier DNA.

Nuclear Extract Preparation—Nuclear extracts were prepared according to Krieg and Melton (1987) using gel-purified probes. Hybridizations were for 12 h at 52°C, and resultant hybrids were digested for 1 h at 37°C with 1500 units/ml RNase T1. Protected fragments were fractionated in 6-8% polyacrylamide, 8 M urea gels.

RESULTS

Substrate Cleavage Using the Exogenous Assay—The two polyadenylation signals focused on in this study are from the bovine growth hormone gene (Woychick et al., 1984) and the SV40 late transcription unit. The BGH signal to our knowledge has not been studied in vitro but has been shown to be highly efficient in vivo (Pfarr et al., 1986). The SVL signal is one of the most thoroughly studied in vitro and in vivo and is also one of the most efficient identified (Carswell and Alwine, 1989; Sheets et al., 1987; Ryner et al., 1989). We first wanted to compare the behavior of each signal in the conventional cleavage assay where capped precursor RNAs, prepared using phage polymerases, are incubated for varying amounts of time in nuclear extracts (Moore and Sharp, 1985). These reactions, when performed in the presence of cordycepin triphosphate, permit visualization of the 5' cleavage product because polymerization of adenosine residues to its 3' terminus is inhibited. 3' cleavage products are generally not seen in these assays because of their relative instability (Moore and Sharp, 1985; Sheets et al., 1987).

The behavior of each cassette in an exogenous cleavage assay containing 1.5 mM Mg2+ is shown in Fig. 1, where odd-numbered lanes contain precursor prior to incubation and even-numbered lanes contain the indicated substrates after a 1-h incubation in extract. The BGH substrate, generated from BGHGEM linearized at the EcoRI site (Fig. 1, lanes 1 and 2), was cleaved to generate a 172-nucleotide product which corresponds to the previously mapped in vivo cleavage site (Woychick et al., 1984). The SVL substrates, generated from SVLGEM linearized at either HindIII (Fig. 1, lanes 3 and 4) or DraI (lanes 5 and 6), were each cleaved after incubation to yield a 172-nucleotide product which corresponds to the previously mapped in vivo cleavage site (Woychick et al., 1984). The SVL substrates, generated from SVLGEM linearized at either HindIII (Fig. 1, lanes 3 and 4) or DraI (lanes 5 and 6), were each cleaved after incubation to yield a 172-nucleotide product which corresponds to the previously mapped in vivo cleavage site (Woychick et al., 1984).
Protection analysis of transcription reactions performed using "Materials and Methods" were incubated for 1 h as described in nuclear extract containing cordycepin triphosphate and 1.5 mM Mg²⁺. After incubation RNA was purified and fractionated in a 6% polyacrylamide, 8 M urea gel which was dried and exposed to generate the autoradiograph shown. In each case, the amount of substrate shown represents one-third of the amount included in each incubation. 

**Fig. 1. Exogenous cleavage of BGH and SVL substrates.** 

Capped ³²P-labeled substrates, prepared as described under "Materials and Methods" were incubated in each band. The 15-min ADAGALBGH time point shown in Fig. 2A, lane 6, contains a transcription reaction which employed the ADAGALBGH template linearized just downstream of the BGH poly(A) cassette and demonstrates that the ratio of full length to cleaved product was essentially the same as in reactions which employed supercoiled templates. Thus a proximal 3' terminus was not required for efficient cleavage in this assay since RNA generated from supercoiled templates was processed as efficiently as that generated from linear templates. Fig. 2A, lane 11, shows the RNase protection pattern of total RNA prepared from cells which were transfected with the ADAGALBGH construct and demonstrates that, in vivo, transcripts were also cleaved to generate an RNA which protected the 432-nt fragment (see also Fig. 3B, lane 5). Fig. 2A, lane 6, shows the protection pattern following in vitro transcription of a similar construct which lacks the ADA promoter and demonstrates that the bulk of transcription was dependent on this promoter. Thus accurate cleavage of endogenously transcribed BGH substrates occurred rapidly, and, unlike the exogenous assay, cleavage efficiency was not influenced by increasing the distance downstream of the cleavage site.

