JOURNAL OF VIROLOGY, Dec. 1986, p. 935–942
0022-538X/86/120935-08$02.00/0
Copyright © 1986, American Society for Microbiology
Vol. 60, No. 3

# Spacing between Simian Virus 40 Early Transcriptional Control Sequences Is Important for Regulation of Early RNA Synthesis and Gene Expression

KAMEL KHALILI,* GEORGE KHOURY, AND JOHN BRADY
*Laboratory of Molecular Virology, National Cancer Institute, Bethesda, Maryland 20892*

Received 4 August 1986/Accepted 17 August 1986

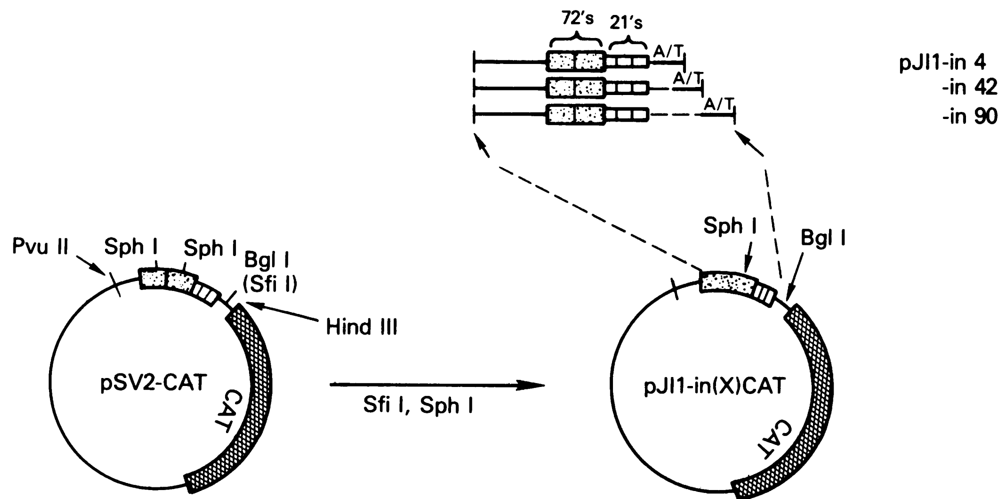
**We have analyzed the effect of insertion mutants between the simian virus 40 (SV40) 21-base pair (bp) repeats and the early-early (EE) TATA sequence. Insertion of 4, 42, or 90 bp of DNA at the SV40 *NcoI* site (map position 37) has been analyzed for its effect on expression of the SV40 early gene and positioning of the RNA 5′ ends. Insertion of 4 bp reduced SV40 early promoter-dependent chloramphenicol acetyltransferase (CAT) expression by six- to eightfold. Increasing the size of the insertion to 42 or 90 bp resulted in a further drop in early gene expression to basal levels. At the RNA level, the 4-bp insertion reduced EE RNA synthesis approximately 10-fold. No concomitant increase in late-early (LE) RNA synthesis was observed. Insertion of 42 or 90 bp of DNA resulted in a decrease of EE RNA synthesis and a stimulation of LE RNA synthesis. Deletion of the SV40 72-bp repeats from the insertion mutants demonstrated that some, but not all, of the LE RNA depends upon the presence of these sequences. These studies suggest that the ability of RNA polymerase II to utilize the EE (TATA-directed) transcriptional control sequence requires an interaction with the upstream 21-bp repeats or the 72-bp repeats or both. That LE RNA levels in pJI1-in42 CAT and pJI1-in90 CAT were equivalent to the level of EE RNA in pJI1-CAT, yet the level of CAT gene expression was decreased >10-fold, suggests that LE mRNA is under translational control and probably prefers a 5′ initiation codon proximal to that of the CAT gene.**

Simian virus 40 (SV40) has two early transcriptional units that are transcribed from a region near the origin of replication (33). The early-early (EE) transcription unit, whose RNA encodes a large-T antigen and small-t antigen, predominates at early times postinfection. The late-early (LE) transcription unit, whose RNA initiation sites are located upstream of the EE TATA box, have the potential to code for both small-t and large-T antigen and a small 2.7-kilodalton protein, the SV40 early leader protein (K. Khalili, J. Brady, and G. Khoury, manuscript in preparation). Normally, the LE transcription unit functions at late times postinfection, after DNA replication (16). The EE SV40 promoter has been studied in detail. There are at least three distinct functional elements which control the initiation specificity and efficiency of early RNA transcription. A Goldberg-Hogness (TATA) box is responsible for positioning the major EE 5′ termini (4, 17). The results of several studies have demonstrated that the SV40 21-base pair (bp) repeats, which contain six guanine-cytosine (G-C)-rich hexanucleotide repeats, directly affect the efficiency of EE RNA synthesis in vitro and in vivo (13–15, 21). Finally, the SV40 72-bp repeats or enhancer element potentiates the efficiency of transcriptional initiation in a position- and orientation-independent manner (1, 20, 28). The regulatory sequences which determine the efficiency and accuracy of LE RNA synthesis have not been identified definitively. In contrast to the EE transcription unit, there is no TATA-like sequence upstream of the LE RNA initiation sites to position LE RNA 5′ ends. In view of the results of Mishoe et al. (27), it is possible that sequences within the 21-bp repeats specify the site of LE RNA initiation. The G-C-rich 21-bp repeat region has been shown to regulate the level of LE RNA synthesis (3). Specifically, Chambon and his colleagues demonstrated that the four G-C repeats most proximal to the 72-bp repeats are important elements of the SV40 LE promoter in vitro and in vivo. These studies further suggested that the LE promoter is weaker than the EE promoter and must compete for common transcription factors. At present, it is not clear whether the 72-bp repeats influence LE promoter function (7, 8, 34).

Transcription analysis of promoter mutants, in vivo and in vitro competition studies, and binding experiments of partially purified transcription factors with the SV40 and other eucaryotic promoters suggest that the upstream *cis*-acting transcription-regulatory sequences directly interact with specific transcription factors (9, 11, 13, 29, 31, 32, 35). Binding of a protein to the eucaryotic TATA box (11, 29), Spl to the G-C-rich sequences in the SV40 21-bp repeats and the herpes simplex virus thymidine kinase promoter (13, 24) and USF or MLTF to the −52 to −63 region of the adenovirus late promoter (9, 31), has recently been reported. Another phenomenon which may contribute to enhanced promoter activity is the physical interaction of the transcription factors (6, 24, 31). In this regard, the actual distance between the upstream elements becomes important.

