Transcription Unit Mapping in Adenovirus: Regions of Termination

S. IWAMOTO,¹ F. EGGERDING,² E. FALCK-PEDERSON,¹ AND J. E. DARNELL, JR.^{1*}

The Rockefeller University, New York, New York 10021,¹ and Department of Pathology, UCLA School of Medicine, Los Angeles, California 90024²

Received 30 January 1985/Accepted 8 April 1986

Using a series of single-stranded clones of adenovirus DNA, we determined the extent of RNA polymerase II transit early in infection for two rightward-reading transcription units. RNA synthesis beginning at the major late promoter (16.5 on the genomic map) continued until approximately 65 to 70 map units so that differential choices of mRNAs within that region were not based primarily on transcriptional decisions but rather on posttranscriptional decisions. Transcription from the major late promoter beginning at 16.5 map units, however, did greatly decrease before approximately 75 map units, ensuring that no mRNAs were formed with sequences beyond approximately 75 map units. Early transcription from E3 then began just past 75 map units (at a higher rate than transcription from the major late promoter); E3 transcripts terminated at least 2 kilobases downstream from a second and final poly(A) site in this transcription unit. The effectiveness of termination in E3 was greater than 95 to 99%.

Regulation of gene expression through regulation of transcriptional termination occurs in a number of bacterial and bacteriophage genes (1, 17, 21, 36). The mechanisms of this regulation vary. In some cases, termed attenuation, the termination event is related to effective translation near the termination site (36). In other cases, proteins bind either to DNA at the site of termination or to newly transcribed RNA near the termination site (17, 21). In animal cells, premature termination by RNA polymerase II (i.e., termination within a few hundred to a few thousand nucleotides of the start site) seems to be a prominent event (10, 14, 23, 31). If a transcription complex escapes premature termination, the polymerase usually then continues across the "useable" portion of a gene; i.e., transcription encompasses all poly(A) sites and proceeds from a few hundred to a few thousand nucleotides further downstream (3, 13, 15, 18-20, 27, 28). Since the portion of a primary transcript that is downstream from a poly(A) site is very unstable (6, 20), it has been difficult to measure. However, because of this instability, examination of newly labeled RNA, not of steady state RNA, is required to successfully define 3' termini of RNA chains caused by termination (12). Steady state RNA in the nucleus is a collection of the most stable processed RNA fragments. No evidence for specific termination sites was found in the one case in which nascent-chain analysis within a termination region was carried out (mouse β -globin in Friend cells [6, 12]).

For several reasons, we returned in these experiments to examination of termination in adenovirus transcription units. First, the transcription rates of these units are fairly high, and several active transcription units are available for analysis on one genome (11). In addition, the major late transcription unit is also active early in infection, albeit producing much less total RNA (2, 4, 28, 33). Late in infection, mRNAs from this transcription unit are produced with 3' ends at five different poly(A) sites, termed L1 and L5 (5, 24, 26, 38), but early in infection, mRNAs accumulate mainly with the L1 poly(A) site (2, 4, 28, 33). It has been proposed (2, 28, 33) that regulated termination in this transcription unit

is at least partly responsible for different patterns of mRNA biogenesis early and late in infection.

With the aid of single-stranded M13 bacteriophage clones of the adenovirus genome (9), we investigated the 3' border of transcription from the major late promoter early in infection. Also, with the single-stranded clones and ³²P-nascent-labeled RNA from isolated nuclei (34), we searched for specific RNA polymerase II termination sites. We found that, early in infection, it is true that no (or very little) mRNA can be produced from L4 and L5 because termination of RNA originating at the major late cap site (37) occurred before transcription of these sequences. However, transcription appeared to proceed across L1, L2, and L3 with no more than a two- to threefold decrease, lending support to the conclusion (28) that RNA processing choices [mainly poly(A) selection], but not RNA termination, and differential mRNA stability are the major reasons that L1 mRNAs are predominant over L2 and L3 early in infection. However, there was a demonstrable and substantial decrease in transcription after the L3 poly(A) site at map unit 62 was reached and before the L4 poly(A) site at map unit 79 was reached. Finally, examination early in infection of RNA from the E3 transcription unit (which begins at map unit 75 and reads rightward) showed that here also transcription occurs past the poly(A) site but decreases greatly about 2 kilobases (kb) beyond the poly(A) site. Again, no definite evidence of specific termination sites was obtained in an assay that was capable of detecting such sites in nascent RNA (6).

MATERIALS AND METHODS

Clones, cells, and virus stocks. HeLa cells were maintained in Joklik's modified essential medium containing 5% fetal calf serum. Adenovirus stocks were grown and purified by methods described by Nevins (25). For experiments, cells were infected with 5,000 particles per cell for 5 h (early infection) or with 2,000 particles per cell for 18 h (late infection).

