Transcription of Adenovirus Type 2 Genes in a Cell-Free System: Apparent Heterogeneity of Initiation at Some Promoters

DAVID C. LEE†* AND ROBERT G. ROEDER

Departments of Biological Chemistry and Genetics, Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri 63110

Received 17 February 1981/Accepted 23 April 1981

We examined the transcription of a variety of adenovirus type 2 genes in a cell-free system containing purified ribonucleic acid polymerase II and a crude extract from cultured human cells. The early E1A, E1B, E1II, and E1IV genes and the intermediate polypeptide IX gene, all of which contain a recognizable TATAA sequence upstream from the cap site, were actively transcribed in vitro, albeit with apparently different efficiencies, whereas the early E1I (map position 74.9) and IVa₂ genes, both of which lack a TATAA sequence, were not actively transcribed. A reverse transcriptase-primer extension analysis showed that the 5' ends of the in vitro transcripts were identical to those of the corresponding in vivo ribonucleic acids and that, in those instances where initiation was heterogeneous in vivo, a similar kind of heterogeneity was observed in the cell-free system. Transcription of the polypeptide IX gene indicated that this transcript was not terminated at, or processed to, the polyadenylic acid addition site in vitro. We also failed to observe, using the in vitro system, any indication of transcriptional regulation based on the use of adenovirus type 2-infected cell extracts.

Accumulating evidence suggests that control of eucaryotic gene expression can be exerted, at least in part, through changes in the number of ribonucleic acid (RNA) polymerase molecules actively transcribing a gene (11, 24, 25, 27, 29, 34, 39). An understanding of the molecular mechanisms by which specific regulatory signals can influence levels of transcription requires a detailed analysis of the components and events governing initiation and elongation of RNA transcripts. The recent development of cell-free transcription systems which direct accurate initiation at specific genes (22, 41, 43), as well as the partial purification of required factors (23, 37), opens the way toward such an analysis. For example, it is now possible to define a region of the *Xenopus* 5S gene critical for its correct transcription in vitro by RNA polymerase III (6, 36) and to show that one of the required factors binds with high affinity to that region (12). Similarly, using in vitro mutagenesis coupled with in vitro transcription by RNA polymerase II, it has been possible to probe the involvement of deoxyribonucleic acid (DNA) sequences around the 5' ends of the chicken conalbumin (10, 40) and adenovirus type 2 (Ad2) major late (10) transcription units, thereby demonstrating the

critical role of the Goldberg-Hogness box (M. Goldberg, Ph.D. thesis, Stanford University, Stanford, Calif., 1978). Finally, in the first clear example of reconstituted regulation involving RNA polymerase II, the D2 tumor antigen has been shown to selectively inhibit in vitro transcription from the simian virus 40 early promoter, consistent with its postulated role in vivo (35).

We have chosen to apply the technique of in vitro transcription to an analysis of viral gene regulation during the course of infection of human KB cells with Ad2. The Ad2 infectious cycle is characterized by a clearly defined temporal program of gene expression; it is conventionally divided into early and late stages, separated by the onset of viral DNA replication roughly 8 h after infection (for a recent review, see reference 45). During both the early and late stages, transcription is by the host RNA polymerase II. There are at least six early transcription units (see Fig. 1A), designated E1A, E1B, E1II through E1IV, and IVa₂, and transcription rate measurements reveal a timed sequence of activation of these regions (27). E1A is expressed within 1 h of infection, followed by E1B, E1II, and E1IV; E1II is then activated after 2 to 3 h. Studies involving the addition of protein synthesis inhibitors (20), as well as the use of viral host

† Present address: Department of Pharmacology, University of Washington, Seattle, WA 98195.

range mutants (3, 18, 19), suggest a role of the product(s) of the region(s) expressed earliest in the infectious cycle in the succeeding activation of other early regions. At 4 to 6 h after infection, the gene coding for polypeptide IX (pIX; a structural component) is expressed (30, 31). The late stage is characterized by reduced transcription of the EII and EIV regions, possibly resulting from the accumulation of a viral protein coded for by the EII region (5, 7, 8, 28), and by a massive increase in the abundance of late transcription unit messenger RNAs (mRNA's). The extent to which the latter is a result of activation or enhancement of expression of the major late promoter is unclear, since recent reports indicate that this promoter is active at early times of infection (9, 39).

Our ultimate goal in these studies is to reproduce the temporal regulation of Ad2 gene expression in vitro, using purified or partially purified components. The wealth of information concerning the structure and sequence of Ad2 genes, the detailed positioning of the 5' ends of the mRNA's, the considerable effort aimed at the purification and characterization of viral proteins, and the relative feasibility of isolating native templates make Ad2 an attractive system for such an analysis. As an initial step towards this goal, we have extended a previous observation from this laboratory that the major late promoter of Ad2 is transcribed selectively and accurately in an in vitro system comprising purified RNA polymerase II and a crude cell-free extract (S100), both prepared from uninfected human KB cells (41). We have analyzed all of the principal promoters expressed early in the infectious cycle, as well as that of the pIX gene, which appears to be expressed at intermediate times (30, 31). Using both cloned viral fragments and whole Ad2 DNA, we find that the promoters for the early EIA, EIB, EIII, and EIV genes, as well as that of the pIX gene, direct transcription in vitro, albeit with apparently different efficiencies. In contrast, the promoters for the early EII and IVa₂ genes, both of which lack obvious Goldberg-Hogness boxes, are recognized in vitro with greatly reduced efficiencies, if recognized at all. Detailed analysis of the 5' ends of the in vitro transcripts suggests that in each case initiation occurs at the same base(s) used in vivo, and in those instances in which the initiation in vivo is heterogeneous, a similar kind of heterogeneity is observed in vitro. Finally, a comparison of extracts prepared from either uninfected KB cells or cells at various times after infection with Ad2 does not reveal any significant difference in the relative efficiencies with which the viral promoters are utilized in vitro.

MATERIALS AND METHODS

Procedures for the preparation of S100 extracts, purification of RNA polymerase II from human cultured (KB) cells, preparation of DNA templates, and in vitro synthesis, purification, and gel electrophoresis of RNA have all been previously described (21, 41). Optimal conditions for in vitro transcription of the various Ad2 genes do not differ significantly and are similar to those previously reported.

Preparation of early and late Ad2 poly(A)⁺ RNA. KB cells (7.5×10^6 per ml) were infected with Ad2 at multiplicities of infection of either 100 or 40 for the preparation of early and late viral RNA, respectively. After incubation for 1 h at 37°C to allow for viral adsorption, cells were diluted to 3×10^5 per ml, and in the case of early RNA preparations cycloheximide was added to 25 µg/ml. Cytoplasmic RNA was isolated by the method of Zieve and Penman (44) at either 6 h for early RNA or 24 h for late RNA. Polyadenylic acid-containing [poly(A)⁺] RNA was selected by oligodeoxythymidylic acid chromatography. RNA was also prepared in a similar manner from uninfected KB cells.

Reverse transcriptase extension procedures. Analysis of 5' ends of unlabeled transcripts was carried out with either total RNA from in vitro transcription reactions or specific transcripts isolated from preparative polyacrylamide gels. In the former case, transcription reactions were carried out as usual (except that no carrier RNA was added when terminating the reaction) and nucleic acid was extracted with phenol and chloroform, precipitated with ETOH, and resuspended in 125 µl of 50 mM tris(hydroxymethyl)aminomethane (pH 7.9)–10 mM MgCl₂–50 µg of iodoacetate-treated deoxyribonuclease I per ml. After incubation for 10 min at 37°C, RNA was again extracted and ETOH precipitated. In the latter case, nucleic acid was extracted and ETOH precipitated, and the resulting runoff RNAs were localized by comparison with labeled markers after polyacrylamide gel electrophoresis. The transcripts were eluted from gel slices, purified by chromatography on methylated albumin-kieselguhr columns (16), and ETOH precipitated with 2 to 3 µg of *Escherichia coli* transfer RNA.

DNA primers were prepared from cloned viral fragments by successive rounds of digestion with restriction enzymes and electroelution. DNA fragments were 5'-end labeled as described before (17).