In this study, we have specifically analyzed the spacing between two transcriptional domains of the SV40 early promoter, the SV40 EE TATA box and the 21-bp repeats, to gain insight into some of the factors which regulate EE and LE RNA synthesis and early gene expression. Our results demonstrate that insertion of as little as 4 bp of DNA dramatically decreases the efficiency of the EE promoter. At the same time, an increase in LE promoter activity is observed. It appears that the G-C-rich 21-bp repeat se-

* Corresponding author.



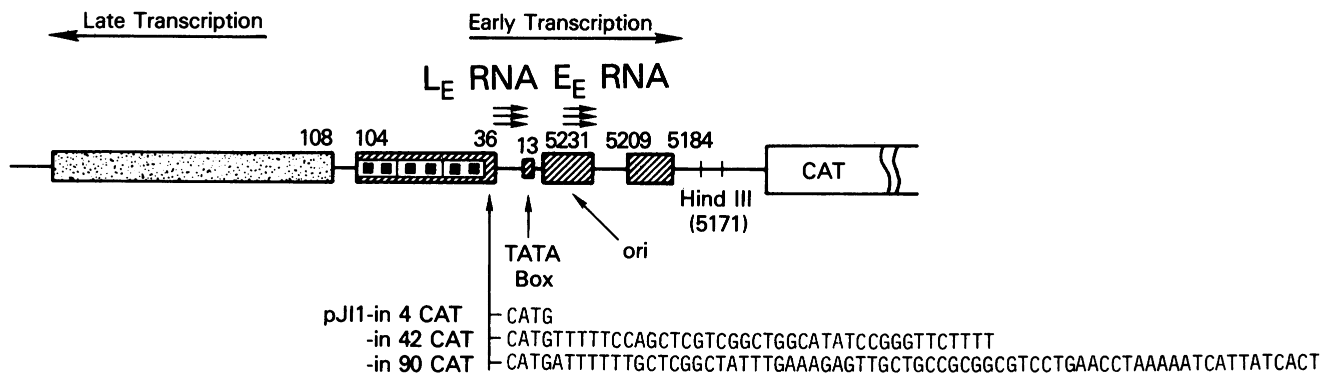




FIG. 1. Structure of the SV40 pJI1-in(X) CAT constructs (A) and regulatory region of the SV40 early promoter in pJIin(X) CAT plasmids (B). (A) The *BglI-SphI* fragment from insertion mutants (23) was ligated into the unique *SfiI-SphI* restriction sites in pSV2CAT, yielding pJI1-CAT, pJI1-in4 CAT, pJI1-in42 CAT, and pJI1-in90 CAT. (B) The origin of replication, T-antigen-binding sites I, II, and III, TATA box, G-C-rich motifs within the 21-bp repeats, two 72-bp enhancer elements, and the 5' ends of EE and LE RNA are depicted. The nucleotide sequence contained in insertion mutants pJI1-in4, -in42, or -in90 CAT, as determined by Innis and Scott (23), is presented below the schematic.

quences are a major determinant of the start sites for early SV40 transcription.

Deletion of the 72-bp repeats from the insertion mutants demonstrated that, while all of the EE RNA seems to depend on the enhancers, the synthesis of some, but not all, of the LE RNAs depends upon the presence of these sequences. Of particular significance is our finding that the levels of LE RNA synthesis and early gene expression do not correlate. This result suggests that post-transcriptional regulation contributes to the level of SV40 early protein synthesis.

## MATERIALS AND METHODS

**Cell culture and transfection.** CV-1 cells were passed into 10-cm culture dishes with Dulbecco modified eagle medium supplemented with 10% fetal calf serum. Twenty-four hours later, when cells were 60 to 80% confluent, they were transfected by the calcium phosphate precipitation technique (19). The total DNA concentration was kept constant (30 μg) with carrier plasmid DNA. Cells were incubated with transfection mixtures for 12 h, washed, and cultured with fresh Dulbecco modified Eagle medium until 48 h posttransfection.

**CAT assays.** Cells were harvested and assayed for chloramphenicol acetyltransferase (CAT) activity as described by Gorman et al. (18).

**S1 nuclease analysis of RNA.** RNA was extracted and analyzed by the S1 nuclease technique as described previously (5). Briefly, RNA was extracted by the hot-acid phenol method as described by Queen and Baltimore (30). Purified RNA (20 μg) and labeled DNA probe were suspended in 20 μl of 0.4 M NaCl–0.04 M PIPES [piperazine-*N*,*N*'-bis(2-ethanesulfonic acid; pH 6.4]–1.25 mM EDTA–80% formamide and heat denatured at 65°C for 15 min. The samples were then transferred immediately to a 37°C water bath and incubated for at least 6 h. Hybridizations were terminated by the addition of 9 volumes of 0.25 M NaCl–0.03 M sodium acetate (pH 4.6)–1 mM $ZnSO_4$, which had been preequilibrated to 4°C. S1 nuclease was added to a final concentration of 800 U/ml, and the reaction sample was incu-

bated for 60 min at 37°C. Protected DNA fragments were then purified and analyzed by electrophoresis in a denaturing acrylamide-urea gel.

**Primer extension analysis of RNA.** A 285-base *Eco*RI-*Hin*dIII DNA fragment of the CAT-coding region was synthesized from an M13mp19 vector containing an *Eco*RI-*Sph*I fragment of pSV2CAT. The probe was homogeneously labeled by synthesis of the DNA strand, using Klenow fragment in the presence of [$^{32}$P]dCTP. Approximately 50,000 cpm of the primer was hybridized with 30 μg of total cytoplasmic RNA for 16 h, as described above. Hybrids were recovered by precipitation with ethanol, and primer extension reactions were carried out with reverse transcriptase as described by Hernandez and Keller (22).

**In vitro transcription.** Nuclear extracts were prepared from uninfected HeLa cells as described by Dignam et al. (12). Nuclear extracts were precipitated with ammonium sulfate and dialyzed as described by Wildeman et al. (35). Transcription reactions were done in a final volume of 10 μl, using 2 to 4 μl of HeLa cell nuclear extract. Unlabeled ATP, CTP, and GTP were added to a final concentration of 0.5 mM and [$^{32}$P]UTP was added to 10 μM in a reaction mixture containing 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.9), 2 mM $MgCl_2$, 50 mM KCl, 20 mM spermidine hydrochloride, 12% glycerol, and 0.2 μg of template DNA. After all of the components of the reaction mixture were mixed on ice, the reactions were carried out for 30 min at 35°C. Samples were sodium dodecyl sulfate-phenol-chloroform extracted, chloroform extracted, ethanol precipitated, and analyzed on 6% polyacrylamide gels containing 8.3 M urea. DNA templates were obtained by *Eco*RI digestion of the plasmids shown in Fig. 1A. Templates were sodium dodecyl sulfate-phenol-chloroform extracted, chloroform extracted, ethanol precipitated, and resuspended at 0.1 μg/ml.