Most of the bacteriophage clones of adenovirus DNA were constructed as described by Eggerding and Pierce (9). The clone spanning L3 (60.7 to 64.2 map units) was prepared by Sarah Hanly; the clone spanning from 76.8 to 78.4 map units was prepared by Der-Hwa Huang.

* Corresponding author.

JM101 cells were grown to a density of 10^8 cells per ml. They were infected with M13 recombinant phage at a multiplicity of infection of 50. After infection at 37°C for 3 to 5 h, the cells were spun down for restriction mapping analysis to ensure the integrity of the clone. To isolate mature phage, the supernatant was precipitated twice in polyethylene glycol 6000, once in 5% polyethylene glycol–0.5 M NaCl, and then a second time in 2% polyethylene glycol–0.5 M NaCl. The phages were then purified by centrifugation through a 25 to 40% cesium chloride gradient in a 10 mM Tris chloride (pH 7.5)–2 mM EDTA buffer at $150,000 \times g$ for 24 h. Phage bands were collected, cesium chloride was removed by dialysis, and the phage DNA was then extracted with phenol-chloroform.

The single-stranded DNA was checked for possible deletions and rearrangements within the cloned region by hybridization of single-stranded clones containing cloned inserts of opposite orientations (22). The double-stranded regions were then measured by electrophoretic sizing of the S1 nuclease (Boehringer Mannheim Biochemicals)-resistant regions or by electron microscopy.

In vivo pulse-labeling of RNA followed by filter hybridizations. The method of tritiated uridine pulse-labeling and RNA extraction has been described previously (10, 11, 25–27). Our pulse-label period was 2 min. In a typical experiment, 2×10^8 cells were suspended in a prewarmed mixture of the following: 12 ml of tritiated uridine, 18 ml of Joklik's modified essential medium containing 5% fetal calf serum, and 0.6 ml of $20\times$ phosphate-buffered saline.

Procedures for quantitative DNA filter hybridizations have been described previously (25–27). The only modification here was that single-stranded DNA solutions were boiled without the addition of NaOH before filtration through nitrocellulose filters.

Isolated nuclear transcriptions and S1 nuclease analysis. Isolated nuclear transcriptions for production of [α - ^{32}P]UTP-labeled nuclear RNA were performed as described previously, with a label time of 10 min (7, 12, 35).

Hybridization of labeled RNA with single-stranded DNA, followed by S1 nuclease digestion was performed as described previously (29, 30), and S1 nuclease-resistant RNA-DNA hybrids were collected according to a protocol kindly provided to us by N. Acheson. The hybridization mixture was first diluted to 0.3 M NaCl, and T1 ribonuclease was added to yield 200 U/ml. A 30- to 500- μg amount of phage DNA with nuclear RNA from 10^8 (early) or 10^7 (late) infected cells was used; this represents a more than 10-fold-greater DNA excess than was used in earlier filter hybridization experiments (25–27). After an incubation period of 30 min at 30°C , proteinase K was added to yield 200 $\mu\text{g}/\text{ml}$, and incubation was continued at 37°C for 30 min. NaCl was added to increase the concentration to 0.75 M, and the solution was then filtered through a nitrocellulose filter (Schleicher & Schuell, Inc.). The nitrocellulose filter was washed with hybridization buffer and transferred to a flat-bottom siliconized glass tube with a diameter slightly larger than that of the nitrocellulose filter. Just enough S1 nuclease buffer (0.3 M NaCl, 0.05 M sodium acetate [pH 4.6] 4.5 mM zinc sulfate) was added to cover the filter. S1 nuclease was then added to yield 10,000 Vogt U/ml, and incubation was carried out for 30 min at 37°C . (Note that 30 Vogt U is equivalent to 1 Ando unit.) The supernatant was removed and kept in a separate tube. The filter was incubated once again with S1 nuclease buffer and enzyme for an additional 0.5 h. This supernatant was pooled with the previous supernatant, and sodium dodecyl sulfate was added to a concen-

FIG. 1. Transcription early in infection from the rightward-reading transcription units. The adenovirus genome (0 to 100, 36 kb) is shown at the bottom of the figure, together with two promoters (arrows) at map units 16.4 (major late promoter) and at 76 (E3) and three of the five late poly(A) sites (upward pointing arrows) at map units 38.5, 49.9 (L2), and 62 (L3). [^3H]uridine-labeled nuclear RNA was prepared after the labeling of 2×10^8 infected HeLa cells (2-min label time, 5 h postinfection). Single-stranded M13 DNA clones containing the regions of the genome indicated by the widths of the bars in the diagram were used to hybridize RNA (also see Table 1). The average of two to four experiments for each clone is given here. The total counts per minute hybridized ranged from 120 to about 3,000 cpm. Background of 30 cpm was subtracted before the counts per minute per base pair was computed.

tration of 0.1%. After phenol-chloroform extraction, 5 μg of carrier tRNA was added, and the nucleic acid was precipitated with ethanol.

