MOLECULAR AND CELLULAR BIOLOGY, Sept. 1984, p. 1900–1914
0270-7306/84/091900-15$02.00/0

Vol. 4, No. 9

Complex Regulation of Simian Virus 40 Early-Region Transcription from Different Overlapping Promoters

ANDREW R. BUCHMAN,† MICHAEL FROMM,‡ AND PAUL BERG*
Department of Biochemistry, Stanford University, Stanford, California 94305

Received 15 March 1984/Accepted 29 May 1984

During simian virus 40 lytic infection there is a shift in initiation sites used to transcribe the early region, which encodes large T and small t antigens. Early in infection, transcription is initiated almost exclusively from sites that are downstream of the origin of DNA replication, whereas transcripts produced later are initiated mainly from sites on the upstream side. We have used mutant virus and specially constructed plasmid DNAs to investigate the factors regulating this transcriptional shift. In our studies simian virus 40 large T antigen appears to mediate the shift in transcription in two ways: first, T antigen represses transcription at the downstream sites late in infection by binding to the region where these RNAs are initiated; second, T antigen promotes transcription from sites on the upstream side by its ability to initiate replication or amplification, or both, of the template DNA. In addition, transcription from the downstream sites is heavily dependent on enhancer sequences located in the 72-base-pair repeat region, whereas transcription from the upstream sites late in infection does not require enhancer sequences. Thus, different overlapping promoters regulate simian virus 40 early-region expression in a manner that apparently coordinates the production of large T antigen with the increase in viral DNA.

At present, the molecular mechanisms controlling gene expression in higher eucaryotes are only poorly understood. One of the best characterized systems which has been used to approach this problem is the early region of simian virus 40 (SV40) (1, 24, 56). The early region codes for two known proteins, large T antigen and small t antigen. Although the function of small t antigen remains obscure, large T antigen is known to be required for the initiation of viral DNA synthesis. These proteins are translated from differentially spliced RNAs which are transcribed, using the same promoter and polyadenylation signals. The SV40 early-region promoter is contained within a set of repeated segments adjacent to the origin of viral DNA replication (*ori*) (Fig. 1). Extensive analysis of this region has demonstrated the presence of three separate elements involved in SV40 early transcription (5, 16). At least one copy of a set of six short guanine-cytosine-rich repeats is necessary for early transcription. In addition, a sequence within the 72-base-pair (bp) repeat has been shown to be an essential element. Finally, a Goldberg-Hogness or TATA box sequence (19; M. Goldberg, Ph.D. thesis, Stanford University, Stanford, Calif., 1978), imbedded within a larger 17-bp adenine-thymine (AT)-rich block, is responsible for positioning the 5′ ends of early RNA.

The SV40 early region is also one of the best characterized examples of regulation mediated by a defined protein. Large T antigen regulates its own synthesis by controlling the level of early-region transcription (33, 53, 57). Furthermore, T antigen binds to SV40 DNA at three adjacent sites overlapping *ori* and the early promoter (Fig. 1). In vitro, site I has the highest affinity for T antigen and site III has the lowest (58, 60). Studies with deletion and point mutations suggest that T-antigen binding at sites I and II is involved in the initiation of DNA replication at *ori* (37, 50, 55). Moreover, studies of transcription in vitro have shown that the binding of T antigen to these sites is correlated with the repression of early-region transcription (25, 49).

Recently, it has been discovered that during the course of lytic infection there is a shift in initiation sites used to transcribe the early region (16, 18, 25). Before the onset of viral DNA replication, transcripts appear with 5′ ends mapping downstream from the center of the 27-bp inverted repeat at *ori*. The positions of the 5′ termini of the downstream RNAs (doRNAs) are determined by a TATA signal within the AT-rich block (Fig. 1) (5, 16, 19). Concommitantly with viral DNA replication, the production of downstream transcripts decreases, a new set of early-region RNAs appear with 5′ ends mapping upstream of the inverted repeat, and upstream RNAs (upRNAs) become the predominant species at late times in infection.

In the present experiments we examine the factors involved in regulating the transcriptional shift from downstream to upstream sites with virus and plasmid DNAs. These studies show that the shift in transcription is mediated by large T antigen. T antigen represses the production of doRNA late in infection, probably by binding to the site where downstream transcription is initiated. The production of upRNAs also depends on T antigen but not on T antigen binding at the promoter. Instead, the increased synthesis of upRNA late in infection results either from a replication intermediate or from amplification of transcription templates.

