Delineation of DNA Sequences That Are Important for In Vitro Transcription from the Adenovirus EIIa Late Promoter

DER-HWA HUANG[†] AND ROBERT G. ROEDER*

Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, New York 10021

Received 30 November 1987/Accepted 2 February 1988

Late in infection, transcription of the EIIa gene is initiated primarily at map unit 72 of the adenovirus genome. A cell-free nuclear extract system was used to determine sequence elements important for the function of this late promoter. In such a system, the transcriptional activity of a circular template was found to be much higher (5- to 10-fold) than that of a linear template. The effect of template topology appeared to be dependent on two distal upstream elements with 5' boundaries located near -265 to -223 and -147 to -133 (in relation to the initiation site), since deletions of these regions reduced transcription of the circular template, in a stepwise fashion, to a level similar to that observed with the linear template. Further deletions revealed an element in the -116 region that appeared to be more important for transcription of the circular template (10-fold reduction) than for transcription of the linear template (3-fold reduction). Lastly, deletion of the TACAAA sequence in the -29 region resulted in further reduction in transcription, indicating that this element functions as a promoter in vitro.

The adenovirus EIIa gene encodes a 72-kilodalton (kDa) DNA-binding protein which has multiple functions in viral nucleic acid metabolism, including DNA replication (6), transcriptional regulation (14, 30), and mRNA stability (1). Early in infection, transcription of this gene is under the control of a promoter (EIIa early) at map unit 75 of the adenovirus genome (Fig. 1). After the onset of viral replication, another promoter (EIIa late) at map unit 72 is activated and presumably accounts for most of the 72-kDa DNA-binding protein synthesized late in infection (7). Although it is not clear how these two promoters are differentially regulated during viral infection, it is possible that such regulation reflects structural and functional differences between these two promoters (or associated elements) and the factors which recognize them.

It has been demonstrated that transcription of the early promoter is positively regulated by E1a proteins encoded by the 13S RNA (3, 16, 22, 29) and possibly the 12S RNA (21) species. In contrast, the late promoter appears to be refractory to the 13S RNA-encoded E1a but repressed by the 12S RNA-encoded product (13, 32). Extensive studies of the structure and function of the EIIa early promoter in vivo and in vitro have demonstrated two proximal (pseudo-TATA) elements and two upstream elements (36, 41, 42), one of which binds an E1a-inducible factor. However, little is known about the structure of the late promoter and the mechanism involved in its activation late in infection. The initiation site of the late promoter is preceded by a TACAAA sequence, which might be expected from its position and sequence similarity to the TATA box consensus (5, 34) to be a major element of the late promoter. Nevertheless, there is as yet no direct evidence to support this notion, and the presence of a C residue on the very highly conserved third position (5) of the presumptive TATA sequence raises questions about the significance of this element.

In an effort to determine both basal and regulatory elements of the late promoter, we used a nuclear extract system to study transcription of this promoter. In this report, we demonstrate that template topology has a marked effect on the in vitro transcriptional activity of this promoter, and we further identify multiple elements important for promoter function. As this manuscript was submitted, Bhat et al. (4) reported a detailed in vivo analysis of a series of linker-scanning mutations constructed in this promoter. Information relevant for the in vitro study will be discussed.

MATERIALS AND METHODS

Enzymes and reagents. All restriction endonucleases, S1 nuclease, T4 ligase, Klenow fragment of *Escherichia coli* DNA polymerase I, and calf thymus topoisomerase I were purchased from Bethesda Research Laboratories. T4 polynucleotide kinase was purchased from P. L. Biochemicals. [γ -³²P]ATP was purchased from New England Nuclear Corp.

Plasmids and construction of deletion mutations. Plasmids pF2 (a gift of C. Goldenberg) and pEII-1 were pBR322 recombinants that contained, respectively, the *EcoRI-HindIII* fragment extending from map units 71 to 73.2 and the *BamHI-EcoRI* fragment extending from map units 59.5 to 76 of adenovirus type 2 DNA (Fig. 1B).

Construction of restriction site deletion mutations in the upstream region of the EIIaL promoter was done with *HindIII*, *PvuII*, *ApaI*, *SstI*, *BalI*, *DdeI*, and *RsaI* restriction endonuclease sites (see Fig. 3A). Appropriate fragments were isolated from pF2 and cloned into the polylinker region of pUC13. The 5' ends generated by *HindIII* or *SstI* digestion were cloned into corresponding restriction sites. The 5' ends generated by other restriction enzymes were cloned into the *SmaI* site (*ApaI* and *DdeI* sites were repaired by T4 DNA polymerase and *E. coli* Klenow enzyme, respectively). The 3' ends of all the DNA fragments were derived from the *EcoRI* site at +330.