**Plasmids.** Plasmids containing insertions between the SV40 early TATA box and the 21-bp repeats were kindly

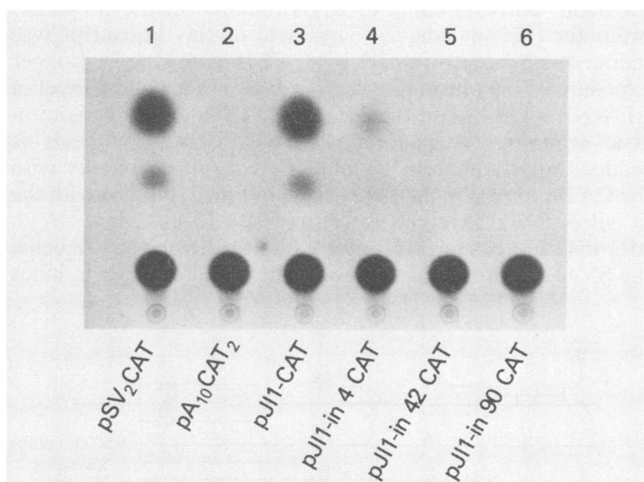



FIG. 2. CAT activity in CV-1 cells transfected with pSV2CAT, pA10CAT2, and pJI1-in(X) CAT constructs. Transfections were performed by the calcium phosphate precipitation method (19), containing 10 μg of each recombinant plasmid and 20 μg of carrier salmon sperm DNA. At 48 h post-transfection, cell extracts were prepared and analyzed according to the protocol of Gorman et al. (18). The percent conversion of chloramphenicol into its acetylated forms are: lane 1, pSV2CAT, 31%; lane 2, pA10CAT2, 0.7%; lane 3, pJI1-CAT, 32%; lane 4, pJI1-in4 CAT, 4%; lane 5, pJI1-in42 CAT, 0.2%; lane 6, pJI1-in90 CAT, 0.4%.

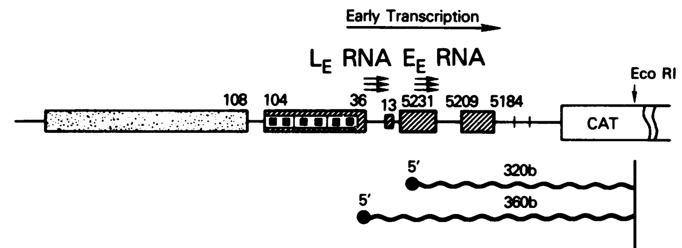



FIG. 3. In vitro transcription of SV40 early promoter insertion mutants. pJI1-CAT, pJI1-in4 CAT, and pJI1-in90 CAT plasmid DNA were restricted with *Eco*RI, purified by phenol-chloroform extraction, and ethanol precipitated. The length of RNA transcripts originating from the SV40 EE and LE RNA initiation sites is represented below the plasmid schematic. Note: The length of LE RNA in plasmid pJI1-in90 CAT may include up to 90 bases of inserted sequence.

provided by Innis and Scott (23). By using standard recombinant DNA techniques, a *Bgl*I-*Sph*I fragment (SV40 map position [mp] 0 to 128) was transferred from each mutant to the parental vector, pSV2CAT.

## RESULTS

**Analysis of the effect of SV40 early promoter insertion mutants on early gene expression in vivo.** We have studied transcription of the early region of SV40 to understand the regulatory mechanisms which control EE and LE RNA synthesis during the lytic cycle. In these studies, we have specifically focused on the effect of increasing the distance between two well-defined transcriptional control domains, the TATA box and the 21-bp repeats. As a first step in assaying transcriptional activity, early-region SV40 promoter mutants with insertions of DNA sequences at the *Nco*I site (mp 37) were fused to the coding sequences of the CAT gene (18). To construct these plasmids, a *Bgl*I-*Sph*I (SV40 mp 0 to 128) fragment of pJI1, -in4, -in42, and -in90 (23) was isolated and inserted into vector pSV2CAT, which had been previously digested with *Sfi*I (SV40 mp 0) and *Sph*I (SV40 mp 128). Figure 1A summarizes the construction and structure of the SV40 mutants used in these analyses. The expanded SV40 early transcriptional regulatory region and the nucleotide sequence of the DNA inserts, as determined by Innis and Scott (23), is presented in Fig. 1B. CV-1 cells were transfected by the calcium phosphate precipitation technique with either the wild-type pJI1-CAT or the three insertion mutants pJI1-in4 CAT, -in42 CAT, and -in90 CAT. The transient expression of CAT was assayed after 48 h. pJI1-CAT, containing one copy of the 72-bp repeat, was as active as the wild-type pSV2CAT, which contains two copies of the enhancer sequence, in inducing CAT enzyme (Fig. 2). The comparison of CAT activity in pJI1-CAT and pSV2CAT extracts also serves as a construction control, demonstrating that the reconstruction of the plasmids did not inherently decrease the expression capacity. Insertion of 4, 42, or 90 bp at SV40 mp 37 results in a dramatic decrease in the level of CAT expression driven by the early promoter. Quantitative comparison of the conversion of chloramphenicol into its acetylated forms showed that CAT activity in pJI1-in4 CAT-transfected cell extracts was six- to eightfold less than the wild-type level of expression (Fig. 2, lanes 1 [31%], 3 [32%], and 4 [4%]). In the extracts prepared from pJI1-in42 CAT and -in90 CAT-transfected cells (Fig. 2, lanes 5 [0.2%] and 6 [0.4%]), the basal level of early CAT activity