5'-End-labeled [³²P]DNA primer (5×10^4 to 10×10^4 cpm) was ETOH precipitated along with 1 to 4 µg of early or late Ad2 poly(A)⁺ RNA, the total RNA isolated from a standard in vitro transcription reaction, or the total gel-purified transcript from a 10× in vitro reaction. The samples were resuspended in 10 µl of 0.1 M NaCl–20 mM tris(hydroxymethyl)aminomethane (pH 7.9)–0.1 mM ethylenediaminetetraacetic acid and then drawn into 10-µl capillary pipettes. The pipettes were flamed sealed, and the samples were boiled for 2 min and then incubated at 60°C for 3 to 5 h. Samples were expelled from the capillaries into an equal volume of 80 mM tris(hydroxymethyl)aminomethane (pH 7.9)–10 mM MgCl₂–4 mM dithiothreitol–0.4 mM each deoxyadenosine, deoxycytidine, deoxyguanosine, and deoxythymidine triphosphates. Reverse transcriptase

was added to 7 U per reaction, and the samples were mixed and allowed to sit on ice for 5 to 10 min. After further incubation for 30 min at 37°C, the samples were extracted with phenol and chloroform, ETOH precipitated, and dried thoroughly. The precipitates were resuspended in 5 μ l of 90% formamide-containing dyes and analyzed on an 8% 0.5 $\times$ tris(hydroxymethyl)aminomethane-borate-ethylenediaminetetraacetic acid thin sequencing gel.

RESULTS

A map of the Ad2 transcription units whose promoters were analyzed in this study is shown in Fig. 1A. Promoters were tested for their ability to direct transcription in vitro, using a system comprising uninfected KB cell S100 and purified KB cell RNA polymerase II. The transcription templates consisted of fragments of Ad2 DNA, containing one or more viral promoters, cloned into plasmid pBR322. Initiation of transcription at specific sites in vitro was initially scored by using a runoff assay, as described previously (21, 41). Briefly, template molecules were cleaved with restriction enzymes at variable distances downstream from the promoter region and then tested for their ability to direct the synthesis of appropriately sized runoff RNAs. Diagrams of the cloned fragments, as well as of the locations of the restriction enzyme sites used to generate the runoff templates, are shown in Fig. 1B. The sizes of the various runoff transcripts, as well as those of the extension products described in the 5'-end analysis, are predicted from DNA sequence information provided by R. Roberts (personal communication).

DNA sequences surrounding the putative Ad2 promoters are shown in Table 1. Comparison of the various sequences indicates that, with the exception of the early EII promoter at map position 74.9 and the IVa₂ promoter, all of the genes examined in this report contain sequences homologous to the Goldberg-Hogness box. In addition, Table 1 also indicates the positions of the various cap sites; as discussed below, C. Baker and E. Ziff (manuscript submitted for publication) have analyzed the 5' ends of Ad2 mRNA's and shown that several of these promoters display microheterogeneity with respect to the cap site (see below).

Analysis of Ad2 early gene transcription in vitro. The activity of the EIA promoter (map position 1.5) was examined with an Ad2 DNA fragment which extends from near the left end of the Ad2 genome to the *Hpa*I site at map position 4.5 and which is cloned into the *Pst*I site of pBR 322 by GC tailing (Fig. 1Bi; kindly provided by M. Wilson and J. Darnell). Cleavage of this template with either *Sma*I or *Xba*I before

transcription should result in runoff transcripts of 510 or 843 nucleotides, respectively. Figure 2A shows a gel analysis of the total products obtained from in vitro transcription of the *Sma*I (lane 1)- or *Xba*I (lane 3)-cleaved templates; comparison of the position of the prominent bands (arrows) with denatured DNA standards indicates that, consistent with the expected values, the *Sma*I and *Xba*I products are approximately 500 and 800 nucleotides in length, respectively (data not shown). A more careful comparison of the position of the *Sma*I-cleaved transcript with marker RNAs of 533, 404, and 197 nucleotides (transcripts derived from the major late promoter; see reference 41) indicates a size of 510 bases (Fig. 3, lane 4). The *Sma*I- and *Xba*I-cleaved transcripts are not present when RNA polymerase II is selectively inhibited by the addition of α -amanitin to the transcription reaction (Fig. 2A, lanes 2 and 4). These results suggest that RNA polymerase II, in the presence of KB cell S100, initiates at or near the putative in vivo EIA promoter and transcribes in the correct direction. Using a similar KB S100 extract, Corden et al. (10) have also recently observed an in vitro transcript whose approximate size is consistent with correct initiation at the EIA promoter. The labeled material at or near the top of the gel in lanes 1 and 3 has not been further characterized; it is, however, invariably present, is presumably generated by RNA polymerase II, and may result from transcription of plasmid sequences (see reference 41).

The activity of the EIB promoter (map position 4.5) was examined with a cloned Ad2 DNA fragment which extends from near 0 at the left end of the Ad2 genome to the *Hind*III site at map position 7.9 (Fig. 1Bii). This fragment was inserted after the addition of an *Eco*RI linker to the left end, into the *Eco*RI and *Hind*III sites of pBR322 (kindly provided by G. Chinnadurai). Cleavage of this template with either *Kpn*I or *Hind*III should yield transcripts of either 347 or 1,105 nucleotides, respectively, if transcription initiates at the putative in vivo promoter. Gel analysis of the products obtained from in vitro transcription of these templates reveals, in each case, a single prominent transcript (Fig. 2B, lanes 1 and 3) whose size relative to DNA markers is consistent with the predicted value (data not shown). A comparison of the *Kpn*I transcript with RNA markers indicates a size of 349 nucleotides (Fig. 3, lane 5). That these transcripts are produced by RNA polymerase II is again demonstrated by their sensitivity to low levels of α -amanitin (Fig. 2B, lanes 2 and 4). Manley et al. (22), using a whole-cell extract,

FIG. 1. Diagrams of Ad2 transcription units and truncated templates used in the analysis of promoter activity in vitro. (A) The approximate positions of the early and late Ad2 transcription units examined in

have recently observed an in vitro transcript whose size is consistent with correct initiation at the EIB promoter.

The Ad2 DNA fragment used to assay for the EIB promoter also contains the EIA promoter at map position 1.5. Digestion of this template with *Sma*I yields a runoff transcript (data not shown) identical in size to that of the presumed EIA product shown in Fig. 2A, lane 1. However, the efficiency of EIA transcription from the larger fragment is lower by approximately two- to threefold. Whether this is the result of (i) competition with the apparently stronger (see below) EIB promoter, (ii) a deletion, during cloning, of sequences at the left-hand end of the fragment which are important for optimal expression of the EIA promoter, or (iii) some other phenomenon related to the site or method of cloning is not known.

The activities of the early EII (map position 74.9) and EIII (map position 76.9) promoters were examined with a cloned *Hind*III-Ad2 DNA fragment which extends from map positions 73.6 to 79.9 (Fig. 1Biii); this fragment was cloned by insertion into the *Hind*III site of pBR 322 (kindly provided by S.-L. Hu). Cleavage of this template with *Bgl*II should yield a runoff transcript from the EIII promoter of 570 nucleotides, whereas cleavage with *Hind*III should generate transcripts of 710 bases from the EII promoter and 1,045 bases from the EIII promoter. Digestion with both *Hind*III and *Bgl*II should yield transcripts of 719 and 570 nucleotides from the EII and EIII promoters, respectively. The results obtained from in vitro transcription of this template are shown in Fig. 4A. Transcription of the *Hind*III-cleaved template yields a single, prominent, α -amanitin-sensitive band (lanes 1 and 2) whose size by comparison with DNA standards is approximately 1,000 nucleotides (data not shown). Transcription of both the

this report are shown as a percentage of the genome length from the left end. Each percentage point represents about 350 base pairs of DNA. The locations of the corresponding cap sites are indicated (●). Diagrams of the various truncated templates used in the analysis of promoter activity in vitro are shown in (B). Ad2 insert DNA is denoted by straight lines, and plasmid sequences are denoted by jagged lines. The positions of cap sites are shown as (▽), and the positions of restriction enzymes used to generate truncated template are also indicated. The numbers immediately below these lines show the approximate map positions in whole Ad2 DNA. The heavy arrows below the restriction maps indicate the direction of transcription; the numbers immediately above the arrows correspond to the distance, in bases, from the putative cap sites to the ends of the coding strands for each restriction site.

TABLE 1. DNA sequence upstream from Ad2 cap sites

Promoter	TATAA			
	-30	-20	-10	cAp
EIA ^a	GTGTATTTTATACCCGGTGAGTTCTCTCAAGAGGCCA	CTCTT		
EIB ^a	AGGGTATATAATGCGCCGTGGGCTAATCTTGGTTA	CATCT		
EII ^{a,b}	TAGTCCTTAAGAGTCAGCGCGCAGTATTTGCTGAAG	AGAG		
EIII ^a	CAGGGTATAACTCACCTGAAAATCAGAGGGCGAGGT	ATTC		
EIV ^a	CCTATATATACTCGCTCTGTACTTGGCCCTTTT	TACACTG		
pIX ^{a,c}	GAATATATAAGGTGGGGGTCTCATGTAGTTTTG	TA TCTGT		
Major late ^{a,d}	GGCTATAAAAGGGGTGGGGGCGCGTTCGTCCTC	ACTCTC		
IVa ₂ ^a	GTGCTGGCCTGGCCTGGACGCGAGCCTTCGTCT	CAGAGTG		

^a Baker and Ziff, submitted for publication.