For construction of Bal 31 deletion mutations, the pUC13 recombinant that contained the *SstI-EcoRI* fragment of the

* Corresponding author.

[†] Present address: Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, Bethesda, MD 20892.

FIG. 1. Organization of transcription units of the adenovirus genome. (A) Entire genome of adenovirus, divided into 100 map units from left to right. The transcription initiation sites of the EIIa early and late genes are located at map units 75 and 72, respectively. The direction of transcription is indicated by arrows. (B) Simplified restriction maps of inserts present in pEII-1 and pF2. Plasmid pEII-1 is a pBR322 recombinant that contains a 4-kb *Bam*HI-*Eco*RI DNA fragment extending from map units 59.5 to 76 of the adenovirus genome. Plasmid pF2 is a pBR322 recombinant that contains a 0.7-kb *Eco*RI-*Hind*III DNA fragment extending from map units 71 to 73.2 of the adenovirus genome. Transcription initiation sites and direction of transcription of both EIIa early and late genes are indicated by the L-shaped arrows above the horizontal line. The S1 probes were the DNA fragment extending from +180 to -285 of the EIIa early gene and the DNA fragment extending from +68 to -412 of the EIIa late gene. The numbers in parentheses indicate the distance (in kilobases for pEII-1, in nucleotides for pF2) between restriction sites and the initiation sites of the EIIa late gene.

EIIaL promoter was first treated with *Sst*I. Following *Bal* 31 exonuclease digestion for various lengths of time, the reaction was stopped and the plasmid was digested with *Eco*RI and *Bgl*II to completion. This mixture was then ligated with a pUC13 vector that had been cleaved by *Eco*RI and *Sma*I. The 5' endpoint of each deletion mutant was determined by the method of Maxam and Gilbert (25). Each mutant was named according to the position of the endpoint in relation to the major initiation site determined by Baker and Ziff (2).

DNA templates. The circular templates referred in the text were supercoiled DNAs prepared by double-banding in an ethidium bromide-cesium chloride gradient. The linear templates were cleaved by restriction endonucleases prior to incubation in the transcription reaction mix. Supercoiled DNA was relaxed with calf thymus topoisomerase I at 37°C in the presence of 50 mM Tris hydrochloride (pH 7.5)-50 mM KCl-10 mM MgCl₂-0.5 mM dithiothreitol-0.1 mM EDTA-30 µg of bovine serum albumin per ml. Supercoiled DNA was not detected in ethidium bromide-stained agarose gels after the treatment of topoisomerase I.

RNA analysis and quantitation. The protocol used for in vitro transcription was similar to that described by Dignam et al. (9). A standard 50-µl reaction mix contained 20 µg of DNA template per ml and 50% (vol/vol) nuclear extract (about 8 mg of protein per ml). Reactions were terminated after 60 min of incubation at 30°C by addition of an equal

volume of stop solution (0.1 M sodium acetate [pH 5.5], 1% sodium dodecyl sulfate, 0.5 mg of yeast RNA per ml). Proteinase K was added to a final concentration of 1 mg/ml, and the mix was incubated for 45 min at 37°C. Following phenol-chloroform extraction and ethanol precipitation, the pellet was dried and suspended in a buffer containing 10 mM Tris hydrochloride (pH 7.9), 1 mM EDTA, and 5 mM MgCl₂. RNase-free DNase I (4.5 U), prepared by the Macaloid absorption method (31), was added and incubated for 15 min at 37°C. RNAs were again extracted and precipitated. Purified RNAs were hybridized in a 20-µl volume with an excess amount of kinase-treated probe (*Kpn*I-*Hind*III fragment from +68 to -418 of the EIIaL promoter) at 45°C for 14 h in the presence of 10 mM PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid), pH 6.4]-0.4 M NaCl-80% formamide. The hybridization mixture was diluted 10-fold with a solution containing 30 mM sodium acetate (pH 4.5), 0.4 M NaCl, and 1 mM ZnSO₄, followed by the addition of 7.5 µg of denatured calf thymus DNA and 300 U of S1 nuclease. After 1 h of incubation at 30°C, RNA was extracted with phenol-chloroform and precipitated by ethanol. Final products were analyzed on a polyacrylamide sequencing gel (25). Appropriate bands in autoradiograms were quantified by densitometric scanning with a Beckman Du8 spectrophotometer. Autoradiograms of different exposures were scanned to ensure accuracy.

RESULTS

Accurate transcription of the EIIaL promoter in vitro.

