Characterization of a Minimal Simian Virus 40 Late Promoter: Enhancer Elements in the 72-Base-Pair Repeat Not Required

WILLIAM S. DYNAN^{1*} AND STEPHEN A. CHERVITZ²

Department of Chemistry and Biochemistry¹ and Department of Molecular, Cellular and Developmental Biology,² University of Colorado, Boulder, Colorado 80309-0215

Received 13 July 1988/Accepted 12 November 1988

A 272-base-pair (bp) portion of the simian virus 40 regulatory region containing the replication origin, Sp1-binding region, and part of the 72-bp direct repeats makes up a minimal late promoter that is able to direct late-direction RNA synthesis in vivo and in vitro. Fourteen linker-scan mutants within this region were characterized. Mutations in the Sp1-binding region decreased late expression both in vivo and in vitro. By contrast, mutations that eliminate genetically defined elements of the early transcriptional enhancer or that prevent binding of the transcription factors AP-1, AP-2, and AP-3 in the 72-bp repeat region had little or no effect on late-direction expression. These results argue that, at least under certain circumstances, the early transcriptional enhancer sequences are not required for simian virus 40 late gene expression.

Simian virus 40 (SV40) is a double-stranded DNA virus, the entire nucleotide sequence of which has been determined. There are two transcription units which proceed in opposite directions from a common control region. The late transcription unit, which is ordinarily not activated until after the onset of DNA replication, specifies differentially spliced 16S and 19S mRNAs. The 16S mRNAs, which are more abundant, have predominant 5' termini at nucleotide position 325 (SV numbering system, see Fig. 1), and encode the virion structural protein VP1 and the nonstructural agnoprotein. The 19S mRNAs have heterogeneous 5' termini and encode the virion structural proteins VP2 and VP3 (for a review, see reference 1).

The DNA sequence elements required for transcription of the late mRNA precursors have been extensively investigated. Various studies have indicated a role for the replication origin (8, 11, 27, 28, 35), the early transcriptional enhancer region (8, 15, 21, 27, 28, 35, 40, 40a), sequences 25 base pairs (bp) upstream from the major late cap site (7, 10, 42), sequences at the late RNA cap sites (11, 45, 53), and sequences downstream of the major initiation site at position 325 (5, 19, 46, 53). The Sp1-binding region of SV40 has been shown to enhance late transcription in some studies (9, 18, 20, 25, 51, 57) and to repress late transcription in others (2, 44) and has recently been proposed to be the binding site for a late-specific transcription factor (37). The switch from early to late transcription has been ascribed to repression of early RNA synthesis by the SV40 early-gene product, large T antigen (3, 6, 26, 36, 48-50, 56) and *trans*-induction of late RNA synthesis by large T antigen (7, 8, 34, 35). Changes in the template accompanying the onset of DNA replication and accumulation of high copy numbers of SV40 DNA template molecules may also play a role, however (for discussion, see reference 1). Antitermination (29, 30), selective RNA degradation (6), and specific repression of late RNA synthesis (8, 24) have also been proposed.

In the face of the complex and sometimes contradictory results that have been obtained in the past, we sought a simpler experimental system to identify and study elements of the SV40 late promoter. Ideally, such a system should show initiation of RNA synthesis at only one or a few sites,

rather than the many that are used in wild-type viral infection. Moreover, to facilitate mutational analysis, the template should have the smallest possible amount of functional redundancy. Finally, late RNA should be preferentially transcribed under conditions similar to those found late in infection, after T-antigen synthesis and DNA replication have occurred.

In the studies described here, we examined synthesis of upstream late RNAs with 5' termini near nucleotide position 170 in the early transcriptional enhancer region. Work from several groups has shown that these RNAs are major products of an in vitro reaction. We sought to better characterize the sequences required for their synthesis. Given the strength of this upstream promoter in vitro, we also wished to investigate whether we could construct a minimal late promoter where these upstream RNAs were expressed as major transcripts.