^b From reference 2.

^c From reference 1.

^d From reference 46.

FIG. 2. Electrophoretic analysis of RNAs transcribed in vitro from cloned Ad2 fragments containing the putative EIA and EIB promoters. The autoradiograms show the gel analyses of RNAs synthesized in vitro from truncated templates containing the putative EIA (A) and EIB (B) promoters. RNAs were synthesized in standard 50- μ l reactions (see text) containing [α -³²P]guanosine triphosphate, uninfected KB cell S100, 150 U of KB RNA polymerase II from uninfected cells, and 0.5 μ g of cleaved plasmid DNA. After incubation for 60 min at 30°C, RNAs were extracted, denatured with glyoxal, and electrophoresed on 4% polyacrylamide gels. For (A), RNAs were synthesized from the EIA template (Fig. 1Bi) cleaved with *Sma*I (lanes 1 and 2) or *Xba*I (lanes 3 and 4); for (B), RNAs were synthesized from the EIB template (Fig. 1Bii) cleaved with *Kpn*I (lanes 1 and 2) or *Hind*III (lanes 3 and 4). For both (A) and (B), the products obtained from standard reactions containing 0.5 μ g of α -amanitin per ml are shown in lanes 2 and 4. The arrows indicate the positions of the run-off transcripts. The numbered lines indicate the size of the transcripts present in lane 1 [(A) and (B)] in each case as determined by comparison with labeled RNA markers (see Fig. 3).

*Bgl*III-cleaved (lanes 5 and 6) and the *Bgl*III- and *Hind*III-cleaved (lanes 3 and 4) templates, on the other hand, produces a single, prominent, α -amanitin-sensitive transcript whose apparent size by comparison with RNA markers (Fig. 3, lane 6) is 550 bases. The sizes of the single, α -amanitin-sensitive transcripts generated by each of the various templates shown in Fig. 4A are consistent with initiation by RNA polymerase II at, or near, the putative EIII promoter. In contrast, the absence in lanes 1 and 3 (*Hind*III-cleaved or *Hind*III- and *Bgl*III-cleaved templates) of an α -amanitin-sensitive transcript intermediate in size between the 550- and 1,000-base transcripts described above suggests that the early EII promoter is not efficiently utilized in the in vitro system. Indeed, we have consistently failed to observe transcription from the early EII promoter with a number of different S100 extracts and various incubation conditions (data not shown). The weaker α -amanitin-resistant bands observed in Fig. 4A that are intermediate in size between the 550- and 1,000-nucleotide transcripts correspond to the so-called end-to-end transcripts described before (21, 41). These bands are not present when RNAs are denatured with glyoxal before electrophoresis (cf. Fig. 3).

Because the EII and EIII promoters are present on the same cloned fragment, it seemed possible that our failure to observe active initiation at the former site is the result of competition with the presumably stronger EIII promoter. To test this possibility, the plasmid bearing the EII and EIII promoters was digested with *Eco*RI, which cleaves at a site (map position 75.9) approximately equidistant between the two promoters and at a single site in the plasmid vector (Fig. 1Biii). The two fragments, each containing a single promoter, were then separated by gel electrophoresis, purified by glass bead adsorption, cleaved with appropriate

FIG. 3. Size analysis of glyoxal-denatured RNAs synthesized *in vitro* from Ad2 promoters. The autoradiogram shows a long polyacrylamide gel analysis of glyoxal-denatured RNAs synthesized *in vitro* in the standard reaction mixture. The templates are: (lane 4) *Sma*I-cleaved EIA plasmid (Fig. 1Bi); (lane 5) *Kpn*I-cleaved EIB plasmid (Fig. 1Bii); (lane 6) *Bgl*III-cleaved EII/III plasmid (Fig. 1Biii); (lane 7) *Sma*I-cleaved EIV plasmid (Fig. 1Biv); (lane 9) *Sma*I-cleaved pIX plasmid (Fig. 1Bv). In lanes 1 to 3, marker RNAs were synthesized with Ad2 late promoter DNA cleaved with *Sma*I (lane 1), *Hind*II (lane 2), or *Hind*III (lane 3) as template (see reference 41). The lengths of the marker RNAs (deduced from sequence information provided by R. Roberts) are indicated in the right margin of the gel. The lengths of the Ad2 runoff RNAs in lanes 4 to 9 (denoted by arrows) were computed by comparison with the marker RNAs, assuming that distance of migration is proportional to the log of molecular weight.

restriction enzymes, and tested for ability to promote specific transcription *in vitro*. The results are shown in Fig. 4B. When the small *Eco*RI fragment (*Eco* RI-a), which contains the EIII promoter, is cleaved with *Bgl*II and tran-

scribed *in vitro*, a prominent α -amanitin-sensitive band is produced (lane 2) which comigrates with the 550-base transcript (lane 1) derived from the *Bgl*II-cleaved plasmid (described above). In addition, when *Eco* RI-a is digested instead with *Hind*III, a prominent α -amanitin-sensitive transcript is produced (lane 4) whose size by comparison with DNA standards is approximately 1,000 nucleotides (data not shown). These results are consistent with predictions based on the restriction map (Fig. 1Biii) and suggest that the EIII promoter is active on the isolated fragment. It is apparent, however, that when the isolated fragment is transcribed, a number of additional bands of both higher and lower molecular weight are produced. These new bands include both α -amanitin-sensitive and -resistant transcripts; since the products shown in Fig. 4B were denatured with glyoxal before electrophoresis, the latter presumably do not correspond to end-to-end products. The reason for this increased background of transcripts is not known. However, it is a commonly observed phenomenon when isolated fragments are transcribed.

In contrast, when the large *Eco*RI fragment (*Eco* RI-b), which contains the putative early EII promoter (approximately 300 nucleotides from the *Eco*RI terminus), is cleaved with *Hind*III and transcribed *in vitro*, no transcripts are observed (lane 6). This result suggests that the failure to observe transcription from the early EII promoter *in vitro* is not due to competition with the EIII promoter.

The activity of the EIV promoter (map position 99.1) was analyzed with a cloned Ad2 DNA fragment which extends from the *Eco*RI site at map position 89.9 to the right end near 100 map units (Fig. 1Biv); this fragment is inserted by GC tailing into the *Pst*I site of pBR322 (kindly provided by J. Nevins). Digestion of this template with *Sma*I, *Hpa*I, or *Hind*III should yield runoff transcripts of 247, 284, and 673 nucleotides, respectively. The results obtained from *in vitro* transcription of this template are shown in Fig. 5A and reveal prominent α -amanitin-sensitive bands of appropriate size; comparison with RNA markers indicates sizes of 248 and 287 nucleotides for the *Sma*I and *Hpa*I runoff transcripts, respectively (Fig. 3, lanes 7 and 8).

Although not apparent in the exposure shown in Fig. 5A, the *Hind*III-cleaved runoff transcript (lane 5), which by comparison with DNA standards is approximately 700 bases long (data not shown), is actually composed of a doublet. We have, in the course of these experiments, also transcribed a clone of Ad2 EIV sequences derived from a transformed cell line (kindly provided by S.-L. Hu). This clone contains, in ad-

FIG. 4. Electrophoretic analysis of RNAs transcribed *in vitro* from cloned Ad2 fragments containing the putative EII and EIII promoters. (A) The autoradiogram shows the gel analysis of RNAs synthesized *in vitro* (see Fig. 2) from truncated templates derived from the EII/EIII plasmid (Fig. 1Biii). The template was cleaved with HindIII (lanes 1 and 2), the combination of HindIII and BglII (lanes 3 and 4), or BglII alone (lanes 5 and 6). After synthesis, transcripts were extracted and analyzed as in the legend to Fig. 2 except that the glyoxalation step was omitted. The products obtained from standard reactions containing 0.5 μ g of α -amanitin per ml are shown in lanes 2, 4, and 6. Arrows indicate the positions in lanes 1 to 6 of the relevant transcripts whose synthesis is sensitive to α -amanitin, and the numbered line indicates the size of the transcript present in lanes 3 and 5 as determined by comparison with labeled RNA markers (see Fig. 3). (B) The autoradiogram shows the gel analysis of RNAs synthesized *in vitro* from BglII-cleaved EII/EIII plasmid (lane 1) or from the gel-purified, isolated EcoRI fragments (see Fig. 1Biii) derived from the EII/EIII plasmid (lanes 2 to 7). Lanes 2 to 5 show the products obtained by *in vitro* transcription of the isolated EcoRI-a fragment cleaved with BglII (lanes 2 and 3) or HindIII (lanes 4 and 5). Lanes 6 and 7 show the products obtained from *in vitro* transcription of the HindIII-cleaved EcoRI-b fragment. The products obtained from standard reactions containing 0.5 μ g of α -amanitin per ml are shown in lanes 3, 5, and 7. Arrows indicate the positions in lanes 1 to 7 of the relevant transcripts whose synthesis is sensitive to α -amanitin.

dition to cellular DNA, Ad2 sequences which extend rightward from the HindIII site at 97.3 map units and contains all but the first two nucleotides of the right-hand terminus of the Ad2 genome. When this template is cleaved with HindIII and transcribed *in vitro*, we observe a single α -amanitin-sensitive transcript which comigrates with the faster-moving component of the doublet described above (data not shown). This result suggests that the slower-moving component does not derive from initiation at the EIV promoter; we have not yet mapped the origin of this latter transcript. In addition, we have observed similar levels of transcription from the EIV promoter using this clone, rather than the clone described above (Fig. 1Biv), where the extent of Ad2 sequences preserved at the right end during cloning is not precisely known. This suggests that the level of initiation observed at the EIV promoter, as demonstrated

in Fig. 5A, is not limited by the loss of sequences upstream (toward the right terminus) from the initiation site.