Transcription of the EIIaL promoter was initially observed in an RNA polymerase II-supplemented cytosolic extract by using a run-off assay with linear template (23). With a similar type of assay, a transcript of the correct size was barely detectable in nuclear extracts (data not shown). The transcriptional efficiency of this promoter appeared to be lower than that of all other adenovirus promoters tested and less than 1% that of the major late promoter. The weakness of this promoter and the lack of a previous demonstration of precise initiation at the nucleotide level necessitated a more sensitive assay for analysis of the in vitro transcripts. As shown in Fig. 2 and in subsequent figures, two major bands were readily detected by S1 nuclease mapping of the RNA generated from circular templates in nuclear extracts. A very similar result was obtained with polyadenylated RNA isolated from infected cells, and a detailed comparison of in vivo and in vitro results (D.-H. Huang, Ph.D. thesis, Washington University, St. Louis, Mo., 1986) indicated that these two major bands correspond to transcription initiation at the two adenine residues (+1 and +3) previously established by fingerprint analysis as the 5' ends of EIIaL mRNA (2). Thus, we conclude that the EIIaL gene is accurately transcribed in vitro.

Transcription of circular versus linear templates. In the course of studying upstream deletion mutations of the EIIaE promoter (Huang et al., manuscript in preparation), we observed that the overall transcription of the EIIaL promoter from supercoiled templates was usually 5- to 10-fold higher than that from linear templates prepared by *Hind*III cleavage (at -418; Fig. 1B). It seemed possible that the topology of the template might have a direct effect on the transcription activity of the EIIaL promoter. Alternatively, the presence of free ends, especially near the promoter, might have had an adverse effect on promoter activity. To examine these possibilities, a more extensive analysis of the activity of the EIIaL promoter in various linear templates was carried out. The pEII-1 plasmid, which contained both EIIaE and EIIaL promoters, was cleaved by restriction endonucleases at several different locations. As shown in Fig. 1B, the cleavage sites analyzed ranged from 1.42 kilobases (kb) upstream (*Eco*RI site) to 2.9 kb downstream (*Bam*HI site) of the initiation site of the EIIaL promoter. Transcription from the EIIaE and EIIaL promoters in these linear templates was compared with transcription from the uncleaved supercoiled template. As shown in Fig. 2A, transcription of the EIIaL promoter from linear templates (lanes 3 to 6) was significantly lower than that from a supercoiled template (lane 2), regardless of the location of the cleavage site. In contrast, transcription of the EIIaE promoter remained constant.

These results clearly indicated that the difference in the transcription activity between supercoiled and linear templates was related to the topology of the template. Nonetheless, there was still a remote possibility that the presence of the stronger EIIaE promoter might have somehow blocked the accessibility of a limiting transcription factor to the EIIaL promoter. Therefore, we also examined transcription of the EIIaL promoter in a plasmid (pF2) which lacked the EIIaE promoter (Fig. 1B). As shown in Fig. 2A (lanes 7 to 10), transcription of the EIIaL promoter was still markedly lower from the linear template than from the circular template. Thus, we conclude that transcription of the EIIaL promoter is indeed more active in the supercoiled template.

FIG. 2. In vitro transcription analysis of circular versus linear templates. (A) In vitro transcripts produced from circular and linear templates of pEII-1 (lanes 2 to 6) or pF2 (lanes 7 to 10) were analyzed by S1 mapping. For lanes 2 to 6, both the EIIaE and EIIaL probes were used. For lanes 7 to 10, only the EIIaL probe was used. The restriction enzymes used to linearize templates are indicated at the top of the gel. The location of these restriction sites in relation to the initiation site of the EIIaL promoter is shown in Fig. 1B. The supercoiled templates are indicated by S. Bands corresponding to the EIIaE or EIIaL transcripts are indicated accordingly. (B) In vitro transcripts produced from supercoiled (S), linear (L), and relaxed (R) pF2 templates were analyzed. The linear template was *Eco*RI-cleaved pF2 DNA. The relaxed template was pF2 DNA treated with calf thymus topoisomerase I. The molecular weight markers (M) were end-labeled *Hpa*II fragments of pBR322.

It has been widely observed that supercoiled DNA templates are rapidly relaxed under the usual transcription conditions. Thus, it seemed possible that the high level of transcription activity observed for the supercoiled template might not be dependent on the template superhelicity. To test this possibility, transcription was analyzed with a template that had been relaxed by calf thymus topoisomerase I. As shown in Fig. 2B, the relaxed template (lane 4) was transcribed as efficiently as the supercoiled template (lane 2). This observation strongly suggested that the circular template is sufficient to support active transcription. For this

reason, the uncleaved plasmid DNA will be referred to as the circular template in later discussion. However, given that a maximum of 5% of the major late promoter-containing templates are involved in transcription under the same conditions, and given the 10- to 20-fold-lower template efficiency of the EIIaL promoter (relative to major late), we estimated that no more than 0.5% of the EIIaL circular templates were actually engaged in transcription. Thus, we cannot completely rule out the possibility that a small fraction of supercoiled DNA was still present after treatment with topoisomerase I or that a small fraction of the relaxed template was subject to gyration in the nuclear extract (16).

