Volume 15 Number 20 1987 Nucleic Acids Research

Upstream regulatory regions required to stabilize binding to the TATA sequence in an adenovirus early promoter

Joseph Garcia, Foon Wu and Richard Gaynor

Division of Hematology-Oncology, 11-242 Louis Factor Building, Department of Medicine, UCLA School of Medicine and Jonnson Comprehensive Cancer Center, Los Angeles, CA 90024, USA

Received June 24, 1987; Revised and Accepted September 1, 1987

ABSTRACT

Of the five early adenovirus promoters, the early region 3 (E3) promoter is one of the most strongly induced by the E1A protein. To identify cellular proteins involved in both the basal and E1A-induced transcriptional regulation of the E3 promoter, DNase I footprinting using partially purified Hela cell extracts was performed. Four regions of the E3 promoter serve as binding domains for cellular proteins. These regions are found between -156 to -179 (site IV), -83 to -103 (site III), -47 to -67 (site II), and -16 to -37 (site I), relative to the start of transcription. Examination of the DNA sequences in each binding domain suggests that site III likely serves as a binding site for activator protein 1 (AP-1), site II for the cyclic AMP regulatory element binding protein (CREB), and site I for a TATA binding factor. The factors binding to either site II or III were sufficient to stabilize binding to the TATA sequence (site I). Mutagenesis studies indicated that both sites II and III, in addition to site I, are needed for complete basal and E1A-induced transcription. These results suggest that multiple cellular factors are involved in both the basal and EIA-induced transcriptional regulation of the E3 promoter, and that either of two upstream regions are capable of stabilizing factor binding to the TATA sequence.

INTRODUCTION

Upstream promoter regulatory sequences are important in the genetic regulation of eucaryotic promoters transcribed by RNA polymerase II. Mutagenesis studies have defined a number of these *cis*-acting sequences, such as the TATA sequence present 25-30 bp upstream of the RNA initiation site, which are important in the transcriptional control of both viral and cellular promoters and serve as DNA-binding domains for cellular transcription factors. Several DNA binding proteins important in promoter regulation have been purified to homogeneity, including SP1, CAAT binding factor, nuclear factor I, and the upstream factor (USF) (1-4). The addition of these purified proteins to *in vitro* transcription systems stimulate transcription of specific genes (1-3). Thus, a study of *cis*-acting regulatory sequences and the proteins which bind to them is important in understanding the regulation of a variety of genes.

Cellular transcription factors appear critical for the induction of both viral and cellular genes. This may occur due to the binding of different transcription factors that are not present in the non-induced state, or by potential alterations in the binding activity of factors that are normally present. The induction of cellular genes such as c-*fos*, β-interferon, or metallothionein may occur in the presence of agents such as serum stimulation, glucocorticoids, poly(I)-poly(C), or phorbol esters (5-7). The induction of many viral genes occurs in the presence of viral proteins known as *trans*-activators (8-16). The mechanism by which viral *trans*-activators induce transcription is unknown. One of the best studied of these viral *trans*-activator proteins is the 289 amino acid (289AA) adenovirus E1A protein (8,9,11-14,17-21). This protein activates five early adenovirus genes (8,9,22). It is also capable of activating several endogenous cellular genes (23,24) and both class II and class III genes newly introduced into the cell by either infection or transfection (8,9,11,25-30). In addition to having the potential to *trans*-activate genes, a domain in the E1A protein has been found which is also involved in transcriptional repression of both cellular and viral genes linked to enhancer elements (31-35). Thus, E1A is capable of both transcriptional induction and repression. This data suggests that the cellular factors that are bound to a promoter may be important in determining whether E1A-mediated transcriptional induction or repression will occur.

Trans-activation of several adenovirus early promoters has been studied in detail (36-50). These studies have identified several important regulatory sequences upstream of the E1A, E2, E3 and E4 promoters (36-50). Although these sequences have been identified, no unique sequences required for E1A activation have been found (36-50). Mutagenesis studies have identified several important transcriptional regulatory sequences in the E2 promoter (38-40,43-45,48,49). DNA binding experiments have shown that one of these regions, containing the sequence TGACG, serves as a binding domain for a cellular protein whose activity does not change during viral infection (43,48). Binding to a different control region of the E2 promoter with homology to sequence TTTCGCGC also found in the E1A enhancer element has been identified (45,49). The activity of this binding protein, E2F, increases markedly in response to E1A (45,49). Thus, multiple cellular factors with different E1A patterns of regulation may be important in regulating the E2 transcriptional unit.

Important conserved sequences involved in transcriptional regulation have also been described for both the E3 and E4 promoters. A sequence motif

(AGATGACTA) has been shown to be important in transcriptional regulation of E3 promoter (42). Studies of the E4 promoter (46,47) have shown that two upstream copies of the sequence (TGACG) are important in its transcriptional regulation. A cellular protein, E4F1, may bind to this concensus sequence and similar sequences found in the E1A (36), E2 (40,43,44,48), E3 (41), and E4 promoters (46,47). Thus, several early adenovirus promoters may share common regulatory sequences and DNA binding proteins important in their transcriptional regulation.

Previous studies suggested several upstream promoter regions were involved in the transcriptional regulation of the E3 promoter (41,42), which is one of the most strongly induced early promoters by the E1A protein (22). Using a DNase I footprinting protocol (51), we began an analysis of the cellular proteins required for E3 transcriptional regulation both in the presence and absence of E1A. Four regions of the E3 promoter, including the TATA sequence, serve as binding sites for cellular proteins. Binding to either of two upstream regulatory sites was capable of stabilizing binding to the TATA sequence. Mutagenesis studies indicate that both of these sites, in addition to the TATA region, are required for complete E1A-induced *trans*-activation. The activity and binding characteristics of these cellular proteins do not appear to be changed by viral infection.

MATERIALS AND METHODS

Cell Lines and Tissue Culture Conditions

HeLa cells were maintained in suspension culture with MEM and 5% newborn calf serum. HeLa cells were infected with Ad2 (multiplicity of infection of 20), as described (52). Cells were harvested at 8 hours post-infection in the presence of cytosine arabinoside (ara-C). HeLa cell lines containing the neomycin resistance gene either in the presence or absence of the E1A gene were made by co-transformation of HeLa cells with the plasmid pSVneo and the E1A/E1B (nucleotides 1 to 3329) containing plasmid BE5. HeLa cells were transfected by a calcium phosphate procedure with 10 μg of each of these plasmids. G418 at 100 μg/ml was added on the second day post-transfection, and the cells were incubated in the presence of G418 for four to six weeks (53). Colonies were picked, expanded, and screened for the presence of E1A-specific mRNA by S_1 analysis and the ability to complement the adenovirus E1A deletion mutant, d1312 (9). Yield of d1312, following a single burst on the E1A-containing cell line, was ten-fold lower than found on 293 cells, an E1A/E1B-containing cell line (54). The cell lines containing the

pSVneo gene were screened for the presence of the neomycin resistance gene by Southern analysis. Transfections using the RSVCAT vector (55) indicated no difference in transfection efficiencies between the E1A-containing cell line and the control cell line. Southern analysis of transfected DNA indicated no difference in DNA stability between the E1A-containing and the control cell line (56).