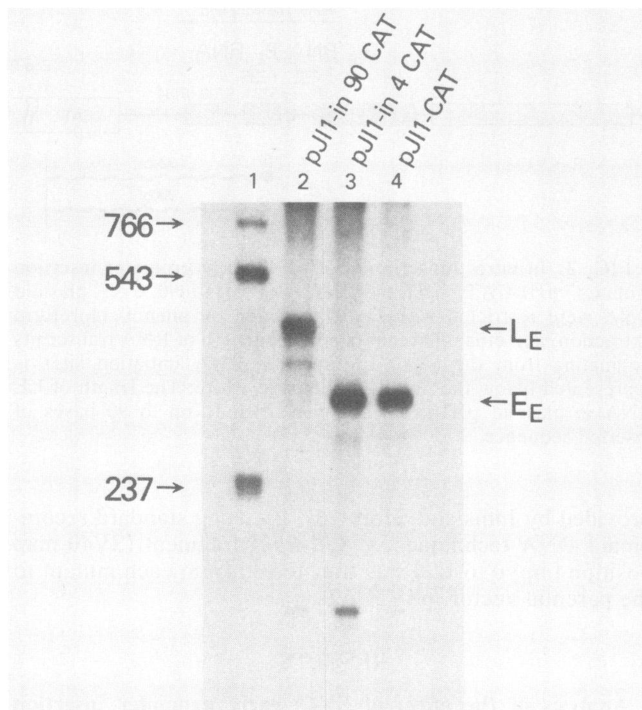



FIG. 4. In vitro transcription of SV40 early promoter insertion mutants. Transcription reactions were done in a final volume of 10 μl, as described in Materials and Methods. Lane 1, SV40 *Hin*fI molecular weight marker; lane 2, pJI1-in90 CAT; lane 3, pJI1-in4 CAT; lane 4, pJI1-CAT. The positions of SV40 EE and LE RNA 5′ ends are indicated by arrows.

was approximately that observed in pA10CAT2-transfected cells (Fig. 2, lane 2 [0.7%]), which was used as a negative control in these analyses. It should be pointed out that similar qualitative and quantitative results have been obtained in multiple transfection assays and with independent DNA preparations. These results suggest that the insertion of minimum lengths of DNA in the SV40 early promoter impairs expression, causing CAT gene expression to drop to a basal level.

**Transcription of SV40 early promoter insertion mutants in vitro.** The experiments described above are an indirect analysis of SV40 early transcriptional activity. We have taken several approaches to directly analyze the transcriptional efficiency and accuracy of RNA initiation of the insertion mutants. The first analysis involves in vitro runoff transcription assays, which allow comparison of promoter activity in a well-controlled system such as the HeLa cell nuclear extract (12). To evaluate SV40 early promoter activity in vitro, template DNAs pJI1-CAT, -in4 CAT, and -in90 CAT were first digested with endonuclease *Eco*RI, which cleaves the plasmid DNA approximately 320 bp downstream of the SV40 EE RNA initiation site (Fig. 3). Purified templates were subsequently transcribed in a HeLa cell nuclear extract. Figure 4 demonstrates the results of a typical in vitro transcription assay. It is clear from the size of the transcript that the majority of RNAs in pJI1-CAT and pJI1-in4 CAT initiate transcription from the SV40 EE RNA initiation site (Fig. 4, lanes 3 and 4). The 4-bp insertion at the *Nco*I site did not change the efficiency or specificity of the SV40 early promoter in vitro. Interestingly, when pJI1-in90 CAT was tested in the nuclear extract, no transcripts were initiated from the EE RNA initiation site (Fig. 4, lane 2). Instead, two transcripts of approximately 350 and 420 bases were detected. The size of the RNA transcripts positions the 5′ end of the minor 350-base RNA at the SV40–insert DNA junction (MP 20 to 30) (Fig. 3). The major 420-base transcript would correspond to an RNA initiated approximately 30 bp downstream of the SV40 21-bp repeats, within the pJI1-in90 CAT insertion sequences. A similar EE to LE shift in RNA initiation sites was obtained with insertion mutant pJI1-in42 CAT (data not shown). The position of these RNA initiation sites correspond to SV40 LE RNAs. This in vitro analysis directly quantitates the efficiency of the SV40 EE and LE promoters and suggests that the interaction between transcription factors binding to control sequences such as the TATA box and upstream sequences (21-bp repeats) are important for efficient EE transcription. When this interaction is disrupted, the SV40 LE promoter is activated to a level which is normally observed for SV40 EE transcription. In contrast to the CAT expression data presented above (Fig. 2), we find very little difference in the level of total early RNA synthesis in the mutants, but do observe a dramatic shift in their 5′ termini in mutant pJI1-in90 CAT (see below).

**Quantitative mapping of SV40 early insertion mutant 5′ ends in vivo.** To determine the level of mRNA which initiated from the SV40 early promoter in CV-1 cells, total RNA was isolated 48 h after transfection and the level of hybrid SV40 CAT RNA was measured by S1 nuclease analysis. Figure 5 diagrams the structure of the 500-base single-stranded probe which was synthesized from a plasmid containing an *Eco*RI-*Bgl*II fragment of pSV2CAT cloned into the M13mp19 vector. The homology of the probe to pJI1-in4 CAT, pJI1-in42 CAT, and pJI1-in90 CAT mRNA extends only to the insertion endpoint (mp 37), so that total RNA will be represented quantitatively in the S1 analysis but not in terms of the position of those 5′ ends which originate upstream of the insertion site. Figure 6 demonstrates that the major RNA synthesized from the pJI1-CAT template transfected into CV-1 cells resulted in the protection of a 316-base DNA fragment, corresponding to RNA initiated at the SV40 EE promoter (Fig. 6, lane 2). The level of this transcript was reduced significantly in pJI1-in4 CAT (Fig. 6, lane 3). Longer exposure of the autoradiogram, in fact, shows a low level of EE RNA synthesis in the pJI1-in4 CAT lane. By densitometric scanning, we estimate that EE RNA synthesis is reduced approximately 10-fold. This result correlates with the CAT expression data presented in Fig. 2, but not with the in vitro RNA levels (Fig. 3 and 4). In pJI1-in4 CAT-, pJI1-in42 CAT, and pJI1-in90 CAT-transfected CV-1 cells, the SV40 EE promoter was completely shutoff (Fig. 6, lanes

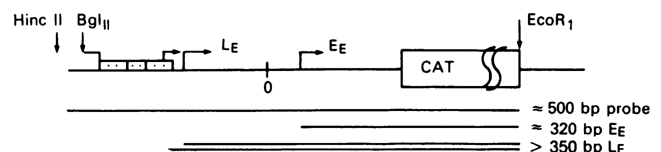



FIG. 5. S1 nuclease analysis of RNA from cells transfected with SV40 early promoter insertion mutants: schematic of S1 probe. An *Eco*RI-to-*Sph*I (converted to *Bgl*II) fragment of DNA from pSV2CAT was cloned into the M13mp19 vector. The DNA strand complimentary to SV40 early RNA was synthesized in the presence of [$^{32}$P]dCTP, restricted with *Hin*cII, heat denatured, and the single-strand probe was isolated following electrophoresis in a 6% acrylamide–urea gel. Hybridization of SV40 EE and LE RNA are expected to protect DNA fragments of approximately 320 or 360 bases, respectively.