The S1 nuclease products were split into two aliquots. One aliquot was run on a 1.5% agarose gel with 5'-end-labeled 123-base-pair and 1-kb DNA marker ladders (Bethesda Research Laboratories, Inc.). Gels were dried and exposed to Kodak XAR-5 film. The other aliquot of S1 nuclease products was run on a 1.5% agarose gel, which was treated with NaOH, and transferred to nitrocellulose as described by Southern (34). The nitrocellulose filters were then hybridized to nick-translated adenovirus DNA probes.

RESULTS

Major late transcript at early infection times: tritiated uridine pulses. A number of examinations have been made, at early times of infection, of RNA complementary to double-stranded genomic fragments representing the major adenovirus late transcript (4, 5, 28, 33, 36). This transcription unit begins at 16.4 map units (11, 16, 34, 36) and continues near to the right end of the genome. Five sites of poly(A) addition are used in late mRNA formation: L1, 38.5, L2, 49.9; L3, 62; L4, 79; and L5, 91.5 (4, 5, 24, 26; Fig. 1 shows positions of L1 through L3). Past hybridization analyses have been complicated by the difficulty in clearly separating measurements of rightward transcription from those of leftward transcription. (For example, the E2 transcription unit that begins at 75 map units [4] and reads leftward is active at early times, complicating measurements in the 50 to 60 map unit region.) Equipped with strand-specific M13 clones, we were able to measure transcription of the stretch of the adenovirus genome from coordinates 14.7 to 78 due solely to the rightward transcription. By examining the RNA extracted from cells infected for 4 h

TABLE 1. DNAs used in band analysis of RNA TRANSCRIPTS

DNA used in (Fig./lane) ^a :	Nucleotide and map units boundaries (map position) ^b		Length	Sequences of note within region
	Left	Right		
3				
1	5,360 (14.7)	7,751 (21.2)	2,391	Major late promoter at 6039 (16.4)
2	11,555 (31.7)	13,636 (37.3)	2,081	
3	13,054 (36.7)	14,456 (40.5)	1,402	L1 poly(A) site at 14,113 (38.5)
4	16,744 (45.9)	18,913 (52.6)	2,169	L2 poly(A) site at 17,948 (49.9)
5	18,913 (52.6)	20,416 (56.9)	1,503	
6	21,816 (60.7)	23,086 (64.2)	1,270	L3 poly(A) site at 22,444 (61.5)
7, 8	25,633 (70.7)	26,369 (72.8)	736	
9	26,369 (72.8)	27,372 (75.9)	1,003	
10	27,372 (76.8)	28,174 (78.4)	802	E3 promoter at 27,609
4				
4, 8	30,046 (84.0)	32,264 (89.7)	2,218	E3 poly(A) site at 30,841 (84.5)
5	32,172 (89.5)	33,090 (91.9)	919	
6, 6a, 6b	33,091 (91.9)	34,933 (97.1)	1,842	
9	32,264 (89.7)	34,933 (97.1)	2,669	
7, 10	34,933 (97.1)	35,937 (100)	1,004	

^a The data for Fig. 4, lanes 1 to 3 were the same as for Fig. 3, lanes 7 to 10.

^b The nucleotide and map positions were taken from the sequence compilation made available from Cold Spring Harbor Laboratories by R. J. Roberts, U. Pettersson, F. Galibert, and J. Sussenbach; also see reference 16.

(early phase of infection), we were able to measure the 3' border of the major late transcript at early-infection times.

This measurement was first carried out by labeling whole infected cells with [³H]uridine for 2 min (10, 27). The labeled RNA was extracted and hybridized to a series of single-stranded cloned DNA samples that were complementary to rightward-transcribed RNA (for some of the exact coordinates, see Table 1). The average amount of labeled RNA (counts per minute per base pair) that hybridized to 11 segments of the genome is given in Fig. 1. There was a continuous decrease in the amount of hybridized RNA moving from the promoter-proximal region to the right until the E3 promoter region was crossed, whereupon an abrupt rise in the amount of hybridizable RNA was seen. Since it was established some time ago (10, 14) that premature termination is a frequent event within the first 2 to 3 kb of the major late transcription unit at late-infection times, it seems reasonable to conclude that the higher amount of newly transcribed RNA complementary to the 14 to 21 and 21 to 28 regions (the two sections to the left in Fig. 1) early in infection is also due to premature termination. Some evidence for premature termination early in infection from this transcription unit has been described previously (23, 33). There was, in addition, a gradual, approximately 2.5-fold reduction of transcription through the 38 to 61 region and a much sharper, 10-fold decline in the 70 to 75 region. This contrasts with the approximately equimolar transcription from about 30 to at least 98 map units observed late in infection (10, 13, 26). Early in infection, the E3 promoter, which is located at 75 map units (4, 11), is active in rightward-reading transcription; this is revealed in Fig. 1 by a sharp increase in hybridized RNA in the 75 to 77 region compared with regions just before map unit 75.