Maps summarizing the constructs and mutants used in these studies are shown in Fig. 1 and 2. We constructed 14 new linker-scan (LS) mutants, which saturate the upstream late control region (nucleotides 65 to 187). The strategy used to obtain these mutants was as follows. Deletion mutants were created by *Bal* 31 digestion of SV40 fragments inserted in the pBRN/B vector (23). These were screened, sequenced, and recombined in vitro to give the LS1 to LS14 series. Additional small deletion mutants (not indicated) were created by in vitro recombination of LS1 and LS4 to give LS1/4, and LS5 and LS7 to give LS5/7. The LS6 × LS14 double mutant (not shown) was created by substituting a restriction fragment bearing the LS14 mutation for the equivalent wild-type fragment in LS6. Selected LS mutants were transferred into the L200 CAT background by substituting restriction fragments bearing the appropriate LS mutations for the equivalent wild-type fragments in L200 CAT.

Linker-scan mutagenesis was used for these studies because the spacing of flanking elements is preserved. In a region containing many promoter elements, this provides a significant advantage. Previous studies of late transcription have, in general, used deletion mutants, and the results therefore may have been influenced by the unintended juxtaposition of control elements flanking the deletions. The LS1 to LS14 series of mutants (Fig. 2, top line) was used for in vitro transcription. The L200 CAT series (second line),

* Corresponding author.

FIG. 1. (A) LS1 to LS14 series of template plasmids. The SV40 segment (dark stippling) extends from the *Hind*III site at SV40 nucleotide 5171 to the *Eco*RI site at 1782 and includes the replication origin, the promoter-enhancer region, and part of the late coding sequences. The direction of late RNA synthesis and the location of the linker-scan substitutions are shown. Plasmid sequences are derived from pBR322 and extend from a novel *Hind*III site at approximately pBR322 nucleotide 2440 to the *Eco*RI site at 4361. (B) L200 CAT template plasmid. The SV40 promoter segment (dark stippling, bottom of circle) extends from the *Hind*III site at SV40 nucleotide 5171 to the *Sph*I site at 200 and includes the replication origin and part of the promoter-enhancer region. An *Sph*I-*Hind*III linker joins the SV40 segment to the *cat* gene. The *cat* expression cassette is the same as in pSV2CAT (22). Plasmid sequences are derived from pBR322 and extend from an *Acc*I site at pBR322 nucleotide 2246 to the *Eco*RI site at 4361. A polylinker is present between the pBR322 and SV40 segments.

which contains SV40 sequences fused to the chloramphenicol acetyltransferase (CAT) gene (*cat*) at nucleotide position 200, was used for in vivo experiments.

In vitro transcription assays. The ability of the mutants LS1 to LS14 to direct late-direction transcription in vitro is shown in Fig. 3. The major band (arrow) mapped to nucleotide 171, as measured by comparison with a dideoxynucleotide sequence ladder (data not shown). The primer used in these experiments detects only the upstream late start sites and does not detect RNA from the major initiation site at nucleotide 325. We know from separate runoff transcription experiments, however, that the nucleotide 171 site shown here accounts for half or more of the total late-direction in vitro transcription (data not shown).

LS1, LS2, LS3, and LS4 each affect one or two GC boxes and are thus expected to lack one or two of the six potential Sp1-binding sites in the region. Each of these mutants show a somewhat greater than 50% decrease in the level of RNA synthesis relative to wild type, suggesting that each GC box contributes to the overall level of activity but that no individual site is essential. In contrast to a previous study with SV40-herpesvirus thymidine kinase fusion promoters, the effect on transcription of the most late-proximal GC box was no greater than the others (20). A recombinant between LS1 and LS4, designated LS1/4, which substitutes 10 bp of linker for 43 bp of SV40 DNA and eliminates all the high-affinity GC boxes in the region showed a strong decrease in transcription. In a separate experiment, 5' deletion mutants were assayed (data not shown). Deletion of upstream sequences to nucleotide position 39 had no effect on transcription, deletion to position 74 caused a small decrease in activity, and deletion to position 85 virtually eliminated transcription. These results show that there is a progressive decline in transcription when GC box sequences are removed and that the two strong Sp1-binding sites at the late-proximal side of the region are the minimum required to support transcription. This is generally consistent with previous studies and demonstrates that Sp1 (or possibly a different protein that binds to the same sequences) has a positive effect on late transcription in vitro (9, 20, 25, 51, 57).