We next wanted to determine if there was any promoter dependence for cleavage of endogenously transcribed BGH substrates. Fig. 2A, lanes 7 and 8, demonstrate the protection pattern generated when the construct used in lane 6, which has a T7 promoter just upstream of the GalK sequences, was incubated in the presence of nuclear extract, as in lane 6, but supplemented with 5 units T7 polymerase without (lane 7) or with (lane 8) GpppG included. In neither case was accurate cleavage observed. These results suggest some type of promoter dependence on cleavage, since similar constructs transcribed by RNA pol II were rapidly processed, whereas those transcribed by the phage promoter were not.

We next investigated whether endogenously transcribed ADAGALSVL substrates were also efficiently cleaved. Fig. 2B, lanes 2–5, demonstrates an RNase protection analysis of an ADAGALSVL transcription time course using a supercoiled template. Full-length transcripts, which protect a 294-nucleotide fragment, were seen within 2 min (lane 3), but accurately cleaved transcripts, which protect a 161-nucleotide fragment were not detected until the 5-min time point (lane 4). This suggests that ADAGALSVL transcripts were not processed as efficiently as ADAGALBGH transcripts. Fig. 2B, lane 7, shows the protection pattern of a transcription reaction performed using ADAGALSVL linearized at the HindIII site, just downstream of the SVL cassette (transcripts generated from the linearized template protect a slightly smaller fragment (287 nt), because the HindIII site lies within the region of probe homology). As for ADAGALBGH, similar levels of processing occurred using either a supercoiled or linearized template, again indicating that a proximal 3' terminus is not required in the coupled assay for maximum efficiency. Fig. 2B, lane 6, shows the protection pattern following transcription of a similar construct lacking the ADA promoter, again demonstrating that the bulk of transcription was dependent on a pol II promoter. Thus rapid and accurate cleavage of endogenously transcribed substrates containing the SVL poly(A) site also occurred.

We next wanted to quantitate the amount of products and precursors present in the reactions shown in Fig. 2 in order to determine the processing efficiencies of endogenously transcribed ADAGALBGH and ADAGALSVL substrates. The gels shown in Fig. 2, A and B, were analyzed using a Betagen instrument which quantitates the disintegrations/min present in each band. The 15-min ADAGALBGH time point shown in Fig. 2A, lane 5, contained 19 amol each of cleaved and full-
Coupled Transcription-Polyadenylation

**A**

![Image A](image1.png)

**B**

![Image B](image2.png)

**Fig. 2.** RNase protection analysis of transcription reactions. 

_A_, ADAGALBGH transcription/cleavage. Transcription reactions performed as described under "Materials and Methods," carried out for the indicated amount of time, were processed and hybridized in reactions containing 5 pmol of the GALBGH probe and digested as described. 

- Lane 1, 5 amol of the 624-nucleotide GALBGH probe.
- Lanes 2–5, time course of supercoiled ADAGALBGH transcription/processing. 
  - Lane 2, 0 min;
  - Lane 3, 2 min;
  - Lane 4, 5 min;
  - Lane 5, 15 min;
  - Lane 6, 15 min time point using supercoiled GALBHGEM (no ADA promoter) as template;
  - Lane 7, as in lane 6, but supplemented with 0.5 mM GpppG;
  - Lane 8, as in lane 7, but supplemented with 0.5 mM GpppG;
  - Lane 9, RNase protection reaction containing 20 μg of Chinese hamster fibroblast RNA; 
  - Lane 10, 15-min time point using HindIII-cut ADAGALBGH as template;
  - Lane 11, RNase protection reaction containing 20 μg of total RNA prepared from Chinese hamster fibroblasts transfected with the ADAGALBGH construct.

_B_, ADAGALSVL transcription/cleavage. Transcription reactions performed as described under "Materials and Methods," carried out for the indicated amount of time, were processed and hybridized in reactions containing 10 pmol of the SVL probe and digested as described. 

- Lane 1, 20 amol of the 294-nucleotide SVL probe;
- Lanes 2–5, time course of supercoiled ADAGALSVL transcription/processing. 
  - Lane 2, 0 min;
  - Lane 3, 2 min;
  - Lane 4, 5 min;
  - Lane 5, 15 min;
  - Lane 6, 15-min time point using supercoiled SVLGEM as template;
  - Lane 7, 15-min time point using HindIII-cut ADAGALSVL as template.