3, 4, and 5). However, two major transcripts initiated from upstream sequences were detected. In one group of the RNAs, the 5' terminus was mapped 10 to 20 nucleotides downstream of the insertion endpoint (SV40 mp 20 to 30). The other transcripts originated further upstream and were not mapped accurately by this probe (see below). Except for pJI1-in4 CAT, these data agreed with the in vitro analysis, demonstrating that insertion of DNA between the SV40 EE TATA box and upstream 21-bp repeats results in a shift in the initiation site of RNA synthesis, but does not significantly alter the steady-state level of RNA. Unlike the in vitro LE RNA levels, DNA insertions of 4 bp (pJI1-in4 CAT) significantly reduced the level of EE SV40 RNA in vivo (which correlates with the reduced CAT activity [Fig. 2]). This in vitro transcription discrepancy with the 4-bp insertion mutant may reflect the lack of chromatin structure on the in vitro templates. The differences in the level of total SV40 early CAT mRNA (Fig. 6, cf. lanes 2, 4, and 5) and CAT enzymatic activity (Fig. 2) in pJI1-in42 and pJI1-in90 CAT transfections suggest that additional regulation of the SV40 early mRNA occurs at the translational level (see Discussion).

**Effect of the SV40 72-bp repeat on EE and LE RNA synthesis in vivo.** To further investigate the regulatory elements affecting the level of upstream promoter activity, the 72-bp repeats were deleted from pJI1-CAT and pJI1-in90 CAT. Removal of the enhancer sequence from pJI1-CAT completely abolished EE RNA transcription (Fig. 7, lanes 2 and 3). In pJI1-in90 CAT, however, deletion of the SV40 72-bp repeats showed a different effect on the various RNA species (Fig. 7, cf. lanes 4 and 5). The synthesis of the LE RNAs initiated at mp 20 to 30 (LE0) decreased dramatically (over 15-fold) in the absence of the SV40 72-bp repeats (Fig. 7, lane 5). In contrast, those RNAs starting at a region further upstream, within the G-C-rich repeats (LE1, LE2), were only partially influenced by the absence of the enhancer sequences (Fig. 7, lane 5). This finding suggested that the LE promoters of SV40 are divided into at least two classes. Enhancer-dependent LE start sites are clustered nearer to the origin of replication and enhancer-independent LE initiation sites are located further upstream, closer to the enhancer sequences.

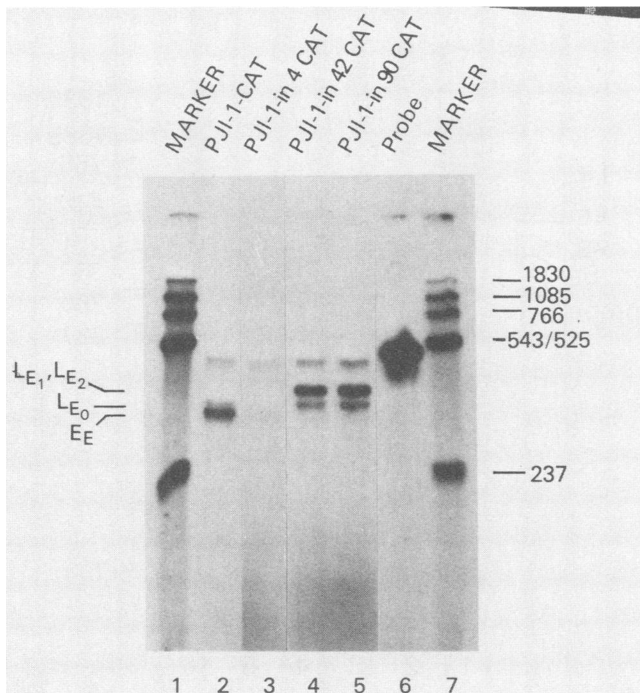



FIG. 6. S1 nuclease analysis of RNA from cells transfected with SV40 early promoter insertion mutants. Transfection of CV-1 cells was performed by the calcium phosphate precipitation method. Each transfection mix contained 10 μg of the recombinant plasmid and 20 μg of carrier salmon sperm DNA. RNA extraction, hybridization, S1 nuclease digestion, and electrophoresis were performed as described previously (5). Lanes 1 and 7, $^{32}$P-end-labeled SV40 *Hin*fI marker DNA; lanes 2, pJI1-CAT; lane 3, pJI1-in4 CAT; lane 4, pJI1-in42 CAT; lane 5, pJI1-in90 CAT; lane 6, control S1 probe.

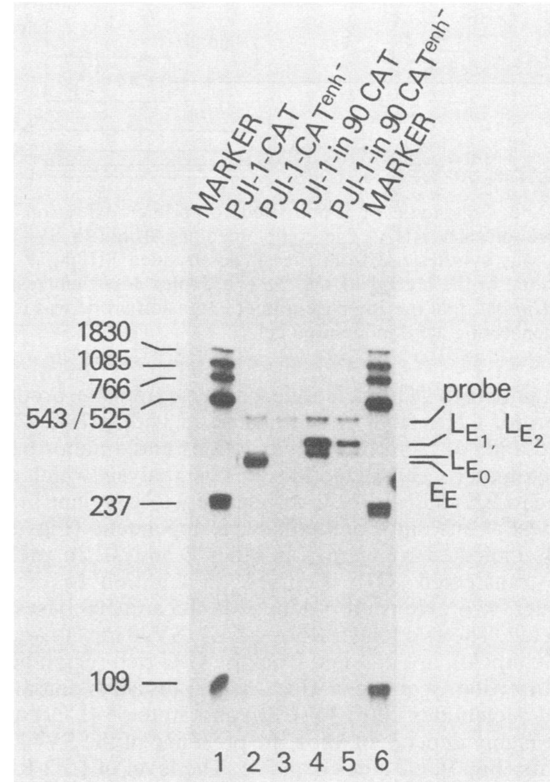



FIG. 7. Effect of deletion of the SV40 72-bp repeats on SV40 EE and LE transcription. The SV40 72-bp repeats were deleted from pJI1-CAT, and pJI1-in90 CAT by restricting the plasmids with single-cut enzymes *Sph*I and *Acc*I. The large DNA fragment was then purified, blunt ended with T4 polymerase, and religated. The resulting plasmids were characterized by restriction enzyme analysis to determine the authenticity of the DNA structure. Transfection, RNA extraction, preparation of single-stranded S1 probe, hybridization, and electrophoresis were carried out as described in the legends to Fig. 5 and 6. Lanes 1 and 6, $^{32}$P-end-labeled SV40 *Hin*fI marker DNA; lane 2, pJI1-CAT; lane 3, pJI1-CAT(enh$^-$); lane 4, pJI1-in90 CAT; lane 5, pJI1-in90 CAT(enh$^-$).