Mutation of a binding site for the transcription factor AP-1

in LS5, LS6, and LS7 gave a decrease in transcription similar to that seen when individual GC boxes were mutated. This effect was reverted by recombining LS5 with LS7, creating a 12-bp deletion that moves the GC box region closer to the transcriptional start. Mutation of a second AP-1 site in LS14 sometimes decreased in vitro transcription, but the effect was usually less than twofold and not as great as in the experiment shown. The double mutant, LS6 $\times$ LS14, which eliminates both AP-1 sites in the 72-bp repeats but leaves the intervening DNA unchanged, gave about the same level of transcription as LS6 alone (data not shown). These results demonstrate that although mutations affecting the AP-1 sites cause a small decrease in transcription, the AP-1 sites are not strictly essential for in vitro transcription in the late direction, and at least in one case the effect of the mutation can be compensated by spacing changes.

The LS9 mutation decreased transcription substantially in an in vitro reaction. This mutation alters a sequence CATA that is about 30 bp from the RNA start site. This sequence may function as a variant TATA box promoter element to increase the level of transcription and to fix the exact position of the start. A similar function for a CATA sequence has been shown in the rat insulinlike growth factor promoter and certain other viral and cellular promoters (16). This role would be consistent with the location of this sequence at -30 to the start in the position normally reserved for a TATA box. It is also consistent with the result in LS5/7, in which the start site remained fixed relative to the CATA sequence but was moved closer to the Sp1 sites because of the deletion. Sequences near the initiation site may also play some role in determining the efficiency of RNA synthesis, as LS10 activates a cryptic start at a position within the linker, and a different construct, pSV01, which has a 3' deletion extending to nucleotide 160, shows essentially no late RNA synthesis (49; W. S. Dynan, unpublished data). The sequence requirements at the initiation site appear to be rather loose, however, and LS12 and LS13, which change the sequences near the initiation site, have essentially wild-type levels of RNA synthesis.

It is interesting that mutations in the known early transcriptional enhancer elements (LS8, LS11, LS12, LS13) (58,

FIG. 3. In vitro RNA synthesis with LS series mutant templates. Preparation of nuclear extracts (12) and in vitro RNA synthesis and analysis by primer extension (13) were done as described previously. (A) The major band (arrowhead) corresponds to RNA synthesized from approximately nucleotide position 170, as determined by comparison with dideoxynucleotide sequence markers in a separate experiment. Lanes are labeled according to the mutant (LS1 to LS14 series) that was used as template. In the lane labeled wt, wild-type late promoter from pSVOL0 was used as template (5). (B) The same RNA was analyzed with a primer that hybridized to viral early RNA. The major bands (arrowheads) correspond to correctly initiated SV40 early RNA.

60) did not have a significant effect on late-direction transcription in vitro. This suggests that the upstream late promoter characterized in these experiments operates independently from the early enhancer, even though the functional elements are interdigitated. In separate experiments (not shown), we were unable to detect an effect of any of the LS1 to LS14 mutants on in vitro RNA synthesis at nucleotide 325, suggesting that the two systems operate independently in vitro. Finally, as expected, in vitro early-direction transcription was not affected by any of the mutations in the 72-bp repeat region.

In vivo expression assays. The level of in vivo gene expression directed by the upstream late promoter was measured with a fusion construct. To concentrate on the upstream start and eliminate the effect of late promoter elements farther downstream, the SV40 regulatory region was truncated at the *Sph*I site at nucleotide 200 and fused with a polylinker to *cat* gene (22). This fusion, L200 CAT, is diagrammed in Fig. 2. The *cat* expression cassette contains a complete eucaryotic transcription unit with a polyadenylation signal.

L200 CAT was introduced into COS-1 cells by the DEAE-dextran transfection method (54), and the levels of CAT enzyme and *cat* RNA were measured after approximately 2 days. Primer extension analysis of the fusion RNA showed several 5' termini (Fig. 4). Two of these (large arrowheads) map to nucleotide positions 167 and 175, which were major

in vitro start sites with this template. Other in vivo start sites (small arrowheads) map to nucleotides 145 and 184, which were minor start sites in the in vitro reaction. This pattern of RNA synthesis is consistent with a model in which the upstream late promoter is active in vivo, but the weak TATA homology (at LS9) has less of a role in specifying the start site, so that minor start sites in the region become relatively more prominent.