Transcription analysis of restriction site deletion mutants. Since an effect of template topology was observed for the EIIaL promoter but not for the EIIaE promoter, it seemed probable that certain gene-specific elements might be present in the 5'-flanking region of the EIIaL promoter. Such elements could be either positive elements that are recognized only in the circular template or negative elements that are recognized only in the linear template. To distinguish between these two possibilities and to localize relevant DNA sequences, we analyzed several deletion mutations constructed in the upstream region of the EIIaL promoter. As shown in Fig. 3, the deletion of upstream sequences from -416 to -265 had only a slight effect (less than 25% reduction; for quantitation, see Fig. 3 legend) on transcription of circular templates. However, further deletions caused more drastic effects on transcription from these templates. For example, the further deletion of sequences (from -265) to -223, to -133, and to -34 reduced transcription to 33, 13, and 1%, respectively, of the wild-type level (the signal of the -34 mutant was only visible in the autoradiogram). Further deletion from -34 to -28 reduced transcription to an undetectable level. In the case of linear templates, the deletion of sequences from -416 to -133 had no deleterious effect on transcription, whereas the deletion of sequences from -133 to -34 reduced transcription to an undetectable level.

It is relevant to note from these results that (i) transcription of the circular template was reduced to a level similar to that observed with the linear template when sequences from -265 to -133 were deleted; (ii) transcription of the circular template was reduced more drastically (10-fold reduction) than transcription of the linear template (3-fold reduction) when sequences from -133 to -34 were deleted; and (iii) further deletion of sequences from -34 to -28 reduced transcription of both circular and linear templates to undetectable levels. These results strongly suggest the presence of at least two functionally distinct domains in the 5'-flanking region of the EIIaL promoter. As deletions of these domains resulted in reduction of transcription, it is clear that they must be positive control elements. One domain whose boundary lay between -265 and -133 was required only for transcription of the circular template, while another domain whose 5' boundary lay between -133 and -28 was important for transcription of both circular and linear templates.

Transcriptional analysis of Bal 31 deletion mutants. To determine whether more than one element was present in each region, we analyzed a series of Bal 31 deletion mutants constructed from the -223 mutant (Fig. 4A). As shown in Fig. 5A, the deletion of sequences from -223 to -147 had no significant effect on transcription of the circular template. However, further deletion of sequences to -133 reduced transcription more than twofold, and this level was unchanged by further deletion to -116 (for quantitation, see Fig. 5 legend; see also data for -223 versus -123 and -116

FIG. 3. Transcriptional analysis of restriction site deletion mutants of the EIIaL promoter. (A) Restriction sites used for construction of upstream deletion mutants of the EIIaL gene. (B) Autoradiograph showing S1 analysis of in vitro transcripts produced from both circular and linear templates of upstream deletion mutants. The deletion mutation endpoints are indicated at the top of the gel. Bands corresponding to the EIIaL-specific transcripts are indicated by arrows. When transcription was analyzed with circular templates, the levels observed for the -306, -265, -223, and -133 mutants were 78, 93, 46, and 19%, respectively, of the wild-type (-418) level. The corresponding values from the averages of three experiments were 78, 77, 33, and 13% and, because of the observed variability, provide a better indication of the effects of sequence deletion. Lane M, size markers.

mutants in Fig. 5B). Thus, in conjunction with data from the preceding section, the present results indicate that the DNA sequence located between -265 and -133 contains two distinct elements which are responsible for active transcription of the circular template. The 5' boundary of one element

FIG. 4. Construction of Bal 31 deletion mutants and DNA sequence of upstream region of the EIIaL promoter. (A) Deletion mutant pEIIaL/-223 was linearized at the unique *Sst*I site and digested with Bal 31 for various lengths of time. Following complete digestion with *Eco*RI and *Bgl*II, the shortened DNA fragments were cloned into the *Eco*RI-*Sma*I sites of pUC13. (B) DNA sequences from -280 to the initiation site are shown. The DNA sequences which are discussed in the text are underlined.

is located between -265 and -223, and the 5' boundary of the other is located between -147 and -133.

The deletion of sequences from -116 to -103 caused a drastic reduction (about 10-fold) of transcription for the circular template (Fig. 5A and B). In contrast, it caused a much smaller (maximally threefold) reduction in transcription of the linear template (Fig. 6). Further deletion from -103 to -34 did not cause any additional reduction in transcription for either circular (Fig. 5B) or linear (Fig. 6) templates. Thus, we conclude that only one element, whose 5' boundary is located between -116 and -103, can be detected in this region. Furthermore, this element appeared to be relatively more important for transcription of the circular template.