The first, or L1, poly(A) site is at map unit 38.5, the 2nd, or L2, site is at map unit 49, and the third, L3, site is at map unit 62. Early in infection, there is in the cytoplasm at least a 10-fold excess of L1 mRNA compared with L2 mRNA (5, 28, 33) and an even greater excess compared with L3 mRNA. Using double-stranded DNA clones, Nevins and Wilson (28) detected relatively more nuclear poly(A)⁺ RNA from L3 than they found in the cytoplasm but still found the

amount of poly(A)⁺ nuclear L3 RNA to be decreased relative to that of L1 RNA. Together with their results, the data in Fig. 1 show conclusively that transcriptional termination cannot be the major reason for the dominance of L1 mRNA over either L2 or L3 mRNA. However, termination of transcription of RNA initiated at map unit 16.4 does explain why no mRNA of the L4 and L5 groups is formed early.

S1 nuclease analysis of RNA-DNA hybrids. Although the technique of pulse-labeling whole cells and hybridizing the labeled RNA to genomic segments is crucial in identifying which regions of DNA are transcribed within cells, the precise location of termination regions or specific termination sites (if they exist) requires analyzing nascent RNA by techniques that offer resolution at the nucleotide level. In an effort to determine whether favored sites for RNA termination could be observed, T1 ribonuclease treatment of labeled RNA-unlabeled DNA hybrids, followed by gel analysis, has been used (6, 30). In the present experiments, an improved S1 endonuclease assay suggested by N. Acheson was used (see Methods and Materials). Many experiments established that nuclei isolated from adenovirus-infected cells will faithfully elongate RNA chains already initiated in vivo (for reviews, see references 7 and 8). Polymerases move a maximum of ~500 nucleotides (19, 34) and thus produce the closest approximation to unprocessed RNA that is available. Therefore, we used isolated nuclei to generate ³²P-labeled nascent RNA chains which were then hybridized to single-stranded DNA probes. The hybrids were trapped on nitrocellulose filters, and the filters were soaked in buffer with S1 nuclease to cut single-stranded RNA and DNA. This step released RNA-DNA duplexes from the filters so that they could be collected and analyzed by gel electrophoresis.

The results of applying this procedure to late RNA complementary to three segments of the adenovirus genome included in the major late transcript are given in Fig. 2. The three regions used (10.9 to 18.2, 14.7 to 21.2, and 31.7 to 37.3) show that the technique is useful in defining discrete-sized bands. The major late promoter is known to lie at nucleotide 6,039 (16). The right ends of the adenovirus DNA were included in the two single-stranded clones used in lanes

FIG. 2. Nascent RNA from late-infected cells analyzed by S1 nuclease band analysis. Nascent ^{32}P -labeled RNA from nuclei of late-infected cells was analyzed by autoradiography after hybridization to single-stranded M13 clones containing adenovirus DNA, S1 nuclease treatment, and gel electrophoresis. Lane M is marker DNA; sizes (in nucleotides) are indicated. Lanes 1, RNA hybridized to M13 with no inserted adenovirus DNA; 2, adenovirus DNA, *Sma*I fragment (map units 10.89 to 18.2, rightward-reading RNA); 3, adenovirus DNA, *Bam*HI fragment (map units 14.7 to 21.2, rightward-reading RNA); 4, adenovirus DNA, *Hind*III fragment (map units 31.7 to 37.3, rightward-reading RNA).

2 and 3, at nucleotides 6,580 and 7,760, respectively. Thus, labeled RNA initiating at the major late cap site hybridized to these cloned single-stranded DNAs and treated with S1 nuclease should produce bands of 541 and 1,721 nucleotides. The observed RNA-DNA duplex bands (lanes 2 and 3) were very close to these sizes, as judged from known DNA markers and recognizing that RNA-DNA duplexes may not run exactly as does DNA. Likewise, the RNA complementary to the single-stranded clone used in lane 4 should match the entire length of the inserted DNA to produce an S1 nuclease-resistant hybrid of about 2 kb. This was also observed (lane 4). In lanes 2 to 4, prominent bands that were less than full length could also be observed. These bands could represent (i) pause sites (Since RNA polymerases move only 500 nucleotides or less, regions that are hard to traverse could be reflected as bands.), (ii) termination sites, or (iii) true RNA processing sites or regions of hybrid susceptible to easy S1 nuclease cleavage. If such bands were

true termination sites, then the full-length band should not be visible, or at least it should not be the most prominent band. With this assay, we proceeded to examine labeled nuclear RNA early after infection for evidence of specific termination sites within the major late promoter region.