Plasmid Constructions

All plasmids were derived from pE3CAT (42) (gift of N. Jones). The EcoRI (-236)/SacI (+31) fragment containing the adenovirus type 5 E3 promoter was subcloned into EcoRI/SacI pUC19. To generate the 5' deletion promoters, this subclone was first cut with a restriction enzyme, treated with either T4 polymerase or S_1 nuclease, cut with SacI (+31), and the fragment gel-isolated [T4 polymerase: EcoRI (-236), Sau96 (-173), BstXI (-130), HaeIII (-106), DdeI (-85), HhaI (-76), MaeIII (-57), SmaI (-40), MnlI (-16); S_1 nuclease: MaeIII (-52)]. These fragments were then subcloned into SmaI/SacI pUC19 for use in DNase I footprinting experiments and CAT (chloramphenicol acetyltransferase) vector construction.

To generate the upstream fragments used in the internal deletion series, the EcoRI/SacI pUC19 subclone was cut with a restriction enzyme, treated with T4 polymerase, cut with EcoRI (-236), gel-isolated, and subcloned into EcoRI/SmaI pUC19 [(BstXI (-134), DdeI (-82), SmaI (-34)]. These subclones were then cut with BamHI in the polylinker, and with EcoRI (-236), the fragments gel-isolated, and ligated with the respective BamHI (polylinker)/SacI (+31) 5' deletion subclone fragments into EcoRI/SacI pUC19. This created a series of internal deletion constructs: 134/85 (EcoRI/BstXI + DdeI/SacI), 134/52 (EcoRI/BstXI + MaeIII/S1/SacI), 82/52 (EcoRI/DdeI + MaeIII/S1/SacI), and 34/16 (EcoRI/SacI + MnlI/Sac I). The polylinker spacer in these constructs was 10 bp. Thus, the overall number of nucleotides deleted in each of these constructs was: 134/85, 39 bp; 134/52, 72 bp; 82/52, 20 bp; and 34/16, 11 bp. These internal deletion constructs were then linearized with BstXI (-130) or EcoRI (-236), treated with T4 polymerase, and recloned into SmaI/SacI pUC19 for use in DNase I footprinting experiments and CAT vector construction.

To construct the CAT vectors, the final 5' and internal deletion pUC19 mutants were cut with SacI (+31) and AccI (polylinker), and cloned into ClaI/SacI pE3CAT. This vector contains additional E3 sequences extending to the Sau3A site (+65) fused directly to the CAT gene. Constructs which fuse

SacI (+31) directly to the CAT gene gave similar results seen with these constructs (data not shown).

Transfection Conditions

For the transfection of the E3 deletion mutants, the E1A-containing or control HeLa cell lines were split on the day prior to transfection, rendering cells between 50 to 75 percent confluent at the time of transfection. Ten micrograms of each of the E3 CAT constructs were transfected onto each cell line by the calcium phosphate transfection procedure, and subjected to glycerol-shock at four hours post-transfection (56). Transfections were harvested at 48 hours, then used for assay of CAT activity, as described (57). Following autoradiography of each CAT assay, both the unacetylated and acetylated chloramphenicol were quantitated by scintillation counting. For S_1 analysis, five plates of the E1A-containing cell lines were transfected with each of the E3 CAT constructs, harvested for cytoplasmic RNA, and 50 µg of this RNA hybridized at 56°C for twelve hours with an end-labeled probe. This probe was made by cutting -236 E3 CAT with EcoRI, which cuts in the CAT gene, end-labelling with gamma ^{32}P, cutting with BamHI (from the polylinker), and gel-isolating the 534 bp fragment. Following S_1 treatment, the samples were run on an 8M urea 8% polyacrylamide gel and autoradiography performed as described (25).

Preparation of Cellular Extracts

For all extracts, a minimum of 7 mls packed cell volume was used. Nuclear extracts were prepared as described (58). Extracts were dialyzed versus a buffer containing 20 mM Tris (pH 7.9), 100 mM KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT, and 20% glycerol. This extract was loaded onto a heparin agarose column, washed with five column volumes of this same buffer, and eluted with 0.5 mM KCl. The extract was dialyzed into 20 mM Tris (pH 7.9), 100 mM KCl, 0.2 mM DTT, 0.2 mM PMSF, and used in DNase I footprinting assays (51).

DNase I Footprinting

To label the coding strand, the deletion mutants were cut with EcoRI, treated with alkaline phosphatase, and end-labeled with gamma ^{32}P. Each clone was cut with PvuII, the fragments gel-isolated, electro-eluted, and used in DNase I footprinting assays (51). One to five nanograms of end-labeled probe was added to each 50 µl reaction, along with extract (0-200 µg), poly dI-dC (3 µg), and final concentrations of 10 mM Tris (pH 7.4), 50 mM KCl, 1 mM EDTA, 1 mM DTT, and 5% glyercol. The DNA and extract were allowed to bind for 30 minutes at room temperature, the reaction volume was increased to 100 µl,

and final concentrations of DNase I (0.4-2.0 µg/ml), 5 mM MgCl, 2.5 mM $CaCl_2$ were added. The reaction was stopped after 30 seconds with phenol-chloroform, ethanol-precipitated, and loaded on a 10% polyacrylamide 8M urea sequencing gel. G+A and C+T Maxam-Gilbert sequencing reactions were performed for each probe. All gels were then subject to autoradiography.

RESULTS

Binding Domains in the E3 Promoters

To analyze cellular proteins which bind to the E3 promoter, DNase I footprinting with partially purified HeLa cell extracts was used. A number of 5' deletion mutants in the upstream promoter region, beginning at either -236, -130, -85, or -52, relative to the E3 transcriptional start site and ending at +31, were tested by DNase I footprinting on both the coding and non-coding strands (Figures 1A-1D). Only the results using the coding strand are shown, since only minimal differences in the binding domains were seen between the two strands (data not shown).