The earlier analysis suggested that DNA sequences located between -34 and -28 were important for transcription of both circular and linear templates. In a further analysis of mutants located in this region (Fig. 5B), we found that the deletion of sequences from -34 to -29 produced no deleterious effect on transcription of the circular template, while deletion of one more nucleotide (-28 mutant) reduced transcription to an undetectable level. Since the first T residue of the TACAAA sequence is present at -29, this observation suggested that this TATA-like sequence was important for transcription. In the case of the linear template, deletion of the TATA-like sequence reduced but did not totally eliminate transcription (Fig. 6). However, the apparent low level of signal observed after deletion of the TATA-like sequence might not have been significant, since it was also present in the -12 mutant and might have been an artifact produced by S1 nuclease (see Fig. 6 legend for further explanation). Therefore, the function of this TA

CAAA sequence appears to be similar to that of the TATA box.

DISCUSSION

Accurate in vitro initiation. Mathis et al. (23) first reported that run-off transcripts from the EIIaL promoter could be detected in an S-100 extract, although definitive evidence for accurate initiation was not provided. Using the S1 nuclease mapping technique, we have demonstrated that the EIIaL promoter can be faithfully transcribed in a nuclear extract prepared from uninfected cells. It has been reported previously that all the other adenovirus genes can also be accurately initiated in cell-free extracts (S-100, whole cell, or nuclear extracts) derived from uninfected cells (9, 11, 18, 28). Along with these previous studies, the present analysis indicates that all of the factors which are required for basal transcription (recognition of basic promoter elements) of adenovirus genes are present in host cells prior to viral infection. This is in fact consistent with the observation that lytic infection can still occur in the absence of known viral regulatory molecules, such as the Ela protein (35). As summarized below, we have used the cell-free system to delineate sequence elements that are important for accurate initiation in vitro. Although our study may be limited by the extractability of certain proteins necessary for transcription and by the possibility that certain sequence elements involved in cooperative interactions may not be revealed by a series of progressive 5' deletion mutations, the physiological relevance of most elements identified here has been substantiated by a recent in vivo study with a series of linker-scanning mutants (4). Thus, the in vitro systems should provide a valuable approach for the further analysis of host factors involved in transcription of adenovirus genes.

Upstream elements that mediate topology-dependent transcription. A surprising and novel finding of this report is that a circular EIIa late template was more actively transcribed than a linear template in vitro. Analysis of upstream deletion mutations enabled us to identify two distal upstream elements (Fig. 7) which enhanced transcription of circular but not linear templates. Although similar studies of other promoters have shown upstream elements whose effects are, under some conditions, more pronounced in circular than in linear templates (15, 27, 33, 37), the present study appears to provide the first example of upstream elements which uniquely enhance transcription from circular templates.

The 5' boundaries of the two distal upstream elements were localized to regions between -265 and -223 and between -147 and -133. Interestingly, both regions contain an inverted CCAAT sequence (one between -233 and -229, the other between -139 and -135). The fact that CCAAT sequences represent the essential part of other small distal control elements (26) suggests that this may also be true for the distal elements described here. However, there are additional sequences with interesting features in both regions. For example, a remarkable G+C-rich sequence, GGGCTGCGGCGGGG, is present between -252 and -233, and the sequence spanning the CCAAT sequence in the second distal region contains two elements of dyad symmetry (TTGCAA and TTGGCCAA). Thus, a further analysis of more defined mutations is necessary to confirm that the CCAAT sequences are indeed involved in the topology-dependent transcription.

Whatever the sequences important in these two distal regions, it should be noted that the effect of template topology is contingent on specific sequences which are

FIG. 5. Transcriptional analysis of circular templates of Bal 31 deletion mutants. (A) Autoradiograph showing S1 analysis of transcripts produced from mutants with endpoints located between -223 and -103 . In the experiment shown, the signal from the -147 mutant was 2.5-fold greater than that from the -133 mutant, and in other experiments this value varied from 1.5- to 2.5-fold depending on the frequency of freezing and thawing of the extract. (B) Autoradiograph showing S1 analysis of transcripts produced from mutants with endpoints located downstream of -123 . The spliced parts on the left were derived from the same gel under the same autoradiographic exposure. The mutant endpoints are indicated at the top of the gel. The linear template is indicated by L. Bands corresponding to the EIIaL-specific transcripts are indicated by arrows. Lane M, Size markers.

distant from the basic promoter element (see below). It is thus very likely that these sequences represent binding sites for factors other than the common factors involved in transcription of the minimal promoters of most class II genes (24, 33; N. Nakajima, M. Horikoshi, and R. G. Roeder, manuscript in preparation). Given that the functions of these elements were maintained when a relaxed circular template was transcribed, it is reasonable to believe that functional interactions between the putative factors and these sequences do in fact require circularity in the template. However, because of the low fractional usage of input templates for transcription and the possibility of gyration of relaxed templates in the extracts (17), the possible involvement of supercoiled templates is not eliminated. It is possible that the

circularity (or superhelicity) of the template may be necessary not only for a protein-DNA interaction per se, but also for the functional consequences of this interaction (perturbation of DNA structure or subsequent interactions with other promoter factors) to be manifest. Obviously, the presumptive protein-DNA interaction might also occur but have no functional consequence in a linear template. Further identification and characterization of putative factors are necessary to determine how the sequence elements function.