**Mapping of SV40 72-bp dependent and independent EE and LE RNA 5' ends in vivo by primer extension.** To determine more precisely the 5' termini of the early RNAs, primer extension analysis of the in vivo CAT RNA was performed. In this analysis, a probe representing the 265-bp *Eco*RI-*Hin*dIII fragment of pSV2CAT was synthesized from the M13mp19 vector (Fig. 8). This probe, homogeneously labeled with [$^{32}$P]dCTP, was hybridized to cytoplasmic RNA and used as a primer for reverse transcriptase. Elongation of the primer to EE RNA initiation sites should yield a DNA fragment approximately 330 nucleotides in length. For pJI1-in90 CAT-derived LE RNAs, depending upon their exact 5'

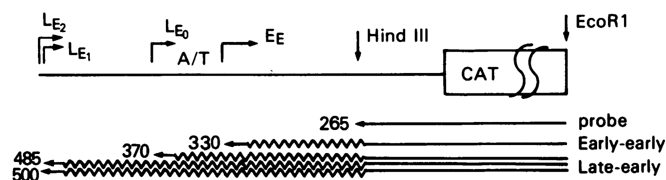



FIG. 8. Schematic of probe used for primer extension 5′-end analysis of early RNA. The complementary strand to SV40 early RNA was synthesized from the single-stranded M13mp19 vector described in the legend to Fig. 5. The duplex was then restricted with *Hin*dIII, and the 265-base primer extension probe was isolated in a denaturing acrylamide-urea gel.

ends, a 370 to 500-nucleotide primer extension product is expected. RNA from cells transfected with pJI1-CAT produced a major band of 330 nucleotides and a minor band of approximately 330 nucleotides in this analysis which correspond to EE and LE RNA, respectively (Fig. 9, lane 2). Both of these transcripts are enhancer dependent (Fig. 9, cf. bands indicated by arrows in lanes 2 and 3). In pJI1-in90 CAT-transfected RNA samples, in addition to the LE0 product, two extension products of 485 and 500 bases were detected. The two longer RNAs, LE1 (SV40 mp 50) and LE2 (SV40 mp 70), are initiated from the G-C-rich region beyond the insertion sequences (Fig. 8). Transcripts initiated at SV40 nucleotides 20 to 30 (LE0) and 50 to 55 (LE1) appear to be highly dependent upon the presence of the SV40 72-bp repeats (Fig. 9, cf. lanes 4 and 5). The level of LE2 RNA is decreased only partially in the 72-bp deletion mutant. This result confirms that the transcripts initiated from the upstream sequences adjacent to the enhancer depend less strongly on the 72-bp repeat element.

## DISCUSSION

The results presented in this manuscript demonstrate that the physical spacing of upstream transcriptional elements has an essential role in the efficiency of SV40 early RNA synthesis. We have specifically demonstrated that insertions of DNA sequences between the 21-bp repeats and the TATA box significantly reduce EE promoter activity in vitro and in vivo. Both of these eucaryotic promoter elements correspond to binding sites for specific transcription factors. Davison et al. (11) have demonstrated that HeLa cell extracts contain a transcription factor that binds to the conalbumin and adenovirus major late TATA box, in the absence of polymerase II, to form stable RNA initiation complexes. Dynan and Tjian (13) have shown that SP1, a transcription factor isolated from uninfected HeLa cells, is required for efficient transcription of the SV40 promoter in vitro and binds specifically to the SV40 21-bp repeats, making contact with their hexanucleotide sequence GG GCGG. Point or deletion mutants in these G-C-rich regions have demonstrated that this interaction is important for EE and LE promoter function in vitro and in vivo (3, 10, 21). Since we obtained similar results in vivo and in vitro (where the 21-bp repeats are the major transcriptional regulatory sequence for SV40 early transcription), the simplest interpretation of our data is that transcriptional factors which recognize the 21-bp repeats need to interact with transcriptional factors bound to other regulatory elements, such as the TATA box, for maximum EE promoter activity. Disruption of this interaction by slight increases in the distance between two major transcriptional elements impairs initiation of EE RNA transcription at the usual start site 25 bases proximal to the TATA box. We suspect that a similar interaction occurs in vivo. Thus it appears that the G-C elements may be the major determinant of transcriptional initiation sites when they are separated from the TATA box. Alternatively, it is possible that the insertion mutants decrease transcription due to blockage of enhancer function.

Two independent studies which also analyzed the effect of spacing between the SV40 early TATA box and upstream 21-bp repeats have recently been published (3, 10). Das and Salzman (10) constructed a set of nine-point deletion or insertion mutants at or near the SV40 *Nco*I site. One mutant, pSVK (+3), which is similar to our pJI1-in4 CAT mutant, decreased early gene expression approximately 10-fold, as measured by CAT enzyme activity. This study, however, did not include an analysis of EE and LE RNA. Barerra-Saldana et al. (3) have reported on a series of insertion mutants analogous to the ones presented in this study. Similar to our results, the authors conclude that the insertion mutants decreased EE RNA synthesis, with a concomitant increase in LE RNA synthesis. By comparison of two sets of insertion mutants, these investigators were able to demon-