The analysis of RNA from early-infected cells that was complementary to the region of the major late transcript is shown in Fig. 3 and Table 1. In lane 1 of Fig. 3, a band of labeled S1 nuclease-resistant RNA-DNA hybrid of about 1,700 nucleotides was observed just as was observed when late RNA was analyzed in this region. This band conforms to the previous finding (4) that the same cap site (16.5) for beginning RNA synthesis is used early and late after infection. In addition to the major band, a weaker band, of approximately 2,400 nucleotides (i.e., full length for the inserted sequence), was observed, which indicates a low level of transcription that may well have originated from the E1A and 1B promoters and read across the entire region.

RNA bands of similar intensity that matched the full length of the cloned DNA inserted into various M13 isolates were observed in the next five lanes (Fig. 3, lanes 2 to 6). This result indicates that RNA polymerases are distributed and continue to transcribe at similar frequency throughout this 16 to 65 region of the genome. Bands smaller than full length were observed in several instances. The shorter band in the L1 region (Fig. 3, lane 3) is the correct size to represent RNA transcribed by a polymerase whose primary transcript has been cut at L1 but which has continued transcription. This phenomenon of RNA synthesis downstream from poly(A) sites at which cleavage has already occurred was described earlier (27) and is of course necessary to explain the approximately equimolar distribution of nascent RNA across the entire genome seen late in infection. The prominent shorter bands in lanes 5 and 6 of Fig. 3 are not explained; neither of them matches a nascent RNA downstream of a poly(A) site. These bands could conceivably be due to unpaired regions in the RNA-DNA hybrids caused by a microheterogeneity in adenovirus DNAs that has been identified during sequencing studies of the adenovirus genome (16) or could be due to deletions in the M13 probe used. We do not think the latter is the case because the single-stranded DNA was checked by sizing the complementary S1 nuclease-resistant DNA formed when two phage DNAs containing inserts of opposite direction were hybridized (26). In some cases, the inserts in replicative form DNA prepared at the same time as single-stranded DNA were also analyzed. Whatever the cause of the shorter bands in lanes 5 and 6 of Fig. 3, the full-length bands in these lanes are as strong as those in lanes 2, 3 and 4; thus, the shorter bands cannot be taken as indications of termination sites.

The transcription of the 70 to 75 region was shown in Fig. 1 to be diminished sharply compared with the preceding DNA regions. This was also observed in Fig. 3, lanes 7, 8, and 9, which represent nascent RNA complementary to map units 70.7 to 72.8 (both lanes 7 and 8) and 72.8 to 75.9. (A similar result will be described in Fig. 4 in a separate experiment.) The band lengths for both the 70.7 to 72.8 and the 72.8 to 75.9 segments match that of the adenovirus DNA inserted into M13, although the intensity of the bands is decreased. Finally, the intense band seen in lane 10 of Fig. 3 corresponds to the length from the known E3 promoter to the right end of the adenovirus insert in the M13 clone used for this lane. In addition, a faint band was seen which represents RNA of length equal to that from the insert and which indicates a low level of continuous transcription from the major late promoter.

FIG. 3. Nascent RNA from early-infected cells analyzed by S1 nuclease band analysis. Shown is the major late transcription unit. The P at the top of the figure denotes the major late promoter at map unit 16.4; L1, L2, and L3 are the three poly(A) sites within the region analyzed. The four dark arrows in the second line of the figure show the position of rightward-reading early promoters E1A, E1B, major late, and E3. The double strand is the adenovirus genome; the single lines underneath are regions present in each single-stranded M13 phage used in experiments shown below. Approximate coordinates for DNA used in each lane are given to the upper left of the figure; see Table 1 for exact coordinates. RNA from early-infected cells was hybridized to these various segments of the adenovirus genome; the length of labeled RNA complementary to each DNA region was determined by the S1 endonuclease assay. The genomic region examined in this figure extends from the region (14 to 21) containing the major late start site at map unit 16.4 (lane 1) to and beyond the junction of the major late transcription unit with the E3 transcription unit at map unit 74 (lane 10). Table 1 gives the precise coordinates of the clones used, as well as the estimated lengths of major S1 nuclease-resistant RNA-DNA duplexes. Lanes 1 to 7 all consisted of equal samples of the same labeled nuclear RNA preparation. Lanes 8 to 10 all consisted of another similar nuclear RNA sample. Lanes 7 and 8 consisted of the same DNA clone. Open arrows next to each lane indicate bands that are full length; i.e., bands in which the protected RNA-DNA duplex matches the length of the DNA insert in the M13 clone. Small filled arrows (lanes 1 and 10) indicate less-than-full-length bands that mark the initiation sites of the major late and E3 primary transcripts within the boundaries of the M13 clones used in lanes 1 and 10. The sizes to the left of the lanes (in nucleotides) were taken from double-stranded DNA markers of known size (see Fig. 2).