Elements required for transcription of both circular and linear templates. The present analysis also revealed a proximal upstream element located near the -116 region (Fig. 7). Unlike the distal elements, this element appeared to be important for transcription of both circular and linear tem-

FIG. 6. Transcriptional analysis of linear templates of Bal 31 deletion mutants. Transcripts produced from mutants with endpoints located downstream of -133 were analyzed. The circular template is indicated by C. Bands corresponding to the EIIaL-specific transcripts are indicated by arrows. Bands marked with asterisks are S1 artifacts which were observed occasionally in other experiments. There are several other nonspecific bands present both below and above the EIIaL-specific bands. The 5' ends of these generally corresponded to either A or T residues, indicative of potential S1 artifacts. Note that the intensities of these bands gradually increased as the deletions approached the cap site, suggesting a general increase of S1 artifacts. Since transcription of the EIIaL gene is initiated at two A residues, it is possible that the residual weak signals present in corresponding positions in the -28, -20, -17, and -12 mutants are also S1 artifacts.

plates and thus probably involves a factor distinct from the factor(s) which recognizes the distal upstream element. Since this element exerted a greater effect in the circular template than in the linear template, the corresponding factor presumably functions more efficiently with the circular template. A close examination of the DNA sequence in this region reveals the presence of a TGACGCA sequence which is highly homologous to the consensus sequence (TGACTCA) for the AP-1 factor (20) and to a related sequence element present in transcription control regions of the adenovirus E1a, EIIaE, EIII, and EIV genes (19).

FIG. 7. Organization of the EIIa late promoter. The initiation site and the direction of transcription are marked by the L-shaped arrow. The upstream elements are indicated by boxes, with the locations of their 5' boundaries shown above.

Interestingly, deletion of the EIIa early upstream element also resulted in a more drastic reduction of transcription from circular templates (Huang et al., manuscript in preparation). It is also relevant to note that transcription from the linear EIIa late template was greater than that from the circular template when the proximal upstream element was deleted. Although the basis for this is not clear, it could be related to the presence of artificial "entry sites" at the ends of the linear template, as suggested previously for the sea urchin histone H2A gene (12). A similar observation was made for an adenovirus IVa2 promoter template that lacked an upstream (between -49 and -38) element (28). It will be interesting to determine whether a common mechanism is involved in this type of reaction.

The present study also demonstrated a low level of accurate initiation in a deletion mutant that retained only 29 nucleotides of upstream sequence. Since the sequence and location of the TACAAA sequence between -29 and -24 are very similar to those of the TATA box found in most class II genes, this observation suggests strongly that the TACAAA element functions as the basic promoter element of the EIIaL gene. It also suggests that the weak activity of the EIIaL promoter (<1% of the major late promoter) is due to a weak basic promoter element. It has been demonstrated previously that either a transition or a transversion at the third nucleotide of the TATA box of the chicken conalbumin gene (39, 40) reduced transcription drastically. Therefore, the presence of the C residue at the third nucleotide of the TACAAA sequence may be the direct cause of the weak promoter strength. This notion has been proven by the recent finding that conversion of the C residue into a T residue markedly stimulates transcription of the EIIaL gene (Huang et al., manuscript in preparation).

Comparison of in vivo and in vitro studies. Bhat et al. (4) recently reported that three regions are important for the maximal in vivo expression of the EIIaL promoter. The common elements detected in both in vivo and our in vitro studies include the TATA-like element and elements located at the -116 and the -147 regions. However, there are three major differences between these two studies. First, the element located at the -265 region was not detected by Bhat et al. This can probably be attributed to the fact that the wild-type construct used in their study only contains the DNA sequence up to -223 (*SacI* site). Second, a functional SP1 site (between -50 and -40) was not detected in nuclear extracts. This probably reflects the presence of low levels of SP1 in our nuclear extracts, since little transcription was detected from the simian virus 40 early promoter (D.-H. Huang, unpublished observation). Lastly, the importance of an inverted CCAAT sequence (at the -80 region) was also not detected in vitro. However, as pointed out earlier, two inverted CCAAT sequences located further upstream may be functional in vitro. It is thus somewhat surprising that the in vitro system was unable to identify this downstream

CCAAT element. Nevertheless, if a cooperative interaction with other upstream elements were necessary for the proper function of this downstream CCAAT element, then it would be expected to be manifest with linker-scanning mutants but not with the progressive 5' deletions used here. Alternatively, since multiple CCAAT-binding factors may exist (10), different factors may be involved in the recognition of these sequences. It is possible that the factor that recognizes the downstream element is again underrepresented in the nuclear extract.