We have also analyzed the *in vitro* activity of the promoter for the IVa₂ gene. As in the case of the early EII promoter, the IVa₂ gene apparently lacks a readily identifiable TATAA sequence (Baker and Ziff, submitted for publication). And, as in the former case (see above), we have failed to observe any significant initiation of transcription from the IVa₂ promoter (data not shown).

In vitro transcription of an intermediate class gene. We have also examined the *in vitro* activity of the pIX gene, which appears not to be activated during the Ad2 infectious cycle until around the time of DNA replication (30, 31). The activity of the pIX promoter (map position 9.7) was analyzed with a cloned Ad2-HindIII fragment which is inserted into the HindIII site of pBR322 (kindly provided by S.-

FIG. 5. Electrophoretic analysis of RNAs transcribed *in vitro* from cloned Ad2 fragments containing the putative EIV and pIX promoters. The autoradiograms show the gel analyses of RNAs synthesized *in vitro* from truncated templates containing the putative EIV (A) and pIX (B) promoters. RNAs were synthesized, extracted, denatured, and electrophoresed as indicated in the legend to Fig. 2. For (A), RNAs were synthesized from the EIV template (Fig. 1Biv) cleaved with *Sma*I (lanes 1 and 2), *Hpa*I (lanes 3 and 4), or *Hind*III (lanes 5 and 6). The products obtained from reactions containing 0.5 μ g of α -amanitin per ml are shown in lanes 2, 4, and 6. Arrows denote the positions of α -amanitin-sensitive transcripts, and the numbered lines indicate the sizes, in bases, of the transcripts present in lanes 1 and 3. For (B), RNAs were synthesized from either *Sma*I-cleaved (lane 1) or *Hind*II-cleaved (lane 2) Ad2 major late promoter DNA (see reference 41) or from pIX plasmid DNA (Fig. 1Bv) cleaved with *Sma*I (lanes 3 and 4), uncleaved (lane 5), or cleaved with *Hph*I (lanes 6 and 7). The products obtained from reactions containing 0.5 μ g of α -amanitin per ml are shown in lanes 4 and 7, and arrows indicate transcripts whose synthesis is sensitive to α -amanitin. The numbered lines indicate the sizes of the marker major late runoff RNAs, as well as that of the *Sma*I-cleaved pIX runoff transcript as determined in Fig. 3.

L. Hu). This Ad2 fragment extends from map positions 7.9 to 17.0 and, therefore, also contains the major late and IVa₂ promoters at approximately 16.4 map units (Fig. 1Bv). Digestion of this template with *Sma*I should yield a runoff transcript from the pIX promoter of 356 bases. Gel analysis of the products obtained from *in vitro* transcription of the *Sma*I-cleaved template reveals a single, prominent, α -amanitin-sensitive transcript (Fig. 5B, lanes 3 and 4) whose size was determined, in the experiment shown in Fig. 3 (lane 9), to be 356 nucleotides. Using a whole-cell extract, Manley et al. (22) have also recently observed a band of the correct approximate size for initiation at the pIX promoter.

Because it is rather short and does not contain intervening sequences (1), the pIX gene provides an assay for termination. Correct termination of the *in vitro* transcript at the poly(A) addition site (map position 11.1; see reference 1) would yield a primary [not including a possible poly(A) tract] RNA product of 483 nucleotides; this product would be intermediate in size between the two markers shown in lanes 1 and 2 of Fig. 5B. However, transcription either of uncleaved

plasmid (data not shown) or of plasmid linearized with *Hind*III, which simply removes the Ad2 insert and does not cut within the pIX gene, fails to yield any identifiable transcript (Fig. 5B, lane 5). However, when the template is digested with *Hph*I, which cleaves 121 base pairs downstream from the poly(A) addition site (and should, therefore, yield a runoff product of 704 bases), and transcribed *in vitro*, a single prominent transcript results, which is significantly larger than the 533-base marker (Fig. 5, lane 6). Size estimates of this latter transcript, using denatured DNA standards, yield a value of approximately 600 nucleotides (data not shown). This result suggests that transcription continues beyond the poly(A) addition site to the *Hph*I site. We have also failed to observe a transcript whose size is consistent with a terminated product when using intact Ad2 DNA as the template (data not shown). These results suggest that termination (or processing) is not occurring to any appreciable extent at the poly(A) addition site or, given the resolution of the gel system, for several hundred nucleotides beyond.

Comparison of *in vivo* and *in vitro* 5'

ends. To more clearly assess the fidelity of the *in vitro* transcription event, we have directly compared the 5' ends of the transcripts with those of the corresponding *in vivo* RNAs. In each case, this was accomplished by hybridizing RNAs to a short 5'-³²P-labeled DNA probe which is complementary to a region of the RNA roughly 30 to 100 bases downstream from the cap site. The DNA probe is then extended (at the 3' end) to the 5' end of the RNA, using reverse transcriptase. In this way, the extension products obtained for a particular *in vivo* RNA and its corresponding *in vitro* transcript can be directly compared on a high-resolution sequencing gel. This approach has now been used to map the 5' ends of a number of transcripts (10, 32, 33). As discussed below, analysis of the 5' ends of Ad2 RNAs is somewhat complicated by the occurrence of microheterogeneity at the cap site (Baker and Ziff, submitted for publication).

The probes used in the 5'-end analysis of Ad2 RNAs are shown in Fig. 6. The probes vary in size from 18 to 92 base pairs in length, whereas the extended products range from 79 to 127 bases in length. *In vivo* RNAs were isolated from Ad2-infected KB cells maintained for 6 h in the presence of cycloheximide (early) or 24 h in the absence of inhibitor (late), and, in both cases, poly(A)-containing RNA was selected by hybridization to oligodeoxythymidylic acid. *In vitro* transcripts were prepared either as total RNA from the transcription reaction or as gel-purified bands, as described in Materials and Methods.

The primer-extension method was first used to analyze the 5' ends of major late transcripts since extensive fingerprint analysis has already established the accuracy with which this promoter is recognized *in vitro* (41). Hybridization of the 92-base-long *Pvu*II-*Eco*RII primer fragment to either late infected *in vivo* RNA or correctly initiated *in vitro* transcript should result in a 125-nucleotide (not accounting for the cap)-long extended product (Fig. 6i). The results obtained from this analysis are shown in Fig. 7A. When the probe is hybridized to late *in vivo* RNA, an extension product of approximately the correct size is observed (lane 2), and this product is absent when the hybridization is carried out in the absence of RNA (lane 1) or with RNA isolated from uninfected KB cells (lane 3). Moreover, when the probe is hybridized to total RNA prepared from an *in vitro* transcription of *Sma*I-cleaved plasmid containing the major late promoter (see reference 41), a prominent band of precisely the same size is observed (lane 4), and this band is clearly absent when α -amanitin is included in the transcription reaction (lane 5). Because of possible differences in the electro-

FIG. 6. Diagrams of probes used in the 5'-end analysis of *in vivo* and *in vitro* RNAs. The diagrams show the restriction enzyme sites used to generate the probes for the 5'-end analysis and indicate the sizes of the probes (a), as well as the expected sizes of the extended products (b). The locations of cap sites are shown (●). In cases of heterogeneous initiation, only the position of the predominant cap site is shown.

phoretic migration of the marker sequence and the extended products, it is not possible to unambiguously size the latter and thereby assign the nucleotide of initiation. However, it is pos-