As shown in Figure 1, when the fragment from -236 to +31 was used, four regions of the E3 promoter exhibit DNase I protection. These include site IV (-156 to -179), site III (-83 to -103), site II (-47 to -67), and site I (-16 to -37) (Figure 1A). The sequence of these regions on the coding strand is shown in Figure 2, with boxes surrounding the binding domains. Site IV contains the sequence GGCCAA, which forms a portion of the CTF/nuclear factor I binding domain (3); site III contains the sequence AGATGACT, which is found in the binding domain for the AP-1 protein (59); site II contains the sequence TGACG on the non-coding strand, which is found in the binding domain for the cyclic AMP response element binding protein (CREB) (60); and site I contains the TATA sequence.

To study potential interactions between these cellular factors, a series of fragments which progressively eliminate the upstream binding sites were used for DNase I footprinting. The fragment from -130 to +31 eliminated binding site IV (Figure 1B), the fragment from -85 to +31 (Figure 1C) eliminated sites III and IV, and the fragment -52 to +31 (Figure 1D) eliminated binding sites II, III, and IV. As shown in Figure 1, deletions that remove binding site IV (Figure 1B) or sites III and IV (Figure 1C) have no effect on the binding of cellular proteins to downstream binding sites, but a deletion of sites II, III, and IV (Figure 1D) has a marked effect on the ability of cellular proteins to bind to site I.

A series of internal deletion mutants were also constructed to study the

effect of eliminating each binding site individually (Figures 1E-1I). The fragment -236Δ 134/85 (Figure 1E), which removed binding site III, gave protection over sites I, II, and IV. The fragment -236Δ 82/52 (Figure 1F), which removed binding site II and moved binding site III 26 bp upstream of site I, gave protection over sites I, III, and IV. The fragment -130Δ 82/52 (Figure 1G) which removed binding sites II and IV, and moved binding site III 26 bp upstream of site I, gave protection over sites I and III. The fragment -236 Δ 134/52 (Figure 1H), which removed binding sites II and III and moved binding site IV 44 bp upstream of site I, gave protection over site IV, but not over site I. The fragment -236Δ 34/16 (Figure 1I), which removed binding site I, gave protection over sites II, III, and and IV.

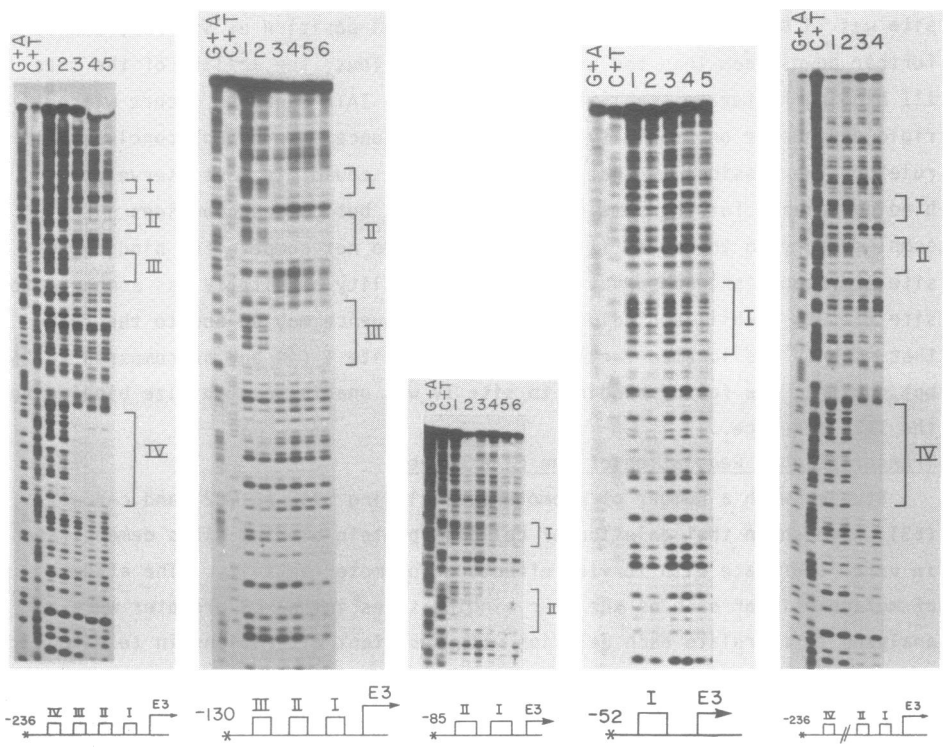
Thus, when both sites II and III were deleted, binding to the TATA sequence did not occur at the protein concentrations tested. However, when either sites II or III are left intact so that proteins bind to either region, binding over the TATA sequence is detected. The proteins binding to sites II, III, and IV bound stably in the absence of binding over site I (Figure 1I). The stabilization of factor binding over the TATA sequence by proteins binding to site III (mutants -236 Δ 82/52 and -130 Δ 82/52) occurred, even though this site was 15 bp further upstream than the normal position of site II, and 20 bp further downstream than the normal site III. Thus, the ability of the site III binding factor to stabilize binding to the TATA sequence occurs without rigid dependence on proximity to the TATA sequence. We cannot conclusively rule out the possibility that sites II and III could, in fact, serve as binding domains for the same cellular protein, but oligonucleotides complementary to the site III binding domain do not compete for binding to site II (unpublished observations). The inability of site IV, as compared to site III, to stabilize binding to the TATA sequence may be due to the fact that site IV was located further upstream of site I (44 bp, as compared to 26 bp), or that the factor binding to site IV was unable to stabilize binding to the TATA sequence.

Transcriptional Regulation of the E3 Promoter

Studies with a number of promoters, including SV40 (61,62) and c-*fos* (63), have shown that deletion of cellular protein binding sites demonstrated *in vitro* correlate with *in vivo* effects on promoter activity. The effects of mutagenesis of each of the four binding sites in the E3 promoter were analyzed to correlate each deletion with resultant alterations in *in vivo* transcription, both in the presence and absence of E1A. Each of the mutants previously discussed was transferred to a plasmid containing the CAT gene

(42). These E3 CAT constructs were then assayed following transfections onto neomycin-resistant HeLa cell lines either containing or lacking the E1A gene. The transfections were harvested 48 hours post-transfection, and CAT activity determined. Each transfection experiment was repeated several times, and the results of each experiment yielded similar data.

As shown in Figure 3A, deletions that which remove binding site IV (-173, -130, and -106) show minimal changes (less than two-fold) in both basal and E1A-induced CAT activity. There was approximately a ten- to fifteen-fold induction with these constructs in the presence of E1A, as has previously been shown (48). However, deletions that remove binding sites III and IV (-76 and -85) show a four- to five-fold decrease in both the basal and E1A-induced CAT activity (Figure 3A). Deletions that remove binding sites II, III, IV (-40, -52, and -57) have an additional seven- to ten-fold decrease in basal and E1A-induced CAT activity (Figure 3A). All mutants remain inducible by E1A, although the level of E1A induction for the -40 and -52 deletions were difficult to quantitate, due to the low basal levels of CAT activity.