In vivo significance of the topological effect. One of the most intriguing observations of this report is the demonstration of multiple upstream elements which are dependent on a certain template topology for optimal function. This raises the possibility that this effect is relevant to the activation of this promoter late in infection. It has been found in superinfection experiments that the addition of DNA synthesis inhibitors after replication of the initial virus restricts expression of certain late genes to the initial virus, with no detectable expression from the superinfecting virus (8, 38). It has been suggested that changes in the physical structure of the DNA templates during replication may be crucial for expression of the intermediate and late genes. The nature of these changes is unknown but could include a topological change which is mimicked by the circular plasmid in the cell-free system and which facilitates activation of the E1aL promoter. Further examination of the structure, during transcription, of the circular plasmid that mediates the topology-dependent transcription and the physical states of the adenovirus template late in infection should provide more useful information on this question.

ACKNOWLEDGMENTS

We thank our colleagues for their encouragement and discussion during this work, especially W. Jack for his valuable comments on this manuscript.

This work was supported by Public Health Service grants CA 42567 and CA 34223 from the National Institutes of Health and by general support from the Pew Trusts to the Rockefeller University.

LITERATURE CITED

- Babich, A., and J. R. Nevins. 1981. The stability of early adenovirus mRNA is controlled by the viral 72 Kd DNA binding protein. *Cell* **26**:371-379.
- Baker, C. C., and E. B. Ziff. 1981. Promoters and heterogeneous 5' termini of the messenger RNAs of adenovirus serotype 2. *J. Mol. Biol.* **149**:189-221.
- 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.
- Bhat, G., L. SivaRaman, S. Murthy, P. Domer, and B. Thimmappaya. 1987. In vivo identification of multiple promoter domains of adenovirus E1a-late promoter. *EMBO J.* **6**:2045-2052.
- Breathnach, R., and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for protein. *Annu. Rev. Biochem.* **50**:349-383.
- Challberg, M. D., and T. J. Kelly. 1982. Eukaryotic DNA replication: viral and plasmid model systems. *Annu. Rev. Biochem.* **51**:901-934.
- 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.
- Crossland, L. D., and H. J. Raskas. 1983. Identification of adenovirus genes that require template replication for expression. *J. Virol.* **46**:737-748.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475-1489.
- Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis. 1987. A multiplicity of CCAAT box-binding proteins. *Cell* **50**:863-872.
- Fire, A., C. C. Baker, J. L. Manley, E. B. Ziff, and P. A. Sharp. 1981. In vitro transcription of adenovirus. *J. Virol.* **40**:703-719.
- Grosschedl, R., and M. L. Birnstiel. 1982. Delimitation of far upstream sequences required for maximal *in vitro* transcription of an H2A histone gene. *Proc. Natl. Acad. Sci. USA* **79**:297-301.
- Guilfoyle, R. A., W. P. Osherooff, and M. Rossini. 1985. Two functions encoded by adenovirus early region IA are responsible for the activation or repression of the DNA-binding protein gene. *EMBO J.* **4**:707-713.
- Handa, H., R. E. Kingston, and P. A. Sharp. 1983. Inhibition of adenovirus early region IV transcription *in vitro* by a purified viral DNA binding protein. *Nature (London)* **307**:545-547.
- Hen, R., P. Sassone-Corsi, J. Corden, M. P. Gaub, and P. Chambon. 1982. Sequences upstream from the TATA box are required in vivo and in vitro for efficient transcription from the adenovirus serotype 2 major late promoter. *Proc. Natl. Acad. Sci. USA* **79**:7132-7136.
- Jones, N., and T. Shenk. 1979. An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proc. Natl. Acad. Sci. USA* **76**:3665-3669.
- Kmiec, E. B., and A. Worcel. 1985. The positive transcription factor of the 5S RNA induces a 5S DNA-specific gyration in *Xenopus* oocyte extracts. *Cell* **41**:945-953.
- Lee, D. C., and R. G. Roeder. 1981. Transcription of adenovirus type 2 genes in a cell-free system: apparent heterogeneity of initiation at some promoters. *Mol. Cell. Biol.* **1**:635-651.
- Lee, K. A. W., and M. R. Green. 1987. A cellular transcription factor E4F1 interacts with an E1a-inducible enhancer and mediates constitutive enhancer function *in vitro*. *EMBO J.* **6**:1345-1353.
- Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* **49**:741-752.
- Leff, T., R. Elkaim, C. R. Goding, P. Jalino, R. Sassone-Corsi, M. Perrecaudet, C. Keding, and P. Chambon. 1984. Individual products of the adenovirus 12S and 13S E1a mRNAs stimulate viral E1a and E1b expression at the transcriptional level. *Proc. Natl. Acad. Sci. USA* **81**:4381-4385.
- Lillie, J. W., M. Green, and M. R. Green. 1986. An adenovirus E1a protein region required for transformation and transcriptional repression. *Cell* **46**:1043-1051.
- Mathis, D. J., R. Elkaim, C. Keding, P. Sassone-Corsi, and P. Chambon. 1981. Specific *in vitro* initiation of transcription of the adenovirus type 2 early and late transcription units. *Proc. Natl. Acad. Sci. USA* **78**:7383-7387.
- Matsui, T., J. Segall, A. P. Weil, and R. G. Roeder. 1980. Multiple factors required for accurate initiation of transcription by RNA polymerase II. *J. Biol. Chem.* **255**:11992-11996.
- Maxam, A., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560-564.
- McKnight, S., and R. Tjian. 1986. Transcriptional selectivity of viral genes in mammalian cells. *Cell* **46**:795-805.
- Mizushima-Sugano, J., and R. G. Roeder. 1986. Cell type-specific transcription of an immunoglobulin k light chain gene *in vitro*. *Proc. Natl. Acad. Sci. USA* **83**:8511-8515.
- Natarajan, V., M. J. Madden, and N. P. Salzman. 1984. Proximal and distal domains that control *in vitro* transcription of the adenovirus IVa2 gene. *Proc. Natl. Acad. Sci. USA* **81**:6290-6294.
- Nevins, J. R. 1981. Mechanism of activation of early viral transcription by the adenovirus E1a gene product. *Cell* **26**:213-220.
- Nevins, J. R., and J. Jensen-Winkler. 1980. Regulation of early adenovirus transcription: a protein product of early region 2 specifically represses region 4 transcription. *Proc. Natl. Acad. Sci. USA* **77**:1893-1897.
- Picard, D., and W. Schaffner. 1983. Correct transcription of a cloned mouse immunoglobulin gene in vivo. *Proc. Natl. Acad. Sci. USA* **80**:1893-1897.