The COS-1 cells used for transfection assays produce SV40 T antigen and are permissive for replication. In this background, the L200 CAT fusion construct expressed an average of 10-fold more CAT enzyme activity than a comparable SV40 early promoter fusion construct, pSV2CAT (Fig. 5). In several experiments with BSC-1 monkey kidney cells, which do not express T antigen, the L200 CAT expression level was at least 10-fold lower than that of pSV2CAT and difficult to distinguish from background (E. Tinkle and W. S. Dynan, unpublished data). These data indicate that the minimal late promoter in L200 CAT has retained its late character, such that the ratio of late to early expression in vivo is greater when T antigen is present. It is likely that template replication plays a role in this activation, but the mechanism has not been investigated in detail and we do not yet know, for example, whether there is also replication-independent transactivation of the type described by Keller and Alwine (34). If wild-type SV40 has several independent mechanisms for switching from early to late

FIG. 4. In vivo and in vitro RNA synthesis with L200 CAT fusion template. Transfection was by the DEAE-dextran method (54). Lane M, pBR322-*Hpa*II markers. Lanes 1 to 3, COS-1 cells were transfected with L200 CAT plasmid DNA. After 56 h, cytoplasmic RNA was isolated and analyzed by primer extension with the *cat* primer as described previously (5) and assayed by primer extension with a synthetic 20-mer oligonucleotide complementary to the *cat* gene coding sequence (gift of S. Nordeen, University of Colorado Health Sciences Center, Denver). Lane 1, 5 µg of RNA; lane 2, 15 µg of RNA; lane 3, 25 µg of RNA. Lane 4, RNA synthesized in vitro with a HeLa cell nuclear extract and L200 CAT template and analyzed by primer extension with the *cat* primer. Lanes C, T, A, G, Dideoxynucleotide sequence markers prepared from L200 CAT template with the *cat* primer. Large arrowheads mark the major in vitro RNA species; smaller arrowheads mark other RNA species seen in vivo and in vitro.

transcription, it is possible that only some, not all, of these mechanisms operate with the L200 construct.

Selected LS mutations were transferred into the L200 CAT background and tested for their ability to support *cat* gene expression (Fig. 5). None of the mutations in the 72-bp repeat region had more than a twofold effect on *cat* expression. Individual linker-scan mutations have been shown to block the binding of enhancer factors AP-1, AP-2 and AP-3 (Fig. 1), and another, LS8, is very likely to affect early transcriptional enhancer function, based on previous genetic analysis. The lack of an effect on CAT synthesis strongly suggests that enhancer-binding proteins are not essential for

FIG. 5. In vivo assays of *cat* gene expression. Plasmids were transfected into COS-1 cells and assayed after 48 h for CAT activity as described previously (43). Plasmids used were L200 CAT, mutant derivatives, the SV40 early promoter-*cat* fusion pSV2CAT, and the promoterless construct pSV0CAT (22). All data were normalized to L200 CAT. Stippled bars represent averages of two (LS1/4 CAT, LS6 CAT, LS10 CAT, LS11 CAT, LS12 CAT, LS13 CAT, LS6 × LS14 CAT), three (LS8 CAT, LS9 CAT, LS14 CAT), or five (pSV2CAT, pSV0CAT) independent experiments. Solid bars indicate range of values from independent experiments. Within each experiment, duplicate or quadruplicate plates of cells were transfected; these were generally in very close agreement. The method used for the CAT assay involves the incorporation of label from [³H]acetyl coenzyme A into a benzene-extractable nonvolatile form, with quantitation by liquid scintillation spectrometry (43). In some cases, the procedure was modified by lysing the cells in 0.1% Nonidet P-40 rather than by sonication. All assays were run in the linear range, and the amount of label incorporated was enough that statistical counting error was insignificant (e.g., 300,000 cpm for L200 CAT). Assay background, measured in the absence of chloramphenicol or the absence of cell extract, was always less than 1% of the value obtained with extracts from cells transfected with L200 CAT and was not subtracted.

late gene expression in this background. LS9, which strongly decreased transcription in vitro, had less of an effect in vivo. Another mutation, LS12, also had a small but consistent effect in vivo. This mutant alters binding by AP-2 (41) but also changes sequences at the nucleotide 170 initiation site, which could affect mRNA processing or stability. It is interesting that LS14 gave consistently greater expression than wild type, which indicates that the AP-1 site affected by LS14 is not required for late expression.

The LS1/4 deletion, which lies outside the 72-bp repeats, was the only mutation with an effect that was consistently

greater than twofold. Although the effect was not as great as in vitro, the observed decrease in *cat* expression suggests that Sp1, or possibly another protein that binds the GC boxes, is a positive effector of late-direction transcription. This agrees with a large body of prior in vitro transcription evidence that the Sp1-binding region contributes to late promoter activity but differs from the recent conclusion of two other groups that the Sp1-binding region acts to repress level of late expression in vivo (2, 44).