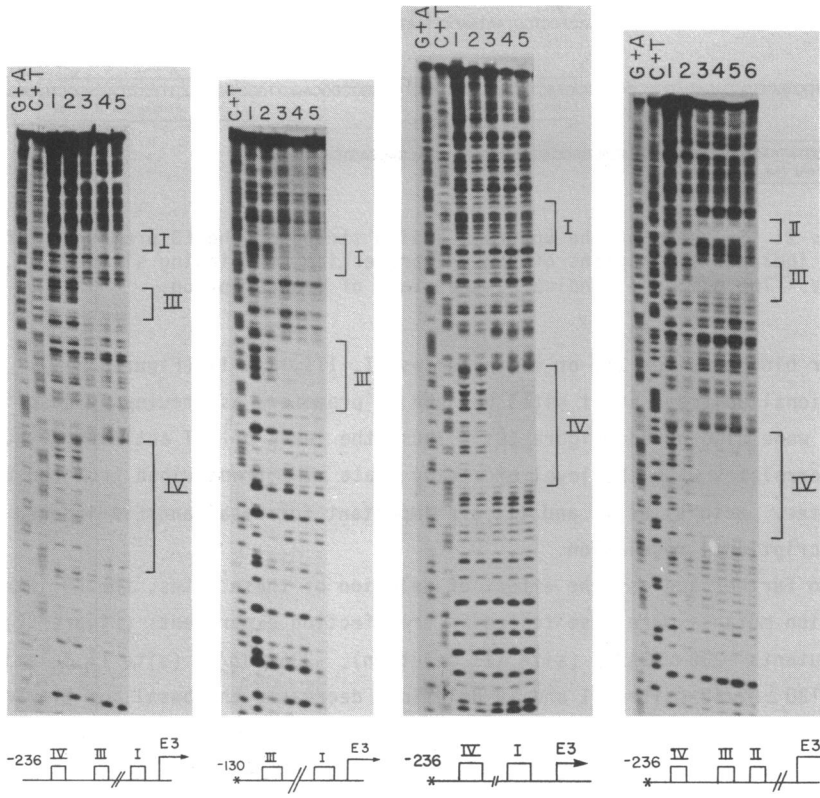


Figure 1: DNase I footprinting of E3 deletion mutants. The following E3 deletion mutants were labeled on the coding strand: (A) -236, (B) -130, (C) -85, and (D) -52. In addition, internal deletion mutants (E) -236 Δ 134/85, (F) -236 Δ 82/52, (G) -130 Δ 82/52, (H) -236 Δ 134/52, and (I) -236 Δ 34/16 were also labeled on the coding strand. Increasing amounts of partially purified HeLa cell extracts were added: Lane 1 contains no added extract; Lane 2, 10 µg; Lane 3, 50 µg; Lane 4, 100 µg; Lane 5, 150 µg; and Lane 6, 200 µg. G+A and C+T are Maxam-Gilbert sequencing lanes. All gels were subjected to autoradiography. The binding domains (I-IV) present in each fragment are indicated, and the // indicates regions of the fragment where deletions were placed.

S_1 analyses of cytoplasmic mRNA from transfections of E3 CAT deletions -236, -85, -76, and -52 onto the E1A-containing HeLa cell line were also performed. As shown in Figure 3B, with a probe labeled at the EcoRI site in the CAT gene, a 317 bp fragment specific for correctly initiated mRNA from the E3 promoter was seen. No detectable E3 CAT mRNA was seen when these mutants were transfected onto the control cell line, which lacks E1A (data not shown). The intensity of the 317 bp band decreased markedly, with deletion of

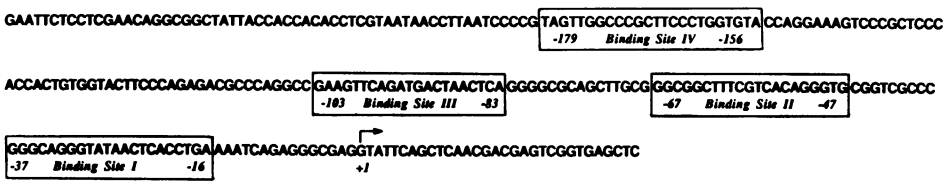


Figure 2: Sequence of the upstream coding strand of the E3 promoter. The boxes indicate the regions of DNase I protection for binding sites I, II, III, and IV. The number +1 indicates the start of transcription.

either binding sites III or IV, or sites II, III, and IV (Figure 3B). Additional upstream start sites for the E3 promoter, as previously described (44), were also noted (Figure 3B). Thus, the level of CAT activity (Figure 3A) correlates with the level of steady-state mRNA transcribed from the E3 promoter. Both sites II and III are important for basal and E1A-induced transcriptional regulation.

To further address the effect of deletion of these sites, the E3 internal deletion mutants were also tested in transfection experiments (Figure 3C). The mutants -236 Δ 134/85 (site III deletion), -236 Δ 82/52 (site II deletion), and -130 Δ 82/52 (sites II and IV deletion) decreased the basal and E1A-induced CAT activity five- to eight-fold (Figure 3C). Deletion mutant -236 Δ 134/52 (sites II and III deletion) decreased the basal and E1A-induced CAT activity at least thirty-fold (Figure 3C). Thus, the elimination of both binding sites II and III gave a decrease of greater magnitude in the level of CAT activity than deletion of either region alone. Mutant -236 Δ34/16 (site I deletion) showed a three- to four-fold decrease in CAT activity, and S_1 analysis demonstrated correct E3 RNA start sites (data not shown).

These results indicate that sites I, II and III are important in both basal and E1A-induced transcriptional regulation of the E3 promoter. The deletion of any of the three sites does not eliminate E1A-induced *trans*-activation, although the presence of all three sites is required for complete *trans*-activation. Table I indicates the percentage of CAT conversion in both the presence and absence of E1A for each of the E3 deletion mutants, and Figure 4 schematically shows the binding sites and level of CAT activity for several of the E3 deletion mutants.