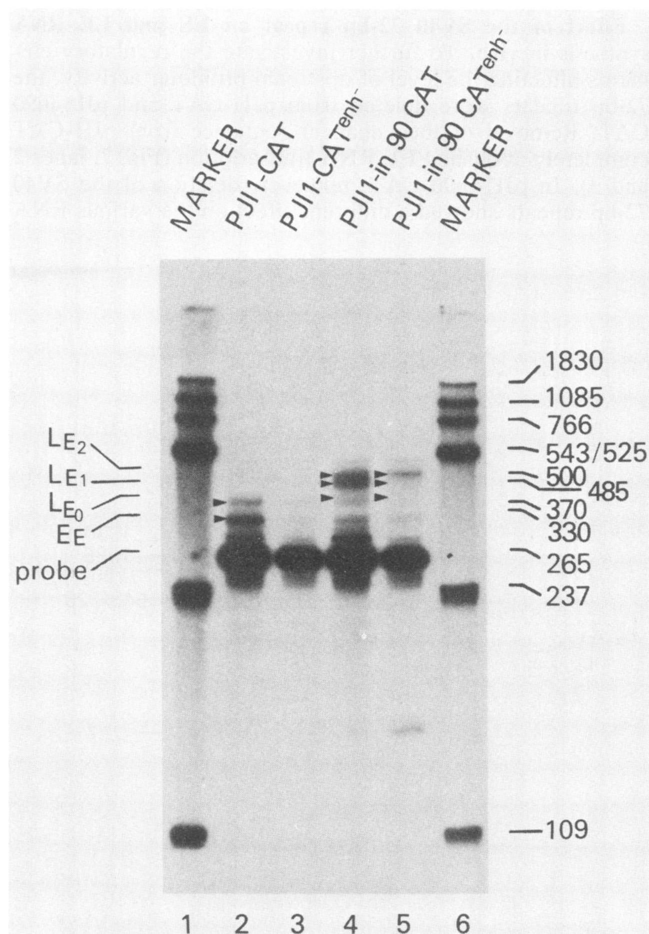



FIG. 9. Primer extension 5′-end analysis of SV40 early RNA. Calcium phosphate transfection, RNA extraction, hybridization, and electrophoresis was carried out as described in Materials and Methods. Primer extension analysis of RNA was performed as described by Hernandez and Keller (22). A 30-μg portion of total cytoplasmic RNA was used for each analysis. Lanes 1 and 6, $^{32}$P-end-labeled SV40 *Hin*fI marker DNA; lane 2, pJI1-CAT; lane 3, pJI1-CAT(enh$^-$); lane 4, pJI1-in90 CAT; lane 5, pJI1-in90 CAT(enh$^-$).

strate that the insertion effect was due to the increased distance between the TATA box and 21-bp repeats and not the TATA box and 72-bp repeats. Our results support this conclusion; in addition, analysis of both RNA 5′ ends and gene expression levels extend previous findings by providing evidence that the switch to LE RNA is accompanied by a decrease in translational efficiency of SV40 T-antigen expression (see below).

Recently, in our studies on the activation of the SV40 late promoter by T antigen, we have shown by in vivo competition studies that efficient binding of *trans*-acting factors requires the presence, in *cis*, of at least two SV40 domains (6). Insertion of increasing lengths of DNA, 42 or 90 bp, between the two domains dramatically reduces their in vivo competition efficiency, suggesting an interaction between the transcriptional factors which bind to the two regulatory regions. A similar conclusion, regarding cooperative interaction of transcriptional factors, was recently reported by Sawadogo and Roeder in their analysis of the binding properties of transcription factors USF and TFIID to the adenovirus major late promoter (31). Dissociation rate measurements suggested a cooperative interaction between USF and TFIID when simultaneously bound to the promoter DNA. Thus, the interaction of transcription factors may be a common mechanism of providing maximum transcriptional activity. It should be noted that the requirement of interaction for inducing RNA synthesis does not preclude a protein from binding to DNA by specific sequence recognition, independent of other transcription factors. In fact, the available data from transcription factors such as SP1, USF, TFIID, etc., would support the concept of independent binding (9, 13, 31).

Concomitant with the decrease in activity of the EE promoter in pJI1-in42 CAT and pJI1-in90 CAT, late early RNA initiating from upstream regions becomes activated. The switch in promoter activity could be due to the new configuration of DNA or chromatin structure making the LE promoter more accessible to the transcriptional machinery or a requirement for initiation adjacent to the G-C-rich sequences. From our analysis, it is not evident that the SV40 LE promoter is weaker than the SV40 EE promoter as suggested previously (2, 3). Based on our in vitro and in vivo studies, under the appropriate conditions the LE promoter is equivalent in activity to the EE promoter. The conditions which stimulate LE transcription also decrease EE transcription, apparently by disrupting the interaction of transcription factors with two domains on the DNA. In this set of experiments, the conditions involve a physical separation of two promoter elements known to associate with DNA-binding proteins. In the SV40 lytic cycle, synthesis of T antigen and concomitant DNA replication lead to a shift in early transcription from the EE to LE promoters. The most likely explanation for this later phenomenon is a physical suppression of EE transcription by T-antigen binding or the structure of newly replicated chromatin or both.

Our mutants which lack the 72-bp repeat sequences demonstrate that the initiation of RNAs from upstream sequences LE0 and LE1 is highly dependent upon the presence of the 72-bp repeats. This observation is in agreement with previous observations of Wasylyk et al. (34). In contrast, the RNA initiated further upstream at LE2 appears to be less dependent on the 72-bp repeats. This result implies that the 72-bp repeats differentially activate the LE promoters. It is of interest that the promoters proximal to the origin of replication are highly dependent on the 72-bp repeats, while the more distal promoters, normally used late in the lytic infection cycle of SV40, are affected to a lesser degree. At this point, it is not clear whether the distance between the promoter and the enhancer sequence is critical to enhancer function; e.g., a minimal distance between the enhancer and promoter sequences may be required to allow enhancer binding factors to interact with downstream regulatory signals. Alternatively, due to the precise set of transcriptional factors required for LE promoter function, this particular RNA initiation site may function in a less enhancer-dependent fashion.

Comparison of the CAT RNA synthesized following transfection of PJI1-CAT, pJI1-in42 CAT, and pJI1-in90 CAT into CV-1 cells with the level of CAT protein, as measured by its enzymatic activity, suggested that additional regulation of the SV40 early gene expression can occur at the translational level. Specifically, our studies suggest that SV40 LE RNA is translated inefficiently compared with EE RNA. In LE RNA, two potential AUG initiation codons are situated 5′ to the coding region of SV40 small-t and large-T antigens. To analyze the effect of the upstream AUGs on translation, in a separate study we have assayed the translational efficiencies of EE and LE RNA transcripts in vitro. The results of these studies indicate that LE RNAs are an inefficient template for translation of T antigens (Khalili et al., in preparation). We suggest, therefore, that the additional AUGs in the leader sequence of LE RNA decrease the translational efficiency of the downstream AUGs coding for early proteins (t/T antigens) in SV40 and the CAT protein in pJI1-in42 CAT and pJI1-in90 CAT.