A minimal late promoter. The present results show that SV40 has an upstream late promoter that is active in vivo and in vitro. The structure of this promoter, which is summarized in Fig. 1, is typical for a promoter transcribed by RNA polymerase II. There is an upstream region containing multiple GC boxes and an AP-1 site. Proteins appear to bind to these sites independently (20; unpublished data), and no single element is essential for transcription. There is also a rather loose requirement for a TATA-like element that is interdigitated with the early transcriptional enhancer elements in the 72-bp repeat region. This element is more important in vitro, where the bulk of the RNA synthesis initiates from a small cluster of sites near nucleotide 170. In vivo, the specificity of initiation is somewhat relaxed, with several different 5' termini present. Correspondingly, mutation of the TATA-like element has less of an effect.

We have not directly compared the activity of pL200 CAT with that of the wild-type late promoter. Such comparisons are not straightforward, because differences in the internal mRNA sequences may lead to differential processing, stability, or translation. An earlier study with a plasmid similar to L200 CAT (pL6A [35]) showed that CAT activity in COS cells was approximately one-fifth of that seen when a nearly full-length late regulatory segment was fused to the *cat* gene. Another rough measure of promoter efficiency is to compare the activity of L200 CAT with that of the early promoter expression construct, pSV2CAT. In COS-1 cells, this ratio was about 10:1, which is only a factor of two less than the ratio of late to early RNA typically seen in wild-type viral infection (1). These ratios based on CAT enzyme activity may actually be low, because the L200 CAT fusion RNA contains an extra AUG start codon within the SV40 sequences, which might decrease the level of *cat* translation.

The apparent lack of a requirement for enhancer sequences for late transcription was interesting in view of several recent studies with somewhat different constructs that have identified specific sequences in the enhancer region that are thought to be required for late transcription (15, 21, 27, 28, 35, 40, 40a, 44, 56). In some cases, the difference in the results is attributable to the different constructs used. For example, the experiments of May et al. (40) measured accumulation of late mRNA body sequences with templates containing the normal late mRNA cap sites. In this case, mutations in "Sph motifs" of the early enhancer (corresponding to LS8 among our mutants) decreased late mRNA synthesis about fivefold. However, RNA synthesis was from a different site and used a much larger late regulatory sequence than the minimal promoter used here.

The present finding that it is possible to construct a minimal late promoter that does not require specific enhancer elements from the 72-bp repeats is consistent with certain prior studies. Viruses lacking almost the entire 72-bp repeat region have been made experimentally and have been shown to produce enough late mRNA to complement late-defective mutants (S-232 and XS13 [18]). A recent, careful study of the XS13 mutant showed that 30 to 100% of the

wild-type levels of late RNA were present 40 h after infection (32). The core elements of the major late promoter, centered around nucleotide 325, were present in the mutant, and it is likely that this site was used for RNA synthesis. It is interesting that because of the deletion in XS13, the GC box region is brought to within about 60 bp of the nucleotide 325 start site. This is approximately the same as distance from the GC box region to the nucleotide 170 start site in wild-type virus. It is known from in vitro studies that the GC box region strongly stimulates transcription from nucleotide 325 when these two regions are fused in close proximity (9, 25, 51, 57), and it is plausible that the same interactions might operate in vivo.

Evolution of late control region. These experiments and ours offer a simple picture of how the current arrangement of late regulatory sequences in wild-type SV40 might have arisen. A more primitive form of the virus might have been missing most of the 72-bp repeat enhancer sequences and was therefore dependent on the minimal late promoter for late mRNA synthesis. This progenitor virus might have used the nucleotide 170 start site, as in L200 CAT, or more likely, the origin and GC box region might have been fused directly to the region of the nucleotide 325 start site, as in XS13.

As a progenitor virus lacking most of the 72-bp repeat region came under selective pressure to express early genes at higher levels and in a variety of cell types, additional enhancer elements might have been gained, separating the upstream late promoter region from the major start and necessitating the development of additional mechanisms to ensure that adequate amounts of late RNA are made. If this view is correct, the upstream initiation site in wild-type SV40 is not an artifact of in vitro transcription, as might have been supposed, but rather the remnant of an earlier transcriptional control system.