Effects of Viral Infection on the Binding of Cellular Protein

The mechanism by which E1A induces transcription is not known, but an increase in the activity or an alteration in binding characteristics of cellular transcription factors remains a possibility (45,49). Studies on the

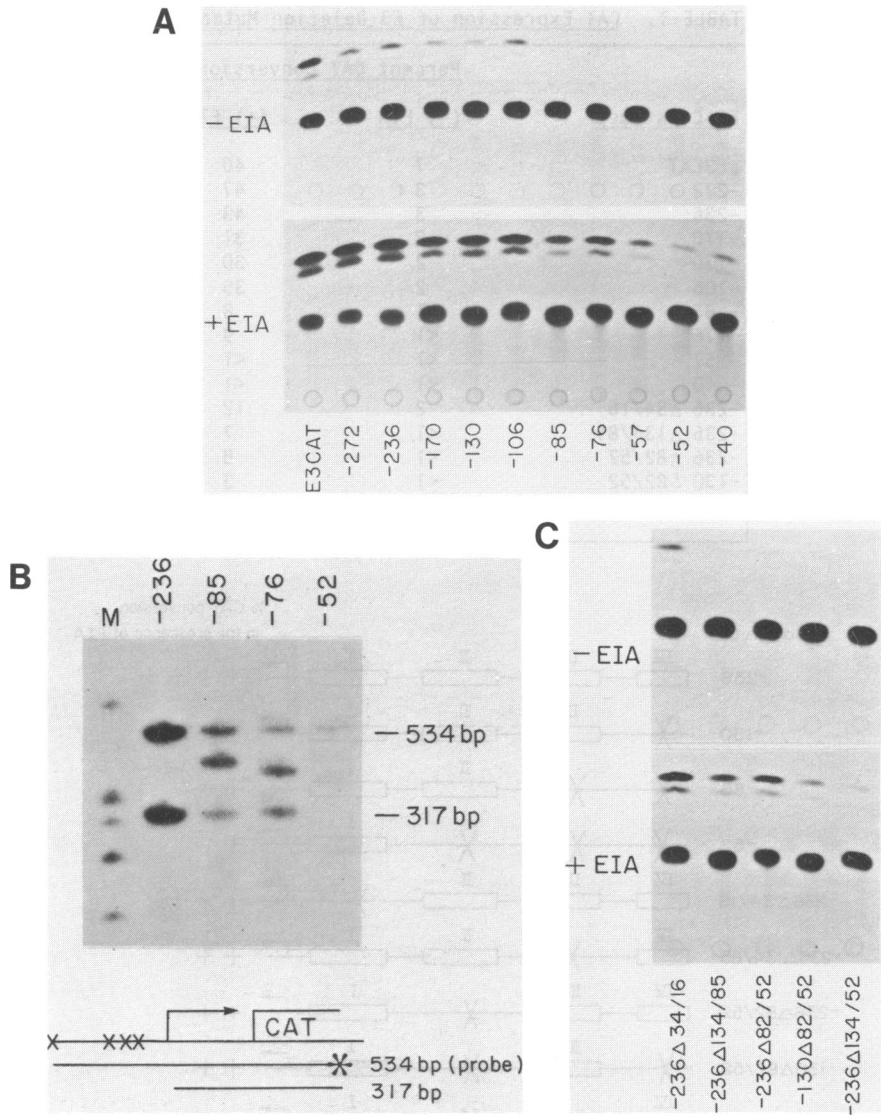


Figure 3: CAT assays of E3 mutants in the presence and absence of E1A.
(A) An E1A-containing HeLa cell line or control HeLa cell line was transfected with 10 μg of the E3 CAT vectors for either pE3CAT, -272, -236, -170, -106, -85, -76, -57, -52, and -40. Transfections were harvested at 48 hours post-transfection, and CAT assays were performed as described.
(B) S_1 analysis of cytoplasmic mRNA from transfections onto the E1A-containing cell line with E3 CAT deletion mutants, -236, -85, -76 and -52.
(C) Both cell lines were assayed for CAT activity following transfections of the internal deletion mutants, -236 Δ34/16, -236 Δ134/85, -236 Δ82/52, -130 Δ82/52, and -236 Δ134/52.

TABLE 1. CAT Expression of E3 Deletion Mutants

E3 Construct	Percent CAT Conversion (-) E1A	(+) E1A
pE3CAT	7	40
-272	3	47
-236	3	43
-170	2	31
-130	2	30
-106	2	35
- 85	<1	8
- 76	<1	9
- 57	<1	<1
- 40	<1	<1
-236 Δ34/16	2	12
-236 Δ134/85	<1	7
-236 Δ82/52	<1	8
-130 Δ82/52	<1	3
-236 Δ134/52	<1	<1

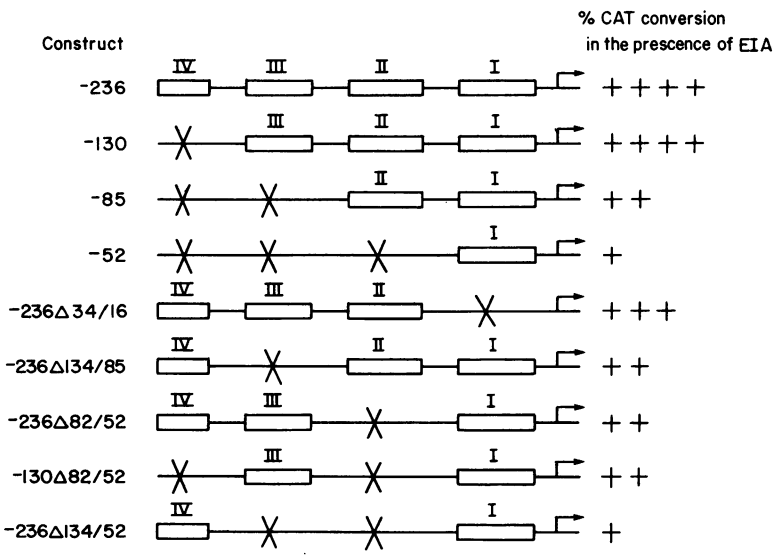


Figure 4: Schematic representation of the E3 deletion mutant binding sites and their effect on E3 transcriptional regulation. E3 deletion mutants removing either no binding sites (-236), binding site IV (-130), binding sites III and IV (-85), or binding sites II, III and IV (-52) are shown. In addition, internal deletion mutants removing binding site I (-236 Δ34/16), binding site III (-236 Δ134/85), binding site II (-236 Δ82/52), binding sites II and IV (-130 Δ82/52), and binding sites II and III (-236 Δ134/52) are shown. The relative amount of CAT conversion in the presence of E1A for each construct is shown.

E2 promoter have shown that one cellular protein has increased binding activity in the presence of E1A (45,49), while another protein is unaffected (43,48). The cellular proteins binding to the E4 promoter have been shown to be unaffected by the presence of the E1A protein (47). For genes transcribed by RNA polymerase III, it has been shown that E1A increases the activity of the III-C factor (28,64).