These latter two points are reemphasized in the results of Fig. 4, in which RNA complementary to three genomic segments (70.7 to 72.8, 72.8 to 75.9, and 75.9 to 77) was analyzed. Again, full-length bands were observed in all three regions with a strong shorter band that maps to the E4 promoter (Fig. 4, lane 3). These results strongly support the conclusion of Fig. 1. Termination of transcription is not effective before L3 has been crossed, and then a gradual, not a single, site-specific type of termination seems to occur for most, but not all, polymerases before the E3 promoter is crossed. E3 transcriptional activity then greatly exceeds any continuing rightward transcription from the major late promoter.

Analysis of the E3 transcription unit. By choosing clones from map unit 75 to the end of the genome to analyze the nuclear RNA, we could explore termination within the E3

transcription unit. Transcription was active beyond the two E3 poly(A) sites, as shown by intense full-length RNA bands after hybridization to the 84 to 89 and 89 to 91 regions (lanes 4 and 5, Fig. 4, and lanes 4' and 5', Fig. 5). Much less hybridization was observed in the 91 to 97 region, and virtually no hybridization was found in the 97 to 100 region. (Compare the level of exposure of lanes 4 to 7 in Fig. 4 with that of lanes 4' to 7' in Fig. 5.) Thus, the E3 primary transcripts extend well beyond the two major poly(A) sites but do not continue indefinitely (lanes 4 and 8 of Fig. 4 are independent experiments using the same DNA region; compare lane 4 with lanes 5 and 10). There were shorter-than-full-length bands in all of the E3 regions downstream from the second E3 poly(A) site, and in fact, when DNA from map units 91 to 97 and 97 to 100 was used, very little full-length RNA was observed (Fig. 4, lanes 5, 6, 7, and 9). The two

FIG. 4. Nascent RNA from early-infected cells analyzed by S1 nuclease band analysis. Shown are the E3 transcription unit and its junction with the major late transcription unit. RNA was prepared and analyzed as described in the legend to Fig. 3. The autoradiograms are placed directly above or below the single lines representing single-stranded cloned DNA that was used for that analysis (e.g., lanes 1, 2, and 3 utilized the three left-most clones indicated below the genomic line). Approximate genomic locations of M13 DNA used are given in numbers to the bottom left of the figure; the exact locations of clones used are given in Table 1. Lanes 1 to 7 are from one experiment. Lanes 6a and 6b consist of RNA from a separate experiment in which the same clone (map units 91 to 97) as in lane 6 was used but in which early (6a)- is compared with late (6b)-infection-labeled nuclear RNA. Lanes 8, 9, and 10 represent another RNA preparation, and lanes 4 and 8 represent analysis with the same DNA clone. Lane 9 represents the region examined by two clones in lanes 5 and 6. Lanes 7 and 10 consisted of the same clone used on different occasions.

FIG. 5. Analysis of RNA complementary to 3' regions of the E3 transcription unit. The autoradiogram is of the same gels as in Fig. 3, lanes 4 to 7, and the lanes are therefore labeled 4' to 7'. Autoradiographic exposure was for less time (4 h instead of 2 days). Comparison of Fig. 4 and 5 and reference to Table 1 show that termination is virtually complete before the position of 92 map units early in infection (compare lanes 5' and 6'). No transcription beyond 97 map units was detectable (compare lane 7', Fig. 5, with lane 7, Fig. 4).

major bands seen in both lanes 6 and 6a of Fig. 4 (two separate experiments), added together, (1,350 plus 460) were equal to a full-length band. This contrasts with RNA from nuclei late in infection (Fig. 4, lane 6b) which gave mostly full-length labeled RNA. It is possible that the RNA transcribed in this 1,800-nucleotide region early in infection was cleaved during transcription, or there may be a preferential S1 nuclease cut site in RNA-DNA hybrids in this region. We do not know the cause for such a cleavage at this site, but it was reproduced in several experiments. As noted above (Fig. 4, lane 6), the cleavage was not apparent in late-infection RNA. Also, when the companion bacteriophage carrying the opposite strand of the same genomic region was used to analyze early-infection RNA originating from the leftward-reading E4 transcription unit, a full-length band was obtained (data not shown). Thus, the most plausible explanation is that the rightward-reading RNA in this region (map units 91 to 97) is simply very easily cleaved at a site 1,350 nucleotides from map unit 91 during early infection. Since the two bands are of equal intensity, they cannot be explained by a termination event in any case. The main conclusion from Fig. 4 and 5 is that hybridization to the 84 to

89 and 89 to 91 regions was much more intense than it was to regions downstream from 90. Thus, it is again clear that a region of termination exists, but no specific stop site was detected that might be responsible for the majority of termination.

DISCUSSION

This study reemphasizes the importance of poly(A) choice early in infection (28) that favors the L1 mRNAs compared with the L2 and L3 mRNAs. The sections of the transcription unit encoding L4 and L5 mRNAs, however, are clearly shown not to be transcribed at significant rates early in infection. Thus, the gradual termination events observed in the ~38 to 75 region are responsible for preventing L4 and L5 mRNA formation early in infection.