This is consistent with the phylogenetic evidence. SV40 is a member of a group of primate viruses that also includes the human viruses BK and JC. These viruses have identical genome organizations and extensive homology in the protein-coding regions but very little sequence similarity in the noncoding regulatory region (17, 52, 59). This pattern is what would be expected if these viruses had developed from a common progenitor by acquisition and amplification of enhancer elements in the intergenic control region.

Moreover, different natural isolates of SV40 and related viruses have different repeat boundaries in the enhancer region. The most commonly used laboratory strain, 776, has 72-bp repeats, but the 800 strain has 85-bp repeats and the Oxman strain has 91-bp repeats (19, 47). Moreover, many new strains with novel repeat patterns have been derived as revertants of viruses with point mutations in the enhancer region (31). These viruses are viable and make late RNA, despite their diversity of sequence organization. This is additional evidence that none of the individual enhancer elements in the 72-bp repeats is critical for late RNA synthesis and suggests that the enhancer region is better viewed as an aggregate of unrelated *cis*-acting elements than as a precisely constructed molecular switch.

We thank A. Siddiqui and S. Jameel, University of Colorado Health Sciences Center, Denver, for the plasmid pPLCAT, which was used as a parent for the L200 CAT series. We thank S. Nordeen, University of Colorado Health Sciences Center, Denver, for the *cat* oligonucleotide. We thank E. Tinkle for carrying out preliminary experiments with *cat* fusion constructs. K. Swaggert, J. Letovsky, and J. Schneringer rendered expert technical assistance, and R.-B. Markowitz provided helpful comments on the manuscript.

This work was supported by Public Health Service grant GM

35866 from the National Institutes of Health. S.A.C. was supported by Public Health Service training grant T32-GM 07135 from the National Institutes of Health.