We have also investigated the dependence of transcription from both the downstream and upstream initiation sites on the 72-bp repeat segment (16). Transcription of doRNA is greatly reduced when the 72-bp repeat region is absent; however, transcription of upRNA is unaffected by the removal of these sequences. Therefore, it appears that transcription of the SV40 early region is mediated by different, yet overlapping, promoters which are regulated by the early-region product, large T antigen.

* Corresponding author.
† Present address: Department of Cell Biology, Stanford University School of Medicine, Stanford, CA 94305.
‡ Present address: Department of Biology, Stanford University, Stanford, CA 94305.

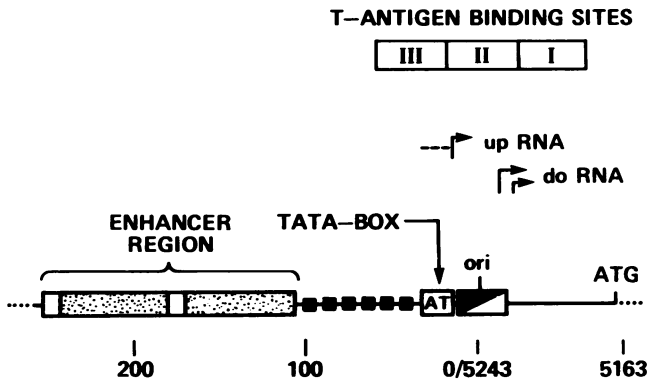


FIG. 1. Diagram of the SV40 segment containing the early-region promoter, the origin of replication, and the large T-antigen binding sites. The rectangle with black and white diagonals, marked *ori*, represents the 27-bp palindrome which is implicated in DNA replication (27). The region immediately adjacent, the 17-bp AT block, has the sequence 5'-TAATTTTTTT*TATTTAT*-3', with the italicized sequence indicating the presumptive Goldberg-Hogness or TATA box. The small black boxes to the left of the AT block represent six repeats of the sequence 5'-PyPyCCGCCC-3'. The partially stippled rectangles further to the left indicate the tandemly repeated 72-bp segments. Unstippled sections of the rectangles show the location of sequences thought to be most responsible for enhancer function (16). The three DNA binding sites for large T antigen and the initiator ATG for the T antigens are also shown. The 5' ends of two sets of early-region transcripts, doRNA and upRNA, are indicated by the arrows pointing rightward. The dashed line adjacent to the arrow representing upRNA indicates the uncertainty in the position of the 5' ends of these RNAs. The nucleotide numbers are those of the SV numbering system (10), which is used throughout.

MATERIALS AND METHODS

Animal cells and viruses. The care and history of the monkey cell line CV-1 and its use for the propagation of SV40 (SVS) and *tsA*58 have been described elsewhere (42). COS cells (21) were grown in Dulbecco modified medium supplemented with 5% newborn calf serum, penicillin, and streptomycin.

Enzymes and reagents. The restriction endonucleases *Eco*RI and *Bam*HI were gifts of J. Carlson and C. Mann (Stanford University), respectively. Other restriction endonucleases used in this study were purchased from New England Biolabs or Bethesda Research Laboratories. T4 DNA polymerase was purchased from P-L Biochemicals. T4 DNA ligase was a gift from S. Scherer (Stanford University). Calf alkaline phosphatase and S1 nuclease were purchased from Boehringer Mannheim Biochemicals. [^{32}P-γ]ATP and T4 polynucleotide kinase were from New England Nuclear Corp. 1-α-D-Arabinofuranosylcytosine (ara-C) was from Sigma Chemical Co.

Transfection of animal cells with plasmid DNA. In transfections performed with plasmid DNA alone, semiconfluent monolayers of either CV-1 or COS cells were treated with 10 μg of plasmid DNA (per 100-mm plate) by a modification by Parker and Stark (51) of the calcium phosphate coprecipitation technique (22).

Coinfection experiments with virus and plasmid DNA were performed as follows. Semiconfluent monolayers of CV-1 cells were first treated with either wild-type SV40 or *tsA*58 at a multiplicity of 10 PFU per cell for 1.5 h at 37°C. The cells were incubated at 37°C for an additional 30 min in normal medium (Dulbecco modified medium plus 5% newborn calf serum) and then transfected with a calcium phosphate precipitate of plasmid DNA as described above.