FIG. 7. Primer extension reactions for major late and pIX RNAs. The autoradiograms show the gel analyses of primer extension products obtained with probes specific for major late (A) and pIX (B) transcripts. Details of the preparation of the probes and RNAs, as well as of the hybridization, extension reaction, and electrophoretic conditions, are given in the text. For (A), the major late probe was hybridized to *E. coli* transfer RNA (lane 1), poly(A)⁺ late RNA from Ad2-infected KB cells (lane 2), poly(A)⁺ RNA from uninfected KB cells (lane 3), or total RNA from an *in vitro* transcription reaction of Ad2 late promoter DNA cleaved with *Sma*I (see reference 41) carried out in the absence (lane 4) or presence (lane 5) of 0.5 μ g of α -amanitin per ml. For (B), the pIX probe was hybridized to poly(A)-containing late RNA from Ad2-infected cells (lanes 1 and 2) and uninfected cell RNA (lane 3) or to total RNA isolated from an *in vitro* transcription of *Sma*I-cleaved pIX DNA (Fig. 1Bv) carried out in the absence (lane 4) or presence (lane 5) of 0.5 μ g of α -amanitin per ml. Lane 1 is a shorter exposure of the result shown in lane 2. For both (A) and (B), the numbered lines denote sizes determined from marker sequence reactions run in adjacent lanes (not shown); only those portions of the autoradiograms which contain relevant information are shown.

sible to state that the extended products are within one or two nucleotides of the predicted size (and, hence, are not likely to be prematurely terminated) and that *in vivo* and *in vitro* RNAs yield the same result. This suggests that initiation *in vitro* occurs at the same nucleotide used *in vivo*, as previously established in this case by direct sequence analysis of the 5' terminus (41).

It should be noted that there is a background of identical extension products common to lanes 4 and 5 (Fig. 7). These presumably reflect initiation events not inhibited by α -amanitin, and we believe them to be the result of random initiation by RNA polymerase III. The latter transcripts

may constitute the α -amanitin-resistant radioactivity which does not enter the usual analytical gel. Several other relevant points need to be made. (i) The large amount of radioactivity present at the bottom of each lane in Fig. 7A corresponds to nonextended primer. (ii) The weakness of the extension product obtained with the late *in vivo* RNA is undoubtedly explained by the fact that only eight bases of the probe actually overlap the first leader sequence of the major late transcript, the remainder falling within nonconserved intervening sequence. (iii) The diffuse band which migrates more slowly than the extension product and which is present in several of the lanes in Fig. 7A (but absent, for example, in lane 2) is an artifact resulting from incomplete denaturation of the probe before electrophoresis. (iv) Shorter exposures of Fig. 7A show that the extension product present in lane 4 is actually composed of a lower predominant band and a minor band immediately above it. Indeed, careful examination of the extension product obtained with *in vivo* RNA reveals a similar pattern. We believe this doublet effect to be explained by incomplete extension at the actual cap. The extent to which this anomaly occurs may be related to the form (I or II) of the cap, which may also differ between *in vitro* and *in vivo* RNAs. Baker and Ziff (submitted for publication) have also seen an additional band(s) relative to the number predicted by fingerprint analysis when examining the 5' ends of *in vivo* mRNA's by reverse transcriptase-primer extension.

The 5' termini of the *in vitro* pIX transcripts were analyzed by using a probe (Fig. 6ii) which should be extended from a length of 33 nucleotides to a length of either 90 or 92 bases (not accounting for the cap) in length, assuming that initiation occurs with the A or G as shown in Table 1. The results obtained after extending the probe hybridized to late *in vivo* RNA are shown in Fig. 7B, lane 2, and in a shorter exposure in lane 1. The extension products, which are not present when the hybridization is carried out with uninfected cell RNA (lane 3), comprise four closely spaced bands (although only three bands may be visible in the photograph); the lowest band is the predominant one, followed by the highest band, and the two intervening bands are of intermediate intensity. The lowest of the four bands appears to be approximately 91 to 92 nucleotides in length, which is consistent with the expected value for initiation with the A. These results are consistent with heterogeneous initiation at the pIX promoter. However, as indicated in Table 1, Baker and Ziff (submitted for publication) find evidence for only two initiation sites, with the A preferred over the G, two

nucleotides upstream (3:1). Our results suggest either that there are two additional initiation sites, presumably adjacent to the A and G, or, more likely, that the four bands are really two doublets resulting from incomplete extension at the cap. In any case, when the probe is hybridized to *in vitro* RNA, an identical pattern of extension products is observed (lane 4); these products are not present when α -amanitin is included in the transcription reaction (lane 5), although, again, there is a background of extension products common to lanes 4 and 5. We conclude that, although there may be some slight differences in the relative frequencies with which the initiation sites are used *in vivo* and *in vitro*, in both cases initiation is heterogeneous and the same initiation sites are used.

The accuracy of the 5' termini of the early EIA, EIB, EIII, and EIV RNAs have been analyzed in a similar fashion, except that, to reduce the background derived from the *in vitro* reactions, gel-purified transcripts have been used rather than total RNA. The control in each case has come from a mock gel purification of transcription products synthesized in the presence of α -amanitin. In the case of each early promoter, we observe extension products which are (i) not produced with uninfected cell RNA or α -amanitin controls, (ii) are of the approximate size expected, and (iii) are identical when produced from either early *in vivo* RNA or specific *in vitro* transcripts. For example, as shown in Fig. 8A, the EIA probe (Fig. 6iii) yields a single extension product of the appropriate size (127 to 128 nucleotides, not accounting for the cap) with both early *in vivo* RNA (lane 1) and the gel-purified *Sma*I-cleaved EIA runoff transcript (lane 3). Similarly, the EIB probe (Fig. 6ii) also yields identical extension products of appropriate size (120 to 121 nucleotides, not accounting for the cap) with both *in vivo* (Fig. 8B, lane 1) and *in vitro* (lane 3) RNAs, although in this case the extension products appear as closely spaced doublets. We again suspect that this is the result of incomplete extension at the cap, but we cannot rule out heterogeneous initiation at neighboring nucleotides. Baker and Ziff (submitted for publication) have found that initiation *in vivo* is essentially limited to a single site at the EIA and EIB promoters (Table 1). (Although they have evidence for an additional minor EIB cap site *in vivo*, this minor site accounts for only 5% of the transcripts.)

In the case of both the EIA and EIB probes, there are additional bands of higher molecular weight in the extension reactions derived from *in vitro* RNAs even when gel-purified transcripts are used. However, these additional bands, whose presence is variable, are simultaneously

present in the α -amanitin controls and are, therefore, not related to specific transcription by RNA polymerase II.

As in the case of the pIX promoter, analysis of EIII and EIV RNAs provides a clear indication of heterogeneous initiation. The EIII probe (Fig. 6v) yields a similar pattern of extension products when hybridized with either early *in vivo* RNA (Fig. 8C, lanes 1 and 2) or a gel-purified, *Bgl*II-cleaved EIII runoff transcript (lane 4), but not when hybridized with uninfected cell RNA (lane 3) or the α -amanitin control (lane 5). The pattern consists of four bands whose relative intensities vary somewhat with each analysis. For example, in the experiment shown in lane 1, hybridization to *in vivo* RNA yields a doublet and a more slowly moving single band, all of roughly equal intensity, separated by a band of reduced intensity; this latter band may be difficult to see in the photograph. The doublet is of the approximate size expected (78 to 79 nucleotides, not accounting for the cap) for initiation with the A (Table 1). In the *in vivo* analysis shown in lane 2, on the other hand, the pattern is reversed, with a smaller, more intense band separated from a more slowly moving doublet by a single faint band. Hybridization of the probe to *in vitro* transcript (lane 4) produces an extension pattern identical to that shown in lane 1. We interpret these results to be indicative of initiation at two sites and assume that the four bands result from incomplete extension at the cap in each case. However, we cannot rule out the possibility of additional initiations. Baker and Ziff (submitted for publication) find that initiation *in vivo* occurs at two sites in the region of the EIII promoter (Table 1). Primary initiation occurs with the A, and secondary initiation occurs with a G located two nucleotides upstream.

The large amount of radioactivity present at the bottom of each lane in Fig. 8C corresponds to unextended primer. The background ladder of bands which emanate up from the primer is not dependent on the addition of RNA and is, therefore, assumed to be the result of end addition reactions. These bands unfortunately present an unavoidable background when the size of the primer is close to that of the extension product.