The 15-min ADAGALSVL time point shown in Fig. 2B, lane 5, contained 12 amol of cleaved product and 26 amol of full-length product. The final ratio of full-length to cleaved product for the SVL poly(A) site was always much higher than that for the BGH poly(A) site, indicating that the BGH site functioned more efficiently in this context. Thus cleavage of endogenously transcribed substrates occurred readily and efficiently and differed from that of exogenous substrates in that a proximal 3' terminus was not required. In addition, a promoter requirement for this processing was observed for the ADAGALBGH substrate which further suggests that in this system transcription and cleavage are coupled in some way.

**Comparison of Coupled and Exogenous Assays**—We next wanted to compare the efficiency of the coupled transcription/cleavage assay with that of the standard exogenous assay. By preparing unlabeled exogenous substrates and analyzing cleavage using the RNase protection assay, we reasoned that comparison of the two reactions was possible, since the probes were of the same specific activity. Since the BGH substrates contained relatively long 3' tails downstream of the cleavage site, comparisons were made between the SVL Dral substrate, the most active in the exogenous assay, and supercoiled ADAGALSVL.

Ideally, such a comparison should be conducted under similar reaction conditions. The main difference between the two assays is the Mg²⁺ concentration; exogenous assays contained 1.5 mM Mg²⁺, whereas the transcription reactions were performed in 4 mM Mg²⁺. Fig. 3A demonstrates the effect of increased Mg²⁺ on cleavage of an exogenous substrate. When
exogenous SVL Dral substrates were assayed for cleavage in 4 mM Mg$^{+2}$ (Fig. 3A, lanes 4 and 5), the yield of cleaved product was depressed 5-10-fold when compared with 1.5 mM Mg$^{+2}$ (Fig. 3A, lanes 2 and 3). Fig. 3, B and C, demonstrates the effect of lowered Mg$^{+2}$ on transcription and cleavage of endogenous substrates. Transcription of the ADAGALBGH template in 4 mM Mg$^{+2}$ (Fig. 3B, lane 2) produced 20 amol of full-length product and 24 amol of cleaved product, whereas transcription in 1.5 mM Mg$^{+2}$ (Fig. 3B, lane 3) resulted in 5.7 amol of full-length product and 9 amol of cleaved product. Thus transcription in 1.5 mM Mg$^{+2}$ produced less total cleaved product than at 4 mM Mg$^{+2}$, because the level of transcription was depressed, although the ratio of cleaved to full-length product was higher at 1.5 mM Mg$^{+2}$, indicating that cleavage was more efficient at the lower Mg$^{+2}$ concentration. Similar results were observed with the ADAGALSVL template (Fig. 3C).

In an attempt to achieve a similar level and rate of cleavage of an exogenous SVL Dral substrate as was seen during the time course of endogenously transcribed ADAGALSVL shown in Fig. 2B, we titrated the amount of substrate needed to produce an equivalent amount of cleaved product after a 15-min incubation. Fig. 3D illustrates an exogenous cleavage time course in which 300 amol of substrate/time point was incubated for 0 min (lane 1), 2 min (lane 2), 5 min (lane 3), and 15 min (lane 4). The amount of cleaved product present after a 15-min incubation was determined by Betagen analysis to be 8.0 amol, which is similar to the amount of product seen after 15 min in the transcription reaction (Fig. 2B, lane 5). The most obvious difference between the two assays was the final ratio of cleaved to uncleaved product. To achieve a similar rate of cleavage for the SVL Dral substrate at 1.5 mM Mg$^{+2}$ as in the transcription reaction, we had to introduce 300 amol of substrate or roughly 10 times the amount of substrate generated in the transcription reaction. Thus although cleavage of endogenous and exogenous substrates can proceed at a similar rate if a vast substrate excess is introduced, endogenously transcribed RNAs were cleaved more efficiently.