To examine the role of viral infection on binding of the cellular proteins to the E3 promoter, extracts prepared from adenovirus-infected or -uninfected cells were tested in DNase I footprinting experiments (Figure 5). The infected extracts were prepared at 8 hours post-infection of HeLa cells with adenovirus (Ad 5). The presence of E1A protein in these infected extracts was confirmed by Western blot analysis, using a polyclonal E1A antisera (data not shown). With increasing amounts of either uninfected (Figure 5, Lanes 3-6) or infected (Figure 5, Lanes 7-10) extracts, there was no difference in either the amount or pattern of clearing for any of the four binding sites of the E3 promoter. Thus, E1A does not appear to change the DNA-binding activity of factors binding to the E3 promoter, as determined by DNase I footprinting assays.

DISCUSSION

Multiple cellular proteins bind upstream of the adenovirus E3 promoter. Deletion analysis has shown that three of these binding sites are important for both basal and E1A-induced transcriptional activation. The sites identified by DNase I footprinting *in vitro* correlate with sites that appear to be important for *in vivo* transcriptional activity. Similar correlation between *in vivo* transcriptional activity and *in vitro* binding data have been made for other promoters, including SV40 (61,62) and c-*fos* (63). Thus, a study of *in vitro* binding sites and their cellular binding proteins likely correlates with the presence of similar regulatory sites and factors *in vivo*.

Previous studies of the E3 promoter have defined several regions important for transcriptional regulation (41,42). Mutagenesis of the E3 promoter indicated that the TATA sequence, a region between -55 and -57, and a region between -111 and -237, were important for both basal and E1A-induced transcriptional activation (41). No unique sequences were found which were required for E1A induction. Another set of E3 mutants were used to identify a region between -82 and -103, important for both basal and E1A-induced transcription (42). This latter region could also function as an upstream E1A-inducible enhancer element for the thymidine kinase gene. Our results

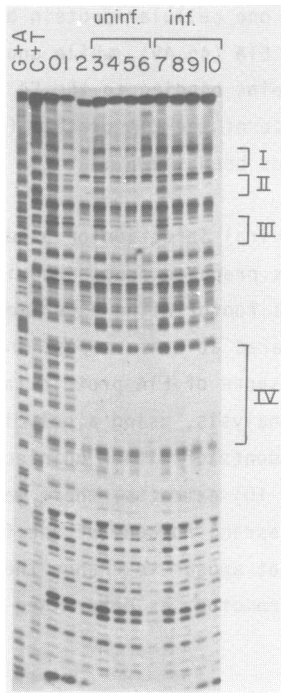


Figure 5: Effect of viral infection on DNase I footprinting of the E3 promoter. The coding strand of the E3 fragment, -236 to +31, was end-labeled and used in DNase I footprinting with partially purified HeLa cell extracts. Lane 0 contains no added protein; Lane 1, 10 µg; and Lane 2, 100 µg of HeLa cell extracts. Lanes 3-6 contain uninfected HeLa cell extracts; Lanes 7-10, infected HeLa cell extracts. Lanes 3 and 7 contains 20 µg of extract; Lanes 4 and 8, 50 µg of extract; Lanes 5 and 9, 100 µg of extract; and Lanes 6 and 10, 150 µg of extract. G+A and C+T are Maxam Gilbert sequencing lanes. The binding domains (I-IV) present in each fragment are indicated.

confirm that the regions important for transcriptional regulation of the E3 promoter identified in previous studies are indeed sites of cellular binding proteins (41,42). However, we do not see a significant effect of the -111 to -237 region on either basal or E1A-induced transcriptional activity.

Three of the four regions which bind cellular proteins appear to be important in both the basal and E1A-induced transcriptional activity of the E3 promoter. Site I includes the TATA sequence, and appears to be important for both basal and E1A-induced transcription. Stable binding to this site requires the presence of either the site II or site III binding domains. Since oligonucleotide competition experiments suggest that sites II and III bind different factors, these results would indicate that alternative upstream

factors can stabilize binding to the TATA sequence. Deletion of site II in these constructs (-236 Δ 82/52 and -130 Δ 82/52) (Figure 1) places site III 26 bp from site I; the position of site II is normally 10 bp from site I. This suggests that there is not a rigid dependence on spacing of these factors to stabilize binding to the TATA sequence. They also suggest that near-integral multiples of the 10.5 bp per turn of B-DNA *in vitro* (65) may not be required to maintain factor interactions, as has been suggested for the SV40 promoter (66). However, since the specific sites of factor binding cannot be accurately determined by DNase I footprinting, further studies will be required, using reagents such as methidiumpropyl-EDTA-Fe (11) to clearly show the distance between these factor binding sites (67).

Results with the SV40 promoter using insertional mutants between the SP1 binding sites and the TATA element suggest that these SP1 sites may stabilize binding of factors to the TATA element (66). A similar effect of SP1 binding sites on the binding of cellular factors to the TATA element has recently been shown for the human immunodeficiency virus (68). In the adenovirus major late promoter, it has been shown that an upstream binding factor (USF) stabilizes factor binding to the TATA sequence (67). This TATA binding factor (II-D) has not yet been purified to homogeneity. A characterization of TATA binding factors in other promoters has been difficult, due to the inability to footprint these factors. Interactions with upstream regulatory elements and the possibility suggested from studies on several yeast promoters that more than one class of TATA elements exist (69) may explain the difficulty in characterizing these factors. The presence of TATA elements in diverse sets of genes transcribed by RNA polymerase II and their critical role in transcriptional regulation (70) suggests that the factors binding to the TATA element may interact with promoter-specific transcription factors in a number of different genes.

The region containing the TATA sequence is not essential for E1A-induced activation of the E3 promoter. A previous study with the E3 promoter has shown that this promoter is induced by E1A in the absence of the TATA sequence (41). The E2 early promoter, which lacks a consensus TATA sequence, is also induced by E1A (40,43,44,48). The TATA sequence has been reported to be necessary for E1A induction of the B-globin promoter and the adenovirus E1B promoter (11,71). Thus, the role of the TATA factor in E1A induction may be a promoter-specific phenomenon.

Site II is important for both basal and E1A-induced transcriptional regulation. This region contains the sequence TGACG on the non-coding strand,

and this sequence is also found in important regulatory regions of the E2 (39,40) and E4 (46,47) promoters, the 21 bp repeats of the HTLV-I and HTLV-II long terminal repeats (72-74), and in cyclic AMP-responsive promoter elements (75,76). A 43 kd (kilodalton) protein (CREB) which binds to the TGACG sequence in the sommatostatin gene (60) has been purified, and may be identical to the recently described E4F1 protein, which binds to the E1A, E2, E3, and E4 promoters (47). CREB may be important in the regulation of a number of inducible genes, and it may serve as a target for different viral *trans*-activator proteins such as E1A and the HTLV *tat* protein (15).