The EIV probe (Fig. 6vi) yields a complex pattern of six or seven extension products after hybridization to either early *in vivo* RNA (Fig. 8D, lanes 1 and 2) or gel-purified *Sma*I-cleaved EIV runoff transcript (lane 4). A typical result with *in vivo* RNA is shown in lane 1: a prominent band whose size is close to the expected 86 nucleotides (not accounting for the cap) for initiation with the A (Table 1) is then followed by

FIG. 8. Primer extension reactions for EIA, EIB, EIII, and EIV RNAs. The autoradiograms show the gel analyses of primer extension products obtained with probes specific for EIA (A), EIB (B), EIII (C), and EIV (D) transcripts. For (A) and (B), the EIA and EIB probes were hybridized to poly(A)-containing early RNA from Ad2-infected KB cells (lane 1) and uninfected-cell RNA (lane 2) or to gel-purified transcripts isolated from 10 $\times$ in vitro reactions carried out in either the absence (lane 3) or the presence (lane 4) of 0.5 μ g of α -amanitin per ml. For these reactions the EIA template was the *Sma*I-cleaved plasmid (Fig. 1Bi), and the EIB template was the *Kpn*I-cleaved plasmid (Fig. 1Bii). For (C) and (D), the EIII and EIV probes were hybridized to poly(A)-containing early RNA from Ad2-infected cells (lanes 1 and 2) and uninfected cell RNA (lane 3) or to gel-purified transcripts isolated from 10 $\times$ in vitro reactions carried out in the absence (lane 4) or presence (lane 5) of 0.5 μ g of α -amanitin per ml. The templates used for these reactions are the *Bgl*II-cleaved EII/EIII plasmid (Fig. 1Biii) and the *Sma*I-cleaved EIV plasmid (Fig. 1Biv). The results obtained with late

five consecutive bands of higher molecular weight; the relative intensities of these five bands are greatest for the largest and smallest. This result suggests the possibility of multiple initiation events. Indeed, Baker and Ziff (submitted for publication) have provided evidence for heterogeneous EIV initiation in vivo (Table 1); their results suggest that initiation occurs primarily with the A, but also secondarily with the five Ts immediately preceding the A. Moreover, the pattern of preference among the Ts is very similar to that suggested by the relative intensities of the bands in lane 1. Similar patterns derived from another experiment are shown in lane 2 for the early in vivo RNA and in lane 4 for the corresponding in vitro transcript. Two relatively prominent bands, visible in lanes 2 and 4, that are immediately below the principal set of extension products are not indicative of additional initiation sites. They are artifactual and derive from the fact that there are two overlapping recognition sites for the restriction enzyme *Fnu*4H-I, an enzyme essential for the generation of the EIV primer. Consequently, there are two forms of the primer: a longer form is responsible for producing the predominant pattern seen in Fig. 8D, whereas a minor, shorter form produces an overlapping pattern of which the lower part is visible as the two anomalous bands described above.

The efficiency with which the Ad2 genes are transcribed in vitro varies. A typical time course for in vitro transcription from the various Ad2 promoters is shown in Fig. 9. An examination of this time course and a comparison of the transcripts shown in Fig. 3, all of which were synthesized under optimal conditions, indicate that the levels to which the various Ad2 transcripts accumulate in vitro vary considerably. When the amount of radioactivity incorporated into each band over a 2-h period is corrected for the size and labeling nucleotide (G) content of the transcript, the apparent extent of accumulation varies over a 10- to 15-fold range (Table 2); the greatest accumulation is seen for the major late transcript, whereas the EIA transcript is produced at only 10% of the level of that of the major late transcript. The accumulation of the EIB and EIII transcripts is about 40 to 50% that of the major late transcript, whereas the EIV and pIX transcripts are produced at a level that is only 25 to 30% that of the major late transcript. The rates of accumulation of the

Ad2 RNA that are shown in lanes 1 and 2 are from separate experiments. For both (A) and (B), the numbered lines denote sizes determined from marker sequence reactions run in adjacent lanes (not shown).

FIG. 9. Kinetics of specific transcription from Ad2 promoters. Transcription reactions were carried out as described (see Fig. 2) for the indicated times, using the following templates: *Sma*I-cleaved EIA, EIV, pIX, and major late (ML) plasmids; *Kpn*I-cleaved EIB plasmid; *Bgl*II-cleaved EII/EIII plasmid. After electrophoresis of the RNAs, specific transcripts were localized from the autoradiograms, excised, and counted by liquid scintillation. Similarly sized adjacent areas from above or below the specific transcripts were excised, counted, and subtracted as background. The specific activity of the [α -³²P]guanosine triphosphate used in this experiment was approximately 10,000 cpm/pmol.

various Ad2 transcripts show characteristic (and reproducible) profiles (Fig. 9). For example, the initial rate of accumulation of the EIA transcript is similar to those of the other early promoters, but then plateaus much more rapidly. In addition, we have consistently observed a lag in the rates of accumulation of the pIX and major late transcripts, but not in the case of the early promoter transcripts. Similarly, Luse and Roeder (21) have observed an initial lag in the accumulation of major late transcript relative to that of mouse β -globin. During the linear phase after the lag, the major late transcript accumulates at a rate significantly greater than that of any of the other transcripts.

We have, during the course of these experiments, also examined the in vitro transcription of these promoters, using whole Ad2 DNA which

had been cleaved with restriction enzymes to generate appropriate runoff transcripts. For example, digestion of Ad2 DNA with *Sma*I simultaneously generates appropriately sized runoff transcripts from the major late, pIX, and EIV promoters. The results obtained from transcribing *Sma*I-cleaved Ad2 DNA are presented in Fig. 10: lanes 1, 2, and 3 show for comparison the major late, pIX, and EIV transcripts, respectively, obtained with *Sma*I-cleaved plasmids as described above. Lanes 4 to 7 show the transcripts obtained with *Sma*I-cleaved Ad2 DNA over a concentration range in which the molarities of the promoters are similar to those used in the plasmid transcriptions. It is apparent that over the concentration range used (which includes the optimal DNA concentrations), there is much less difference in the extent to which the major late, pIX, and EIV transcripts accumulate than is observed when cloned fragments are transcribed. This difference appears to be largely the result of a reduction in the strength of the major late signal. In addition, transcription from the EIV promoter appears to show a lower concentration optimum than that of the pIX and major late promoters; at a DNA concentration of 50 μ g/ml (lane 6), transcription from the EIV promoter has essentially ceased, whereas the pIX and major late activities are still close to optimal. The intense band at the bottom of lane 7 had been identified as VA₁ RNA, which is transcribed from the VA₁ gene at about map position 30 by the RNA polymerase III present in the S100 extract. Termination signals are recognized by RNA polymerase III in the in vitro transcription reactions; hence, VA₁ RNA synthesis is observed even when uncleaved Ad2 DNA is used as the template. In vitro transcription of VA₁ RNA shows a significantly higher DNA concentration optimum than

TABLE 2. Relative levels of transcription of Ad2 promoters

Promoter	Radioactivity level (cpm) ^a	Transcript size (bases)	% G	Relative level (%)
Major late	24,000	533	33	100
EIB	6,120	347	29	45
EIII	9,500	570	27	45
EIV	2,640	247	30	26
EIA	1,800	510	28	9
pIX	3,700	356	27	28

^a The levels of radioactivity incorporated into each transcript are determined from the 2-h time points in Fig. 9 and are corrected for the size and G content of the transcripts.

FIG. 10. Transcription of whole Ad2 DNA cleaved with *Sma*I. The autoradiogram shows the gel analysis of RNAs synthesized *in vitro* from *Sma*I-cleaved whole Ad2 DNA (lanes 4 to 7). RNAs were synthesized in standard transcription reactions at DNA concentrations of 10 (lane 4), 20 (lane 5), 50 (lane 6), and 100 (lane 7) μ g/ml and extracted, denatured, and electrophoresed as described in the legend to Fig. 2. Major late, pIX, and EIV transcripts, all from *Sma*I-cleaved plasmids (see Fig. 1B), are shown in lanes 1 to 7 for comparison.

that of any of the polymerase II promoters and, in fact, does not reach optimal levels until the DNA concentration exceeds 100 to 150 μ g/ml (data not shown).

It should be pointed out that *Sma*I cleavage of Ad2 DNA also yields an EIA runoff transcript which comigrates with the major late RNA shown in Fig. 10. If Ad2 DNA is instead digested with the combination of *Sma*I and *Hind*III, the major late transcript is reduced in size to 197 nucleotides, and it is, therefore, possible to simultaneously examine runoff transcripts from the EIA, pIX, EIV, and major late promoters (data not shown). It is apparent from this analysis that the band identified in lanes 4 to 7 as major late transcript is, in fact, primarily major late transcript and contains only small amounts of EIA RNA. It is also possible to examine transcription from the EIB and EIIB promoters by cleaving Ad2 DNA with either *Kpn*I or *Bgl*II, respectively (data not shown). Transcription from these promoters shows a DNA concentration optimum similar to that of the pIX and

major late promoters described above.

We have compared, assaying both cloned fragments and whole Ad2 DNA, the relative efficiencies with which the various promoters are transcribed in S100 extracts made from either uninfected KB cells or cells infected for various periods of time with Ad2. We have thus far failed to observe any clear difference, attributable to the use of infected-cell extracts, in the relative levels at which these promoters are transcribed (data not shown).