**DISCUSSION**

We have demonstrated the rapid and accurate cleavage of endogenously transcribed polyadenylation substrates using the HeLa S3 nuclear extract of Dignam *et al.* (1983). The cleavage we observed differed from that seen when exogenous substrates were assayed in three major ways: 1) cleavage was not dependent upon a proximal 3' terminus; 2) the reaction was more tolerant of increased Mg$^{+2}$ levels; and 3) cleavage was more efficient when coupled to transcription, as judged by the ratio of cleaved product to precursor. In addition, we have found the BGH poly(A) site to be very active in vitro, which supports previous in vivo studies (Pfarr *et al.*, 1986; Gimmi *et al.*, 1989). Our results differ from those of Moore
and Sharp (1984), who published a similar study. The major differences between our findings and theirs are the lag time between appearance of full-length substrate and processed product which was minutes or less in our study but hours in theirs and the final ratio of full-length to cleaved products. One explanation for this discrepancy is that the previous study employed whole cell extracts, and it was later found by Moore and Sharp (1985) that nuclear extracts are more proficient in exogenous assays. Furthermore, the previous study employed the adenovirus L3 poly(A) site which we have not studied.

The results presented suggest that transcription and polyadenylation communicate in some fashion in this system, since cleavage of endogenously transcribed substrates was more efficient than that of exogenous substrates. This conclusion is further supported by the fact that endogenously transcribed substrates using T7 RNA polymerase were not detectably processed. Since the templates used in this particular experiment were supercoiled, we would not expect them to be processed via the "exogenous pathway," because the 3' terminus would be too long. Whether this finding was the consequence of actual interaction between the transcription and processing machineries or simply due to differential transcription rates of endogenous pol II and T7 polymerase is presently not clear. As mentioned previously, in vivo studies have demonstrated that RNA polymerase I transcripts containing efficient sites cannot be polyadenylated, possibly due to the non-nucleolar location of the polyadenylation apparatus (Smale and Tjian, 1985). Lewis and Manley (1986) have shown that only 50% of RNA polymerase III transcripts containing an efficient site are polyadenylated in vivo. In this study, the majority of transcripts which read through the poly(A) site terminated at a pol III termination site only 110 nt downstream of the cleavage site. It is possible that these substrates were treated as exogenous substrates, since they contained a proximal 3' terminus, which could explain their relatively low level of cleavage. The availability of an in vitro system to address whether pol II- or pol III-initiated transcripts are processed as efficiently as pol II-promoted transcripts should help resolve the question of whether transcription and polyadenylation are coupled.

It is equally plausible that transcription and polyadenylation machineries do not communicate but that in situ transcription permits the substrate to assume a favored topology (Gimmi et al., 1989; Ahmed et al., 1991), or allow proper hnRNP phasing (Economides and Pederson, 1983) which would favor efficient cleavage. That efficient polyadenylation in vitro reflects a competition between nonspecific and specific RNA-binding proteins to cis-acting elements is supported by studies which have shown that the downstream U-rich regions of several poly(A) sites mediate binding of hnRNP C proteins which stabilize the initial interaction between PF2 and the hexanucleotide, perhaps by attracting other necessary processing factors and excluding nonspecific protein-RNA interactions (Wiuisz and Shenk, 1990; Weiss et al., 1991). In support of this is the finding that in some cases, the downstream element is dispensable when using partially purified processing fractions in vitro (Ryner et al., 1989).

The availability of an in vitro system in which transcription and polyadenylation are coupled will not only facilitate studies of the promoter dependence on 3' processing but should also shed light on the process of transcription termination by RNA polymerase II. Numerous in vivo studies have demonstrated that polyadenylation is a prerequisite for this process (Whitelaw and Proudfoot, 1986; Logan et al., 1987; Lanoix and Acheson, 1988; Connelly and Manley, 1988), but in vivo analysis does not permit accurate mapping of termination sites due to the instability of 3' cleavage products. A soluble in vitro system should allow detection of such products and permit the manipulation of reaction conditions to gain insight into the mechanisms of this process.

Acknowledgments—We are grateful to Drs. Mitchell Reff and Susan Berget for providing the Dpl BGH and SVES6 plasmdids, respectively. We also thank Dr. Berget for sharing her expertise and advice concerning in vitro polyadenylation assays. We thank Dr. Jeffrey Innis, who established the in vitro transcription system utilizing a murine ADA promoter, for providing extracts used initially, and for advice. We also acknowledge the help of Mitchell Frederick, who performed the Betagen analyses.

REFERENCES