Evidence that the E3 transcription unit behaves similarly to the major late transcription unit is also provided. Transcription decreased gradually, beginning about 1 to 2 kb beyond the final poly(A) site in that transcription unit, and had ceased almost entirely before the last 2,500 to 3,000 nucleotides at the right end of the genome. One possible event that could color the quantitative aspects of these experiments is RNA accumulation of leftward-reading RNA in the nucleus from regions downstream of poly(A) sites that would in hybridization reactions form RNA hybrids with rightward-reading RNA. Such a result could hypothetically lead to an interpretation of termination when in reality the RNA simply could not be observed. This hypothesis seems highly unlikely because the most prominent steady state leftward-reading RNA would be the stable E4 and E2A mRNAs. These would produce discrete bands of nascent labeled RNA hybridized with them. No such bands were ever seen, and we therefore discount the possibility that leftward-reading RNA interferes substantially with these experiments.

As we have noted before, no specific strong termination sites are known for any eucaryotic RNA polymerase II transcription units (6, 12). Rather, after a poly(A) site has been chosen, there is a region of gradual decrease. In the β -major globin transcription unit, this region appears to stretch over about 500 nucleotides. In the virus transcription units discussed here, the region of gradual decrease is much larger (~3 to 10 kb, between map units 40 and 70 for the major late transcription unit and between map units 85 and 95 for the E3 transcription unit). From the present experiments, together with our early observations on premature termination, first in adenovirus (10, 14), and then in whole cells (32), and the recent work with β -major globin transcription termination (6, 12), the following picture of RNA polymerase II termination seems plausible. The polymerase complex can assemble either with or without an antitermination or elongation factor. Without this putative factor, premature termination occurs; with the factor, premature termination is avoided, and termination will not then occur until after a poly(A) site has been selected and acted upon. Transcription then continues (in the cases studied here, for very long stretches) until the event of termination can occur. This second step appears to require specific sequences based on results with the β -major globin termination region (6, 12), even though there may be no single favored nucleotide site for termination. With the identification of segments of the adenovirus genome (map units 72 to 75 and 89 to 95) in which termination seems prominent, it would now be reasonable to determine whether these regions do in fact confer ability for termination in recombinant DNAs.

ACKNOWLEDGMENTS

We thank M.-T. Hsu, N. Reich, A. Chaudhuri, and M. Salditt-Georgieff for discussions. S. Iwamoto was a predoctoral trainee on PHS AI07233-10.

This work was supported by Public Health Service grants CA16006-12, CA18213-10, and CA25545 from the National Institutes of Health and by grant MV2710 from the American Cancer Society.