DISCUSSION

In this report we have demonstrated that human RNA polymerase II will, in the presence of a crude soluble extract (S100) from uninfected human KB cells, transcribe a variety of Ad2 genes, including those of the EIA, EIB, EIIB, and EIV regions, as well as the region encoding pIX. We have shown by a reverse transcriptase-primer extension method that initiation of transcription *in vitro* is accurate even for weak transcription units and that, in every case, the 5' ends of *in vivo* and *in vitro* RNAs are identical. In addition, preliminary fingerprint analysis has further confirmed that the 5' ends of *in vivo* and *in vitro* RNAs are identical and that, in every case, the *in vitro* transcripts are capped (data not shown). Moreover, we have found that in those instances where initiation appears to be heterogeneous *in vivo* (EIIB, EIV, and pIX genes), there exists a similar heterogeneity in the *in vitro* transcripts. These results suggest that the heterogeneity observed with the *in vivo* RNAs is not the result of processing variability before capping, but rather is an inherent part of the initiation process that is faithfully duplicated in the *in vitro* system. Of course, it is possible in the case of both the *in vivo* and *in vitro* transcripts that there is an initiation event upstream from the cap site followed by rapid processing and capping such that only those sequences downstream from the cap site are conserved. That the *in vitro* system is apparently devoid of processing activity and that we never observe discrete RNA polymerase II-dependent transcripts larger than the stable transcripts characterized, even with very short transcription times (data not shown), argue against this possibility. However, a rigorous identification of initiation sites will ultimately require the demonstration of a conserved nucleoside di- or triphosphate residue on the 5' termini. Finally, there is also the formal possibility with the reverse transcriptase-primer extension method that the bands observed are prematurely terminated and do not represent complete extensions. However, that the extension products are, in

each case, of the approximate size expected argues against this possibility.

An important implication of our results is the confirmation of the presence of functional class II RNA polymerase promoters at the *in vivo* cap sites of the EIA, EIB, EIII, EIV, and pIX genes as predicted by nascent chain analysis (13), ultraviolet transcription mapping (4, 42), and 5,6-dichloro-1- β -D-ribofuranosyl-benzimidazole inhibition of transcription (38). In contrast, we have failed to observe detectable transcription of the early EII and IVa₂ genes, despite evidence from *in vivo* studies that these genes contain functional promoters (4, 13). Since both genes lack a recognizable TATAA sequence upstream from the cap site, and since Corden et al. (10) have demonstrated its importance with respect to *in vitro* transcription, it is tempting to correlate the lack of transcription *in vitro* with the lack of a TATAA box. Since the early EII gene, in fact, becomes one of the more active of the early transcription units *in vivo* (27), this suggests either that the optimal conditions have not yet been devised for expression of promoters which lack the TATAA sequence or that different initiation mechanisms may operate at promoters which do, and do not, have TATAA boxes. It should be noted that we have also failed to observe initiation from the early EII promoter with S100's and whole-cell extracts (22) prepared from Ad2-infected cells (data not shown). At later times of infection, EII mRNA's are also initiated from an additional site around map position 72 (9). We have not yet examined this region for *in vitro* promoter activity.

We have observed that the apparent efficiencies with which the Ad2 genes are transcribed *in vitro* vary considerably. The levels of activity attained *in vitro* may correlate roughly with maximal rates of transcription observed *in vivo* in that, in both cases, the EIA and EIII are the weakest and strongest early transcription units, respectively (27). However, since the rate of accumulation of a transcript depends on both initiation and elongation events, differences in levels of accumulation cannot necessarily be interpreted as differences in promoter strength. Indeed, differences in the kinetics of transcription of the various promoters *in vitro* suggest that the process is complex. True assessments of promoter strengths will ultimately require, for example, the application of appropriate competition and abortive initiation assays with purified initiation factors. In addition, the results obtained from transcription of whole Ad2 DNA raise questions regarding the significance of accumulation levels observed with cloned fragments and suggest that the efficiency of pro-

motor utilization may be affected by the particular DNA background present. Finally, we have recently transcribed a plasmid containing the EIA and major late promoters of Ad3 and have observed much less difference in their levels of transcription than is the case with Ad2 (data not shown).

We have not observed any extract-dependent regulation when transcribing Ad2 genes *in vitro*. For example, the pIX gene, which is not expressed *in vivo* until 4 to 6 h after infection, is transcribed in an extract prepared from uninfected cells. (The major late promoter is also active in uninfected cell extracts; however, as indicated above, recent evidence indicates that it is one of the earliest to be expressed during an infection *in vivo*.) In addition, with the exceptions of the early EII and IVa₂ promoters, all of the early genes tested are active in uninfected cell extracts, and in no case do we observe any significant difference when Ad2-infected cell extracts are used. The relative levels of transcription of the various genes do not appear to change in any clear manner when transcribed in either S100's or whole-cell extracts (22) prepared at various times after Ad2 infection. These results suggest that, with the possible exceptions of the early EII and IVa₂ genes, viral products are not absolutely required for initiation at early and intermediate promoters on purified DNA templates and that, if regulation does occur at the level of initiation in any of these cases (as opposed to elongation or processing), it may be controlled through the accessibility of promoters. A lack of positive control with respect to *in vitro* initiation has now been seen in a number of instances (for a discussion, see reference 21); a major future goal will involve attempts to superimpose regulation on the *in vitro* transcription systems. A possible caveat with respect to the above interpretations, however, is that the efficiency of the *in vitro* transcription systems is still quite low (0.01 to 0.2 transcript per template per h). Consequently, it is possible that we are measuring a background level of initiation and that factors which would modulate or enhance this process are not present (or active) in sufficient titers in the infected-cell extracts.

As described above, transcription of the EIV region *in vivo* declines during the late phase of infection. Nevins and Winkler (28) have suggested that this decline is the result of repression by the 72,000-dalton DNA binding protein, a product of the EII region. We have, using the whole Ad2 DNA assay, occasionally observed what appears to be a specific reduction in the extent of EIV transcription relative to the opti-

mal levels shown in Fig. 10 (data not shown). However, this effect depends on the extract used, and it has been observed with a variety of both uninfected and Ad2-infected cell extracts. We do not believe it to be the result of regulation *per se* of the EIV promoter and feel instead that it is related to the DNA concentration phenomenon shown in Fig. 10. Thus, transcription of the EIV gene is particularly sensitive to changes in the concentration ratios of DNA and initiation factors. These observations point to the danger of possible artifacts in attempting to reconstitute regulation in the *in vitro* transcription systems. For example, whole-cell extracts (22) made from late in the course of viral infection may contain significant amounts of viral DNA. The presence of this viral DNA might then cause a specific reduction in the strength of the EIV signal in a manner that is unrelated to the possible presence of putative repressors in the extract. In this regard, it may be worth noting that we have thus far failed to observe any specific effect on EIV transcription *in vitro* of adding to the reaction partially purified, 72,000-dalton DNA binding protein (a gift from M. Green; data not shown). Finally, the observation that EIV transcription is particularly sensitive to DNA concentration suggests the naive possibility that the decline in activity of this region late in infection is a gene dosage phenomenon related to the replication of viral DNA. However, given the vast excess of cellular DNA present, it is difficult to envision how such a mechanism might operate.

We have failed to observe termination (or processing) at the poly(A) addition site of the pIX gene. A possible explanation for this finding is that termination (or processing) factors are not present in sufficient quantities in our extracts or that termination does not occur at the poly(A) addition site *in vivo*. Recent observations with the mouse β^{maj} -globin (J. Darnell, personal communication), chicken ovalbumin (P. Chambon, personal communication), and several Ad2 transcription units (14, 15, 26) indicate that, in each case, transcription continues for 1 to 2 kilobases beyond the poly(A) addition site. Whether termination is then a discrete or random event is not known. We are presently attempting to determine whether, in the case of the pIX gene, transcription terminates at a discrete site beyond the poly(A) addition site.

ACKNOWLEDGMENTS

We thank J. Nevins, M. Wilson, J. Darnell, S.-L. Hu, and G. Chinnadurai for gifts of plasmids (see Results) containing cloned Ad2 fragments and N. Proudfoot for introducing us to the primer-extension technique. In addition, we are particularly grateful to C. Baker and E. Ziff for helpful discussions and for providing, prior to publication, essential sequence information on the cap sites of Ad2 mRNA's. We also thank

S. Sandmeyer and E. Slattery for critical reading of the manuscript.

These studies were supported by research grants from the Public Health Service National Cancer Institute (CA16640 and CA23615) and the American Cancer Society (NP-284) to R.G.R. and by Public Health Service institute support grant CA16217 to Washington University. D.C.L. is a Helen Hay Whitney Fellow. R.G.R. is a Camille and Henry Dreyfus Teacher-Scholar Awardee.