- Sci. USA **80**:417-421.
32. **Rossini, M.** 1983. The role of adenovirus early region 1A in the regulation of early regions 2A and 1B expression. *Virology* **131**:49-58.
33. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43**:165-175.
34. **Shenk, T.** 1981. Transcriptional control regions: nucleotide sequence requirements for initiation by RNA polymerase II and III. *Curr. Top. Microbiol. Immunol.* **93**:25-46.
35. **Shenk, T., N. Jones, N. W. Colby, and D. Fowlkes.** 1979. Functional analysis of adenovirus 5 host-range deletion mutants defective for transformation of rat embryo cells. *Cold Spring Harbor Symp. Quant. Biol.* **44**:367-375.
36. **SivaRaman, L., and B. Thimmappaya.** 1987. Two promoter-specific host factors interact with adjacent sequences in an E1A-inducible adenovirus promoter. *Proc. Natl. Acad. Sci. USA* **84**:6112-6116.
37. **Sive, H., N. Heintz, and R. G. Roeder.** 1986. Multiple sequence elements are required for maximal *in vitro* transcription of a human histone H2B gene. *Mol. Cell. Biol.* **6**:3329-3340.
38. **Thomas, G. P., and M. B. Mathews.** 1980. DNA replication and the early to late transition in adenovirus infection. *Cell* **22**:523-533.
39. **Wasylyk, B., and P. Chambon.** 1981. A T to A base substitution and small deletions in the conalbumin TATA box drastically decrease specific *in vitro* transcription. *Nucleic Acids Res.* **9**:1813-1824.
40. **Wasylyk, B., A. Derbyshire, A. Gay, D. Molko, A. Roget, R. Teoule, and P. Chambon.** 1980. Specific *in vitro* transcription of conalbumin gene is drastically decreased by single-point mutation in TATA box homology sequences. *Proc. Natl. Acad. Sci. USA* **77**:7024-7028.
41. **Yee, A. S., R. Reichel, I. Kovesdi, and J. R. Nevins.** 1987. Promoter interaction of the E1A-inducible factor E2F and its potential role in the formation of a multi-component complex. *EMBO J.* **6**:2061-2068.
42. **Zajchowski, D. A., H. Boeuf, and C. Keding.** 1985. The adenovirus-2 early E1a transcription unit possesses two overlapping promoters with different sequence requirements for E1a-dependent stimulation. *EMBO J.* **4**:1293-1300.