LITERATURE CITED

- Adhya, S., and M. Gottesman. 1982. Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell* 29:939-944.
- Akusjarvi, G., and H. Persson. 1981. Controls of RNA splicing and termination in the major late adenovirus transcription unit. *Nature (London)* 292:420-422.
- Amara, S. G., R. M. Evans, and M. G. Rosenfeld. 1984. Calcitonin/calcitonin gene-related peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites. *Mol. Cell. Biol.* 4:2151-2160.
- Baker, C. C., and E. B. Ziff. 1981. Promoters and heterogeneous 5'-termini of the messenger RNAs of adenovirus type 2. *J. Mol. Biol.* 149:189-222.
- Chow, L. T., and T. Broker. 1978. The spliced structure of Ad2 fiber message and other late mRNAs. *Cell* 15:497-508.
- Citron, B., E. Falck-Pedersen, M. Salditt-Georgieff, and J. E. Darnell, Jr. 1984. Transcription termination occurs within a 1,000 base pair region downstream from the poly(A) site of the mouse β -globin (major) gene. *Nucleic Acids Res.* 12:8723-8731.
- Cross, F. R., and J. E. Darnell, Jr. 1983. Cycloheximide stimulates early adenovirus transcription if early gene expression is allowed before treatment. *J. Virol.* 45:683-692.
- Darnell, J. E., Jr. 1982. Variety in the level of gene control in eukaryotic cells. *Nature (London)* 297:365-371.
- Eggerding, F. A., and W. C. Pierce. 1983. Construction of a cloned library of adenovirus DNA fragments in bacteriophage M13. *J. Biol. Chem.* 258:10090-10097.
- Evans, R., J. Weber, E. Ziff, and J. E. Darnell. 1979. Premature termination during adenovirus transcription. *Nature (London)* 278:367-370.
- Evans, R. M., N. Fraser, E. Ziff, J. Weber, M. Wilson, and J. E. Darnell. 1977. The initiation sites for RNA transcription in Ad2 DNA. *Cell* 12:733-739.
- Falck-Pedersen, E., J. Logan, T. Shenk, and J. E. Darnell, Jr. 1985. Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse β -major globin terminator element. *Cell* 40:897-905.
- Fraser, N. W., J. R. Nevins, E. Ziff, and J. E. Darnell. 1979. The major late adenovirus type-2 transcription unit: termination is downstream from the last poly(A) site. *J. Mol. Biol.* 129:643-656.
- Fraser, N. W., P. B. Sehgal, and J. E. Darnell. 1979. Multiple discrete sites for premature RNA chain termination late in adenovirus-2 infection: enhancement by 5,6-dichloro-1- β -ribofuranosylbenzimidazole. *Proc. Natl. Acad. Sci. USA* 76:2571-2575.
- Frayne, E. G., E. J. Leys, G. F. Crouse, A. G. Hook, and R. E. Kellems. 1984. Transcription of the mouse dihydrofolate reductase gene proceeds unabated through seven polyadenylation sites and terminates near a region of repeated DNA. *Mol. Cell. Biol.* 4:2921-2924.
- Gingras, T. R., D. Sciaky, R. E. Gelinos, J. Bing-Dong, C. E. Yen, M. M. Kelly, P. A. Bullock, L. O'Neill, and R. J. Roberts. 1982. Nucleotide sequences from the adenovirus-2 genome. *J. Biol. Chem.* 257:13475-13491.
- Greenblatt, J. 1981. Regulation of transcription by the N gene of bacteriophage lambda. *Cell* 24:8-9.
- Hagenbuchle, O., P. K. Wellauer, D. L. Cribbs, and U. Schibler. 1984. Termination of transcription in the mouse α -amylase gene *Amy-2^a* occurs at multiple sites downstream of the polyadenylation site. *Cell* 38:737-744.
- Hofer, E., and J. E. Darnell, Jr. 1981. The primary transcription unit of the mouse β -major globin gene. *Cell* 23:585-593.
- Hofer, E., R. H.-W. Hofer, and J. E. Darnell, Jr. 1982. Globin RNA transcription: a possible termination site and demonstration of transcription control correlated with altered chromatin structure. *Cell* 29:887-893.
- Holmes, W. M., T. Platt, and M. Rosenberg. 1983. Termination of transcription of *E. coli*. *Cell* 32:1029-1032.
- Iwamoto, S., and M.-T. Hsu. 1983. The twist and handedness of a 39 base pair segment of DNA in solution. *Nature (London)* 305:70-72.
- 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.
- McGrogan, M., and H. J. Raskas. 1978. Two regions of the adenovirus 2 genome specify families of late polysomal RNAs containing common sequences. *Proc. Natl. Acad. Sci. USA* 75:625-629.
- Nevins, J. R. 1980. Definition and mapping of adenovirus-2 nuclear transcription. *Methods Enzymol.* 65:765-785.
- Nevins, J. R., and J. E. Darnell. 1978. Groups of adenovirus type 2 mRNAs derived from a large primary transcript: probable nuclear origin and possible common 3' ends. *J. Virol.* 25:811-823.
- Nevins, J. R., and J. E. Darnell, Jr. 1978. Steps in the processing of AD-2 mRNA: poly(A)⁺ nuclear sequences are conserved and poly(A) addition precedes splicing. *Cell* 15:1477-1493.
- Nevins, J. R., and M. C. Wilson. 1981. Regulation of adenovirus-2 gene expression at the level of transcription termination and RNA processing. *Nature (London)* 290:113-118.
- Padgett, R. A., S. F. Hardy, and P. A. Sharp. 1983. Splicing of adenovirus RNA in a cell-free transcription system. *Proc. Natl. Acad. Sci. USA* 80:5230-5234.
- Salditt-Georgieff, M., and J. E. Darnell. 1983. A precise termination site in the mouse β major globin transcription unit. *Proc. Natl. Acad. Sci. USA* 80:4694-4698.
- Salditt-Georgieff, M., M. Harpold, S. Chen-Kiang, and J. E. Darnell, Jr. 1980. The addition of 5' cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped. *Cell* 19:69-78.
- Salditt-Georgieff, M., M. Sheffery, K. Krauter, J. E. Darnell, Jr., R. Rifkind, and P. A. Marks. 1984. Induced transcription of the mouse β -globin transcription unit in erythroleukemia cells. *J. Mol. Biol.* 172:437-450.
- Shaw, A. B., and E. B. Ziff. 1981. Transcripts from the adenovirus 2 major late promoter yield a single family of 3' co-terminal mRNA and five late families. *Cell* 22:905-916.
- Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Weber, J., W. Jelinek, and J. E. Darnell. 1977. The definition of a large viral transcription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei. *Cell* 10:611-616.
- Yanofsky, C. 1981. Attenuation in the control of expression of bacterial operons. *Nature (London)* 289:751-758.
- Ziff, E., and R. M. Evans. 1978. Coincidence of the promoter and capped 5' terminus of RNA from the adenovirus type-2 major late transcription unit. *Cell* 15:1463-1475.
- Ziff, E., and N. Fraser. 1978. Adenovirus type 2 late mRNAs: structural evidence for 3'-coterminal species. *J. Virol.* 25:897-906.