LITERATURE CITED

1. Alestrom, P., G. Akusjaryi, M. Perricaudet, M. Mathews, D. F. Klessing, and U. Pettersson. 1980. The gene for polypeptide IX of adenovirus type 2 and its unspliced messenger RNA. *Cell* 19:671-681.
2. Baker, C. C., J. Herisse, G. Courtois, F. Galibert, and E. Ziff. 1979. Messenger RNA for the Ad2 DNA binding protein: DNA sequences encoding the first leader and heterogeneity at the mRNA 5' end. *Cell* 18:569-580.
3. Berk, A. J., F. Lee, T. Harrison, J. Williams, and P. A. Sharp. 1979. Pre-early adenovirus 5 gene product regulates synthesis of early viral messenger RNAs. *Cell* 17:935-944.
4. Berk, A. J., and P. A. Sharp. 1977. Ultraviolet mapping of the adenovirus 2 early promoters. *Cell* 12:45-55.
5. Blanton, R. A., and T. H. Carter. 1979. Autoregulation of adenovirus type 5 early gene expression. III. Transcription studies in isolated nuclei. *J. Virol.* 29:458-465.
6. Bogenhagen, D. F., S. Sakonju, and D. D. Brown. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription. II. The 3' border of the region. *Cell* 19:27-35.
7. Carter, T. H., and R. A. Blanton. 1978. Possible role of the 72,000-dalton DNA-binding protein in regulation of adenovirus type 5 early gene expression. *J. Virol.* 25:664-674.
8. Carter, T. H., and R. A. Blanton. 1978. Autoregulation of adenovirus type 5 early gene expression. II. Effect of temperature-sensitive early mutations on virus RNA accumulation. *J. Virol.* 28:450-456.
9. Chow, L. T., T. R. Broker, and J. B. Lewis. 1979. Complex splicing patterns of RNAs from the early regions of adenovirus-2. *J. Mol. Biol.* 134:265-303.
10. Corden, J., B. Wasyluk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon. 1980. Promoter sequences of eukaryotic protein-coding genes. *Science* 209:1406-1414.
11. Curtis, P., A. C. Finnigan, and G. Rovera. 1980. The β major and β minor globin nuclear transcripts of Friend erythroleukemia cells induced to differentiate in culture. *J. Biol. Chem.* 255:8971-8974.
12. Engelke, D. R., S.-Y. Ng, B. S. Shastri, and R. G. Roeder. 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* 19:717-728.
13. Evans, R. M., N. Fraser, E. Ziff, J. Weber, M. Wilson, and J. E. Darnell. 1977. The initiation sites for RNA transcription in Ad 2 DNA. *Cell* 12:733-739.
14. Fraser, N. W., and M.-T. Hsu. 1980. Mapping of the 3' terminus of the large late Ad-2 transcript by electron microscopy. *Virology* 103:514-516.
15. 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.
16. Gegenheimer, P., and D. Apirion. 1978. Processing of rRNA by RNase P: spacer tRNAs are linked to 16S rRNA in an RNase P RNase III mutant strain of *E. coli*. *Cell* 15:527-539.
17. Green, M. R., and R. G. Roeder. 1980. Definition of a novel promoter for the major adenovirus-associated virus mRNA. *Cell* 22:231-242.
18. Jones, N., and T. Shenk. 1979. Isolation of adenovirus

- type 5 host range deletion mutants defective for trans-formation of rat embryo cells. *Cell* 17:683-689.
19. Jones, N., and T. Shenk. 1979. An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proc. Natl. Acad. Sci. U.S.A.* 76:3665-3669.
 20. Lewis, J. B., and M. B. Mathews. 1980. Control of adenovirus early gene expression: a class of immediate early products. *Cell* 21:303-313.
 21. Luse, D. S., and R. G. Roeder. 1980. Accurate transcription initiation on a purified mouse β -globin DNA fragment in a cell-free system. *Cell* 20:691-699.
 22. Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gelfer. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. U.S.A.* 77:3855-3859.
 23. Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder. 1981. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.* 255:11992-11996.
 24. McKnight, G. S., D. C. Lee, and R. D. Palmiter. 1980. Transferrin gene expression: regulation of mRNA transcription in chick liver by steroid hormones and iron deficiency. *J. Biol. Chem.* 255:148-153.
 25. McKnight, G. S., and R. D. Palmiter. 1979. Transcriptional regulation of the ovalbumin and conalbumin genes by steroid hormones in chick oviduct. *J. Biol. Chem.* 254:9050-9058.
 26. Nevins, J. R., J.-M. Blanchard, and J. E. Darnell. 1980. Transcription units of adenovirus type 2: termination of transcription beyond the poly(A) addition site in early regions 2 and 4. *J. Mol. Biol.* 144:377-386.
 27. Nevins, J. R., H. S. Ginsberg, J.-M. Blanchard, M. C. Wilson, and J. E. Darnell. 1979. Regulation of the primary expression of the early adenovirus transcription units. *J. Virol.* 32:727-732.
 28. Nevins, J. R., and J. J. Winkler. 1980. Regulation of early adenovirus transcription: a protein product of early region 2 specifically represses region 4 transcription. *Proc. Natl. Acad. Sci. U.S.A.* 77:1893-1897.
 29. Nguyen-Huu, M. C., K. J. Barrett, K. Giesecke, T. Wurtz, A. E. Sippel, and G. Schutz. 1978. Transcriptional regulation of chicken egg white genes during secondary stimulation with estrogen. *Hoppe-Seyler's Z. Physiol. Chem.* 359:1307-1313.
 30. Persson, H., B. Oberg, and L. Philipson. 1978. Purification and characterization of an early protein (E14K) from adenovirus type 2-infected cells. *J. Virol.* 28:119-139.
 31. Persson, H., U. Pettersson, and M. B. Mathews. 1978. Synthesis of a structural adenovirus polypeptide in the absence of viral DNA replication. *Virology* 90:67-79.
 32. Proudfoot, N. J., M. H. M. Shander, J. Manley, M. L. Gelfer, and T. Maniatis. 1980. Structure and *in vitro* transcription of human globin genes. *Science* 209:1329-1342.
 33. Reddy, V. B., P. K. Ghosh, P. Lebowitz, M. Piatak, and S. M. Weissman. 1979. Simian virus 40 early mRNAs. I. Genomic localization and 3' and 5' termini and two major splices in mRNA from transformed and lytically infected cells. *J. Virol.* 30:279-296.
 34. Ringold, G. M., K. R. Yamamoto, J. M. Bishop, and H. E. Varmus. 1977. Glucocorticoid-stimulated accumulation of mouse mammary tumour viral RNA: increased rate of synthesis of viral RNA. *Proc. Natl. Acad. Sci. U.S.A.* 74:2879-2883.
 35. Rio, D., A. Robbins, R. Myers, and R. Tjian. 1980. Regulation of simian virus 40 early transcription *in vitro* by a purified tumor antigen. *Proc. Natl. Acad. Sci. U.S.A.* 77:5706-5710.
 36. Sakonju, S., D. F. Bogenhagen, and D. D. Brown. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: 1. The 5' border of the region. *Cell* 19:13-25.
 37. Segall, J., T. Matsui, and R. G. Roeder. 1981. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.* 255:11986-11991.
 38. Sehgal, P. B., N. W. Fraser, and J. E. Darnell. 1979. Early Ad-2 transcription units: only promoter-proximal RNA continues to be made in the presence of DRB. *Virology* 94:185-191.
 39. Shaw, A. R., and E. B. Ziff. 1980. Transcripts from the adenovirus-2 major late promoter yield a single early family of 3' coterminal mRNAs and five late families. *Cell* 22:905-916.
 40. Waslylyk, B., C. Keding, J. Corden, O. Brison, and P. Chambon. 1980. Specific *in vitro* initiation of transcription on conalbumin and ovalbumin genes and comparison with adenovirus-2 early and late genes. *Nature (London)* 285:367-373.
 41. Weil, P. A., D. S. Luse, J. Segall, and R. G. Roeder. 1980. Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18:469-484.
 42. Wilson, M. C., N. W. Fraser, and J. E. Darnell. 1979. Mapping of RNA initiation sites by high doses of uv irradiation: evidence for three independent promoters within the left 11% of the Ad-2 genome. *Virology* 94:175-184.
 43. Wu, G.-J. 1978. Adenovirus DNA-directed transcription of 5.5S RNA *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* 75:2175-2179.
 44. Zieve, G., and S. Penman. 1976. Small RNA species of the HeLa cell: metabolism and subcellular localization. *Cell* 8:19-31.
 45. Ziff, E. 1980. Transcription and RNA processing by the DNA tumour viruses. *Nature (London)* 287:491-499.
 46. Ziff, E. B., and R. M. Evans. 1978. Coincidence of the promoter and capped 5' terminus of RNA from the adenovirus 2 major late transcription unit. *Cell* 15:1463-1475.