LITERATURE CITED

- Acheson, N. 1980. Lytic cycle of SV40 and polyoma virus, p. 125-204. In J. Tooze (ed.), *DNA tumor viruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Alwine, J. C., and J. Picardi. 1986. Activity of simian virus 40 late promoter elements in the absence of large T antigen: evidence for repression of late gene expression. *J. Virol.* **60**:400-404.
- Alwine, J. C., S. I. Reed, J. Ferguson, and G. R. Stark. 1975. Properties of T antigens induced by wild-type SV40 and tsA mutants in lytic infection. *Cell* **6**:529-533.
- Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin. 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell* **49**:729-739.
- Ayer, D., and W. S. Dynan. 1988. Simian virus 40 major late promoter: a novel tripartite structure that includes intragenic sequences. *Mol. Cell. Biol.* **8**:2021-2033.
- Birkenmeier, E. H., N. Chiu, M. F. Radonovich, E. May, and N. P. Salzman. 1979. Regulation of simian virus 40 early and late gene transcription without viral DNA replication. *J. Virol.* **29**:983-989.
- Brady, J., J. B. Bolen, M. Radonovich, N. Salzman, and G. Khoury. 1984. Stimulation of simian virus 40 late gene expression by simian virus 40 tumor antigen. *Proc. Natl. Acad. Sci. USA* **81**:2040-2044.
- 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.
- Brady, J., M. Radonovich, M. Thoren, G. Das, and N. P. Salzman. 1984. Simian virus 40 major late promoter: an upstream DNA sequence required for efficient *in vitro* transcription. *Mol. Cell. Biol.* **4**:133-141.
- Brady, J., M. Radonovich, M. Vodkin, V. Natarajan, M. Thoren, G. Das, J. Janik, and N. P. Salzman. 1982. Site-specific base substitutions and deletions that enhance or suppress transcription of the SV40 major late RNA. *Cell* **31**:625-633.
- Contreras, R., D. Gheysen, J. Knowland, A. van de Voorde, and W. Fiers. 1982. Evidence for the direct involvement of DNA replication origin in synthesis of late SV40 RNA. *Nature (London)* **300**:500-504.
- Dynan, W. S. 1987. DNase I footprinting as an assay for mammalian gene regulatory proteins, p. 75-87. In J. K. Setlow (ed.), *Genetic engineering*, vol. 9. Plenum Publishing Corp., New York.
- Dynan, W. S., and R. Tjian. 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* **32**:669-680.
- Dynan, W. S., and R. Tjian. 1983. The promoter-specific factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**:79-87.
- Ernault-Lange, M., F. Omilli, D. R. O'Reilly, and E. May. 1987. Characterization of the simian virus 40 late promoter: relative importance of sequences within the 72-base-pair repeats differs before and after viral DNA replication. *J. Virol.* **61**:167-176.
- Evans, T., T. DeChiara, and A. Efstratiadis. 1988. A promoter of the rat insulin-like growth factor II gene consists of minimal control elements. *J. Mol. Biol.* **199**:61-81.
- Frisque, R., G. L. Bream, and M. T. Cannella. 1984. Human polyomavirus JC virus genome. *J. Virol.* **51**:458-469.
- Fromm, M., and P. Berg. 1982. Deletion mapping of DNA regions required for SV40 early region promoter function *in vivo*. *J. Mol. Appl. Genet.* **1**:457-481.
- Ghosh, P. K., M. Piatak, J. E. Mertz, S. M. Weissman, and P. Lebowitz. 1982. Altered utilization of splice sites and 5' termini in late RNAs produced by leader region mutants of simian virus 40. *J. Virol.* **44**:610-624.
- Gidoni, D., J. T. Kadonaga, H. Barrera-Saldaña, K. Takahashi, P. Chambon, and R. Tjian. 1985. Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. *Science* **230**:511-517.
- Gong, S. S., and K. N. Subramanian. 1988. Functional anatomy of the simian virus 40 late promoter. *Virology* **163**:481-493.
- Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
- Haltiner, M., T. Kempe, and R. Tjian. 1985. A novel strategy for constructing clustered point mutations. *Nucleic Acids Res.* **13**:1015-1025.
- Handa, H., and P. A. Sharp. 1980. Expression of early and late simian virus 40 transcripts in the absence of protein synthesis. *J. Virol.* **34**:592-597.
- Hansen, U., and P. A. Sharp. 1983. Sequences controlling *in vitro* transcription of SV40 promoters. *EMBO J.* **2**:2293-2303.
- Hansen, U., D. G. Tenen, D. M. Livingston, and P. A. Sharp. 1981. T antigen repression of SV40 early transcription from two promoters. *Cell* **27**:603-612.
- Hartzell, S. W., B. J. Byrne, and K. N. Subramanian. 1984. Mapping of the late promoter of simian virus 40. *Proc. Natl. Acad. Sci. USA* **81**:23-27.
- Hartzell, S. W., B. J. Byrne, and K. N. Subramanian. 1984. The simian virus 40 minimal origin and the 72-base-pair repeat are required simultaneously for efficient induction of late gene expression with large tumor antigen. *Proc. Natl. Acad. Sci. USA* **81**:6335-6339.
- Hay, N., and Y. Aloni. 1985. Attenuation of simian virus 40 mRNA synthesis is enhanced by the agnoprotein and is temporally regulated in isolated nuclear systems. *Mol. Cell. Biol.* **5**:1327-1334.
- Hay, N., H. Skolnik-David, and Y. Aloni. 1982. Attenuation in the control of SV40 gene expression. *Cell* **29**:183-193.
- Herr, W., and J. Clarke. 1986. The SV40 enhancer is composed of multiple functional elements that can compensate for one another. *Cell* **45**:461-470.
- Hertz, G. Z., and J. E. Mertz. 1988. The enhancer elements and GGGCGG boxes of SV40 provide similar functions in bidirectionally promoting transcription. *Virology* **163**:579-590.
- Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* **83**:5889-5893.
- Keller, J. M., and J. C. Alwine. 1984. Activation of the SV40 late promoter: direct effect of T antigen in the absence of viral replication. *Cell* **36**:381-389.
- Keller, J. M., and J. C. Alwine. 1985. Analysis of an activatable promoter: sequences in the simian virus 40 late promoter required for T-antigen-mediated trans-activation. *Mol. Cell. Biol.* **5**:1859-1869.
- Khoury, G., and E. May. 1977. Regulation of early and late simian virus 40 transcription: overproduction of early viral RNA in the absence of a functional T-antigen. *J. Virol.* **23**:167-176.
- Kim, C. H., C. Heath, A. Bertuch, and U. Hansen. 1987. Specific stimulation of simian virus 40 late transcription *in vitro* by a cellular factor binding the simian virus 40 21-base-pair repeat promoter element. *Proc. Natl. Acad. Sci. USA* **84**:6025-6029.
- Lee, W., A. Haslinger, M. Karin, and R. Tjian. 1987. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. *Nature (London)* **325**:368-372.
- Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor AP-1 binds TPA-inducible enhancer elements. *Cell* **49**:741-752.
- May, E., F. Omilli, M. Ernault-Lange, M. Zenke, and P. Chambon. 1987. The sequence motifs that are involved in SV40 enhancer function also control SV40 late promoter activity. *Nucleic Acids Res.* **15**:2445-2461.
- Mermod, N., T. J. Williams, and R. Tjian. 1988. Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription *in vitro*. *Nature (London)* **332**:557-561.
- Mitchell, P. J., C. Wang, and R. Tjian. 1987. Positive and negative regulation of transcription *in vitro*: enhancer binding protein AP-2 is inhibited by SV40 T antigen. *Cell* **50**:847-861.