ACKNOWLEDGMENTS

We express our gratitude to W. Scott for providing the insertion mutant plasmids, to S. Amini, J. Bolen, P. Howley, and R. Weinmann for their helpful suggestions and critical reading of the manuscript and to Janet Duvall for excellent technical assistance.

LITERATURE CITED

1. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27**:299–308.
2. **Barrera-Saldana, H., K. Takahashi, M. Vigneron, A. Wildeman, I. Davidson, and P. Chambon.** 1985. All six GC-motifs of the SV40 early upstream element contribute to promoter activity in vivo and in vitro. EMBO J. **4**:3839–3849.
3. **Baty, D., H. A. Barrera-Saldana, R. D. Everett, M. Vigneron, and P. Chambon.** 1984. Mutational dissection of the 21 bp repeat region of the SV40 early promoter reveals that it contains overlapping elements of the early-early and late-early promoter. Nucleic Acids Res. **12**:915–932.
4. **Benoist, C., and P. Chambon.** 1981. In vivo sequence requirements for the SV40 early promoter region. Nature (London) **290**:304–309.
5. **Brady, J., and G. Khoury.** 1985. *Trans*-activation of the simian virus 40 late transcription unit by T-antigen. Mol. Cell. Biol. **5**:1391–1399.
6. **Brady, J., M. R. Loeken, and G. Khoury.** 1985. Interaction between two transcriptional control sequences required for T-antigen mediated SV40 late gene expression. Proc. Natl. Acad. Sci. USA **82**:7299–7303.
7. **Buchman, A. R., and P. Berg.** 1984. Unusual regulation of simian virus 40 early-region transcription in genomes containing two origins of DNA replication. Mol. Cell. Biol. **4**:1915–1928.
8. **Buchman, A. R., M. Fromm, and P. Berg.** 1984. Complex regulation of simian virus 40 early-region transcription from different overlapping promoters. Mol. Cell. Biol. **4**:1900–1914.
9. **Carthew, R. W., L. A. Chodosh, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**:439–448.
10. **Das, G., and N. P. Salzman.** 1985. Simian virus 40 early

promoter mutations that affect promoter function and autoregulation by large T-antigen. J. Mol. Biol. **182**:229–239.

11. **Davison, B. L., J.-M. Egly, E. Mulvihill, and P. Chambon.** 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature (London) **301**:680–686.
12. **Dignam, J. D., R. M. Lebowitz, 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.
13. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence specific DNA-binding proteins. Nature (London) **316**:774–778.
14. **Everett, R. D., D. Baty, and P. Chambon.** 1983. The repeated G-C rich motifs upstream of the TATA box are important elements of the SV40 early promoter. Nucleic Acids Res. **11**:2447–2464.
15. **Fromm, M., and P. Berg.** 1982. Deletion mapping of DNA region required for SV40 early promoter function in vivo. J. Mol. Appl. Genet. **1**:457–481.
16. **Ghosh, P. K., and P. Lebowitz.** 1981. Simian virus 40 early mRNAs contain multiple 5′ termini upstream and downstream from a Hogness-Goldberg sequence: a shift in 5′ termini during the lytic cycle is mediated by T-antigen. J. Virol. **40**:224–240.
17. **Ghosh, P. K., P. Lebowitz, R. J. Frisque, and Y. Gluzman.** 1981. Identification of a promoter component involved in positioning the 5′ termini of simian virus 40 early RNAs. Proc. Natl. Acad. Sci. USA **78**:100–104.
18. **Gorman, C. M., L. F. Moffat, and B. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
19. **Graham, F. L., and A. J. Van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
20. **Gruss, P., R. Dhar, and G. Khoury.** 1981. Simian virus 40 tandem repeated sequences as an element of the early promoter. Proc. Natl. Acad. Sci. USA **78**:943.
21. **Hansen, U., and P. Sharp.** 1983. Sequences controlling *in vitro* transcription of SV40 promoters. EMBO J. **2**:2293–2303.
22. **Hernandez, N., and W. Keller.** 1983. Splicing of in vitro synthesized RNA precursors in HeLa cell extracts. Cell **35**:89–99.
23. **Innis, J., and W. Scott.** 1984. DNA replication and chromatin structure of simian virus 40 insertion mutants. Mol. Cell. Biol. **4**:1499–1507.
24. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42**:559–572.
25. **Keller, J. M., and J. Alwine.** 1984. Activation of the SV40 late promoter: direct effects in the absence of DNA replication. Cell **36**:381–389.
26. **McKnight, S. L., R. Kingsbury, A. Spence, and M. Smith.** 1984. The distal transcription signals of the herpes virus tk gene share a common hexanucleotide control sequence. Cell **37**:253–262.
27. **Mishoe, H., J. N. Brady, M. Radonovich, and N. P. Salzman.** 1984. Simian virus 40 guanine-cytosine-rich sequences function as independent transcriptional control elements in vitro. Mol. Cell. Biol. **4**:2911–2920.
28. **Moreau, P., R. Hen, B. Wasylyk, R. Everett, M. P. Gaub, and P. Chambon.** 1981. The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res. **9**:6047.
29. **Parker, C. S., and J. Topol.** 1984. A Drosophila RNA polymerase II transcription factor contains a promoter region specific DNA binding activity. Cell **36**:357–369.
30. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33**:741–748.
31. **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**:167–175.
32. **Scholer, H. R., and P. Gruss.** 1984. Specific interaction between enhancer containing molecules and cellular components. Cell **36**:403–411.
33. **Tooze, J. (ed.).** 1980. The molecular biology of tumor viruses, 2nd ed, part 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Wasylyk, B., C. Wasylyk, H. Matthes, M. Wintzerith, and P. Chambon.** 1983. Transcription from the SV40 early-early and late-early overlapping promoters in the absence of DNA replication. EMBO J. **2**:1605–1611.
35. **Wildeman, A. G., P. Sassone-Corsi, T. Grundstrom, M. Zenke, and P. Chambon.** 1984. Stimulation of *in vitro* transcription from the SV40 early promoter by the enhancer involves a specific *trans*-acting factor. EMBO J. **3**:3129–3133.