Site III is also important for both basal and E1A-induced transcriptional regulation. This region contains the sequence AGATGACT, which is also found in the SV40 enhancer (59,77,78), the metallothionein gene (77,78), and in yeast genes regulated by the GCN4 protein (79). This sequence has been shown to serve as the binding site for a 47 kd protein, AP-1, in both the SV40 enhancer and metallothionein promoters (59,77,78). The AP-1 protein and the CREB protein each bind upstream of a number of cellular and viral genes, and alterations in either the level, or interactions of these proteins, may be important in the process of transcriptional induction. Since these proteins bind to similar DNA sequences and are of similar size, it is possible that they may be members of a family of DNA binding proteins.

Site IV contains a sequence, GGCCAA, which comprises a portion of a low affinity CTF/nuclear factor I binding site (3). This region appears to have only slight effects on the transcriptional regulation of the E3 promoter in HeLa cells. An upstream binding domain also lacking marked transcriptional regulatory effects has recently been described for the E2 promoter (48). Whether either of these factors stabilize or alter the binding of E2 or E3 transcription factors is not known. Both may serve to block transcription from these opposing transcriptional units.

Studies of several adenovirus promoters with combined *in vivo* and *in vitro* analysis lead to a number of general conclusions about the cellular factors involved in E1A *trans*-activation. First, multiple cellular proteins are involved in the transcriptional regulation of these promoters (43,45,47-49). Second, although there are no unique sequences required for E1A *trans*-activation, early adenovirus promoters appear to share several common regulatory sequences, and possibly, common cellular proteins may be involved in their transcriptional regulation (36-50). Third, the E1A protein may have both direct and indirect effects on cellular transcription factors (28,43,45,47-49,64).

It has been shown that for at least two transcription factors, III-C (28,64) and E2F (45,49), E1A increases their activity. For a variety of other transcription factors involved in E2 (43,48), E3, and E4 (47) regulation, the role of the E1A protein remains less clear. E1A may increase the interactions of cellular proteins by post-translational modifications, which are difficult to assay *in vitro* by current techniques. It may also directly interact with or modify a complex of cellular proteins, leading to an increased number or more active stable transcriptional complexes (80-82). The role of the E1A protein in transcriptional activation will require the purification of these cellular transcription factors and their *in vitro* analysis in both the presence and absence of the E1A protein.

ACKNOWLEDGEMENTS

We would like to thank Arnie Berk for suggestions in preparing this manuscript, Danny Hong and Janna Chao for technical assistance, and Wendy Aft for preparation of this manuscript. This work was supported by a grant CA 30981 from the NIH, and JFRA-146 from the American Cancer Society. J.G. was supported by NIH grant GM0-80942 in the UCLA Medical Scientist Training Program.

REFERENCES

1. Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1986) *Science* 234, 47-52.
2. Chodosh, L.A., Carthew, R.W. and Sharp, P.A. (1986) *Mol. Cell. Biol.* 6, 4723-4733.
3. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. and Tjian, R. (1987) *Cell* 48, 79-89.
4. Rosenfeld, P.J. and Kelly, T.J. (1986) *J. Biol. Chem.* 261, 1398-1408.
5. Goodbourn, S., Zinn, K. and Maniatis, T. (1985) *Cell* 41, 509-520.
6. Greenberg, M.E. and Ziff, E.B. (1984) *Nature* 311, 433-438.
7. Imbra, R.J. and Karin, M. (1987) *Mol. Cell. Biol.* 7, 1358-1363.
8. Berk, A.J., Lee, R., Harrison, T., Williams, J. and Sharp, P.A. (1979) *Cell* 17, 935-944.
9. Jones, N. and Shenk, T. (1979) *Proc. Natl. Acad. Sci. USA* 76, 3665-3669.
10. Nevins, J.R. (1981) *Cell* 26, 213-220.
11. Green, M.R., Treisman, R., and Maniatis, T. (1983) *Cell* 35:137-148.
12. Moran, E., Zerler, B., Harrison, T.M. and Mathews, B. (1986) *Mol. Cell. Biol.* 6, 3470-3480.
13. Moran, E., Grodzicker, T., Roberts, R.J., Mathews, M.B. and Zerler, B. (1986) *J. Virol.* 57, 765-775.
14. Ferguson, B., Krippl, B., Andrisani, O., Jones, N., Westphal, H. and Rosenberg, M. (1985) *Mol. Cell. Biol.* 5, 2653-2661.
15. Chen, I.S.Y., Cann, A.J., Shah, N.P. and Gaynor, R.B. (1985) *Science* 230, 570-573.
16. Sodroski, J.G., Rosen, C.A. and Haseltine, W.A. (1984) *Science* 225, 381-385.