42. Nandi, A., G. Das, and N. P. Salzman. 1985. Characterization of a surrogate TATA box promoter that regulates in vitro transcription of the simian virus 40 major late gene. *Mol. Cell. Biol.* **5**:591-594.
43. Nordeen, S. K., P. H. Green, and D. M. Fowlkes. 1987. A rapid, sensitive, and inexpensive assay for chloramphenicol acetyltransferase. *DNA* **6**:173-178.
44. Omilli, F., M. Ernoult-Lange, J. Borde, and E. May. 1986. Sequences involved in initiation of simian virus 40 late transcription in the absence of T antigen. *Mol. Cell. Biol.* **6**:1875-1885.
45. Piatak, M., P. K. Ghosh, L. C. Norkin, and S. M. Weissman. 1983. Sequences locating the 5' ends of the major simian virus 40 mRNA late forms. *J. Virol.* **48**:503-520.
46. Piatak, M., K. N. Subramanian, P. Roy, and S. M. Weissman. 1981. Late messenger RNA production by viable simian virus 40 mutants with deletions in the leader region. *J. Mol. Biol.* **153**:589-618.
47. Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, and S. M. Weissman. 1978. The genome of simian virus 40. *Science* **200**:494-502.
48. Reed, S. I., G. R. Stark, and J. C. Alwine. 1976. Autoregulation of simian virus 40 gene A by T antigen. *Proc. Natl. Acad. Sci. USA* **73**:3083-3087.
49. Rio, D., A. Robbins, R. Myers, and R. Tjian. 1980. Regulation of simian virus 40 early transcription in vitro by a purified tumor antigen. *Proc. Natl. Acad. Sci. USA* **77**:5706-5710.
50. Rio, D. C., and R. Tjian. 1983. SV40 T antigen binding site mutations that affect autoregulation. *Cell* **32**:1227-1240.
51. Rio, D. C., and R. Tjian. 1984. Multiple control elements involved in the initiation of SV40 late transcription. *J. Mol. Appl. Genet.* **2**:423-435.
52. Seif, I., G. Khoury, and R. Dahr. 1979. The genome of the human papova virus BKV. *Cell* **18**:963-977.
53. Somasekhar, M. B., and J. E. Mertz. 1985. Sequences involved in determining the locations of 5' ends of the late RNAs of simian virus 40. *J. Virol.* **56**:1002-1013.
54. Sompayrac, L. M., and K. J. Danna. 1981. Efficient infection of monkey cells with DNA of simian virus 40. *Proc. Natl. Acad. Sci. USA* **78**:7575-7578.
55. Sturm, R., T. Baumruker, B. R. Franza, and W. Herr. 1987. A 100 Kd HeLa cell octamer binding protein (OBP 100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. *Genes Dev.* **1**:1147-1160.
56. Tegtmeyer, P., M. Schwartz, J. K. Collins, and K. Rundell. 1975. Regulation of tumor antigen synthesis by simian virus 40 gene A. *J. Virol.* **16**:168-178.
57. Vigneron, M., H. A. Barrera-Saldaña, D. Baty, R. E. Everett, and P. Chambon. 1984. Effect of the 21-bp repeat upstream element on in vitro transcription from the early and late SV40 promoters. *EMBO J.* **3**:2373-2382.
58. Weiher, H., M. König, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219**:626-631.
59. Yang, R. C. A., and R. Wu. 1979. BK virus DNA: complete nucleotide sequence of a human tumor virus. *Science* **206**:456-462.
60. Zenke, M., T. Grundstrom, H. Matthes, M. Wintzerith, C. Schatz, A. Wildeman, and P. Chambon. 1986. Multiple sequence motifs are involved in SV40 enhancer function. *EMBO J.* **5**:387-397.