17. Perricaudet, M., Akusjarvi, G., Virtanen, A. and Peterson, U. (1979) *Nature* 281, 694-696.
18. Ricciardi, R.P., Jones, R.L., Cepko, C.L., Sharp, P.A. and Roberts, B.E. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6121-6125.
19. Montell, C., Fisher, E.F., Caruthers, M.H. and Berk A.J. (1982) *Nature* 295, 380-384.
20. Winberg, G. and Shenk, T. (1984) *EMBO J.* 3, 1907-1912.
21. Glenn, G.M. and Ricciardi, R.P. (1987) *Mol. Cell. Biol.* 7, 1004-1011.
22. Osborne, T.F., Gaynor, R.B. and Berk, A.J. (1982) *Cell* 29, 139-148.
23. Nevins, J. (1982) *Cell* 29, 913-919.
24. Stein, R. and Ziff, E.B. (1984) *Mol. Cell. Biol.* 4, 2792-2801.
25. Gaynor, R.B., Hillman D. and Berk, A.J. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1193-1197.
26. Svenson, C. and Akusjarvi, G. (1984) *EMBO J.* 3, 789-794.
27. Berger, S.L. and Folk, W.R. (1985) *Nucleic Acids Res.* 13, 1413-1428.
28. Hoeffler, W.K. and Roeder, R.G. (1985) *Cell* 41, 955-963.
29. Gaynor, R.B., Feldman, L.T. and Berk, A.J. (1985) *Science* 230, 447-450.
30. Lillie, J., Green, M., and Green, M.R. (1986) *Cell* 46, 1043-1051.
31. Borrelli, E., Hen, R. and Chambon, P. (1984) *Nature* 312, 608-612.
32. Velcich, A.F., Kern, F.G., Basilico, T.C. and Ziff, E.B. (1986) *Mol. Cell. Biol.* 6, 4019-4025.
33. Hen, R., Borrelli, E. and Chambon, P. (1985) *Science* 230, 1391-1394.
34. Velcich, A.F. and Ziff, E. (1985) *Cell* 40, 705-716.
35. Stein, R. and Ziff, E. (1987) *Mol. Cell. Biol.* 7, 1164-1170.
36. Hearing, P. and Shenk, T. (1983) *Cell* 33, 695-703.
37. Weeks, D.L. and Jones, N.C. (1983) *Mol. Cell. Biol.* 3, 1222-1234.
38. Elkaim, R., Goding, C. and Kedinger, C. (1984) *Nucleic Acids Res.* 11, 7105-7117.
39. Imperiale, M.J., Hart, R.P. and Nevins, J.R. (1985) *Proc. Natl. Acad. Sci. USA* 82, 381-385.
40. Murthy, S.C.S., Bhat, G.P. and Thimmappaya, R. (1985) *Proc. Natl. Acad. Sci. USA* 82, 2230-2234.
41. Leff, T., Gorden, J., Elkaim, R. and Sassone-Corsi, P. (1985) *Nucleic Acids Res.* 13, 1209-1221.
42. Weeks, P.L. and Jones, N.C. (1985) *Nucleic Acids Res.* 13, 5389-5402.
43. Siva Raman, L., Subramanian, S. and Thimmappaya, R. (1986) *Proc. Natl. Acad. Sci. USA* 83, 5914-5918.
44. Zajchowski, D.A., Bouef, H. and Kedinger, C. (1985) *EMBO J.* 4, 1293-1300.
45. Kovesdi, I., Reichel, R. and Nevins, J.R. (1986) *Cell* 45, 219-228.
46. Gilardi, P. and Perricaudet, M. (1986) *Nucleic Acids Res.* 14, 9035-9049.
47. Lee, K.A.W. and Green, M.R. (1987) *EMBO J.* 6, 1345-1353.
48. Boeuf, H., Zachowski, D.A., Tamura, T., Hauss, C. and Kedinger, C. (1987) *Nucleic Acids Res.* 15, 509-527.
49. Kovesdi, I., Reichel, R. and Nevins, J. (1987) *Proc. Natl. Acad. Sci. USA* 84, 2180-2184.
50. Kingston, R.E., Kaufman, R.J. and Sharp, P.A. (1984) *Mol. Cell. Biol.* 4, 1970-1985.
51. Galas, D. and Schmitz, A. (1978) *Nucleic Acids Res.* 5, 3157-3170.
52. Gaynor, R.B., Tsukamoto, A., Montell, C. and Berk, A.J. (1982) *J. Virol.* 44, 276-285.
53. Southern, P.J. and Berg, P. (1982) *J. Mol. Appl. Genet.* 1, 327-341.
54. Graham, F.L., Smiley, J., Russel, W.C. and Nairn, P. (1977) *J. Gen. Virol.* 36, 59-64.
55. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and de Cromborghe, B. (1982) *Proc. Natl. Acad. Sci. USA* 79, 6777-6781.
56. Alwine, J.C. (1985) *Mol. Cell. Biol.* 5, 1034-1042.

57. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) *Mol. Cell. Biol.* 2, 1044-1051.
58. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* 11, 1475-1489.
59. Lee, W., Haslinger, A., Darin, M. and Tjian, R. (1987) *Nature* 325, 368-372.
60. Montminy, M.R. and Bilezikgian, L.M. (1987) *Nature* 328, 175-178.
61. Wildeman, A.G., Zenke, M., Schatz, C., Wintzerith, M.M., Grundstrom, T., Matthes, H., Takahashi, K. and Chambon, P. (1986) *Mol. Cell. Biol.* 6, 2098-2105.
62. Zenke, M., Grundstrom, T., Matthes, H., Winterith, M., Schatz, C., Wildeman, A. and Chambon, P. (1986) *EMBO J.* 5, 387-397.
63. Gilman, M.Z., Wilson, R.N. and Weinberg, R.A. (1986) *Mol. Cell. Biol.* 6, 4305-4316.
64. Yoshinaga, S., Dean, N., Han, M. and Berk, A. (1986) *EMBO J.* 5, 343-354.
65. Wang, J. (1979) *Proc. Natl. Acad. Sci. USA* 76, 200-203.
66. Takahashi, K. Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. and Chambon, P. (1986) *Nature* 319, 121-126.
67. Sawadogo, M. and Roeder, R.G. (1985) *Cell* 43, 165-175.
68. Garcia, J., Wu, F., Mitsuyasu, R. and Gaynor, R.B. (1987) *EMBO J.*, in press.
69. Struhl, K. (1986) *Mol. Cell. Biol.* 6, 3847-3853.
70. Breathnach, A. and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349-383.
71. Wu, L., Rosser, D.S.E., Schmidt, M.G. and Berk, A.J. (1987) *Nature* 326, 512-515.
72. Fujisawa, J., Seiki, M., Sato, M. and Yoshida, M. (1986) *EMBO J.* 5, 713-718.
73. Shimotohno, K., Takano, M., Terunchi, T. and Miwa, M. (1986) *Proc. Natl. Acad. Sci. USA* 83, 8112-8116.
74. Brady, J., Jeang, K.T., Duvall, J. and Khoury, G. (1987) *J. Virol.* 61, 2175-2181.
75. Comb, M., Birnberg, N.C., Seashotz, A., Herbert, E. and Goodman, H.M. (1986) *Nature* 323, 353-356.
76. Montminy, M.R., Sevarino, K.A., Wagner, J.A., Mandel, G. and Goodman, R.H. (1986) *Proc. Natl. Acad. Sci. USA* 83, 6682-6686.
77. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R.J., Rahmsdorf, H.J., Jonat, C., Herrlich, P. and Karin, M. (1987) *Cell* 49, 729-739.
78. Lee, W., Mitchell, P. and Tjian, R. (1987) *Cell* 49, 741-752.
79. Hope, I.A. and Struhl, K. (1985) *Cell* 43, 177-188.
80. Bogenhagen, D.F., Wormington, W.M. and Brown, D.D. (1982) *Cell* 28, 413-421.
81. Davison, B.L., Egly, J.-M., Mulvihill, E.R. and Chambon, P. (1983) *Nature* 301, 680-686.
82. Gaynor, R.B. and Berk, A.J. (1983) *Cell* 33, 683-693.