Isolation and characterization of *tsA*58 and plasmid-derived early RNAs. The methods used for the extraction of cytoplasmic RNA, treatment with DNase I and chromatography on oligo deoxythymidylate-cellulose, have been described in detail elsewhere (61). The positions of the 5' ends of *tsA*58 early RNA were determined by the Weaver and Weissmann modification (63) of the Berk and Sharp S1 nuclease procedure (7) as described previously (16). This involves hybridization of the isolated RNA to an SV40 DNA fragment that is ^{32}P labeled (40) at the 5' end of an *Ava*II restriction site in the early region (nucleotide 5118). The DNA fragment used as probe extends from this *Ava*II site through the SV40 early-region promoter to a *Pvu*II restriction site (nucleotide 270), which is unlabeled. The RNA-DNA hybrids are then digested with S1 nuclease, followed by electrophoresis of the protected DNA fragments (protected fragments) on an 8 M urea–8% polyacrylamide gel (40). The bands visualized by autoradiography of the gel identify fragments whose length measures the distance from the protected ^{32}P-end to a discontinuity in the RNA-DNA pairing. Generally, the discontinuity is at the 5' end of the RNA; however, in the case of upRNA, the RNA-DNA hybrids are digested back to a position at nucleotide 15 due to the S1 nuclease sensitivity of RNA-DNA hybrids ending near or within the 17-bp AT-rich region (16, 25, 39). In this type of analysis the multiple bands indicate fragments differing in length by a single nucleotide, presumably produced by "nibbling" of the terminal residue by S1 nuclease (63). Primer extension experiments (18, 25) and studies reported in the accompanying paper (9) indicate that the 5' ends of upRNA must lie somewhere between nucleotides 15 and 44 and most likely between nucleotides 30 and 34.

The 5' ends of plasmid-derived polyadenylated [poly(A)$^+$] RNAs were determined, using as probe appropriate restriction fragments that were labeled with ^{32}P at one or both 5' ends. Cytoplasmic poly(A)$^+$ RNA was mixed with the ^{32}P-labeled (10^6 cpm/μg) DNA probe and precipitated with ethanol. The precipitate was dissolved in 30 μl of buffer containing 80% formamide, 400 mM sodium chloride, 40 mM PIPES (1,4-piperazine diethanesulfonic acid [pH 6.4]), and 1 mM EDTA, heated at 70°C for 10 min, and then incubated at 49°C for 3 h. The hybridized samples were diluted into 300 μl of cold buffer (0°C) containing 250 mM sodium chloride, 30 mM sodium acetate, 1 mM zinc acetate, 10 μg of denatured salmon sperm DNA per ml, and 3,000 U of S1 nuclease per ml. After incubation at 30°C for 2 h, 10 μg of tRNA was added, and the samples were precipitated by the addition of 2 volumes of ethanol. The products of the S1 nuclease reactions were separated by electrophoresis, either in 1.5% alkaline agarose gels or in 8 M urea–4% polyacrylamide gels. The gels were dried, and autoradiographs were prepared with Kodak XR5 film for 1 to 5 days at −70°C with a Dupont Cronex Lightning-Plus intensifying screen. By using these electrophoretic conditions, the fragments differing by only a few nucleotides that result from nibbling by S1 nuclease are not resolved.

Construction of plasmid DNA molecules. The general procedures used for the separation and purification of DNA fragments, as well as the propagation and purification of plasmid DNA, in *Escherichia coli* have been described previously (16). *E. coli* HB101 (8) was used as the host for all plasmid DNAs. The constructions of several plasmids used in these studies are described elsewhere. These include pSβ-IVS2 (A. R. Buchman and P. Berg, manuscript in prepara-

tion), pSV2-*gpt* (44), and pSV2-*gpt* S-232 (J. Sklar and P. Berg, unpublished data).

(i) Construction of pSβ-*gpt*. In plasmid pSβ-*gpt* the Ecogpt gene (*gpt*) is expressed, using the eucaryotic transcription and processing signals present in pSβ-IVS2 (Fig. 2). Accordingly, pSβ-IVS2 DNA was digested with *Hin*dIII and *Bam*HI endonucleases, and the large 4,345-bp fragment was isolated by gel electrophoresis. A 1.0-kilobase-pair fragment containing *gpt* was obtained by digestion of pSV1GT5-*gpt* plasmid DNA (46) with *Hin*dIII and *Bam*HI endonucleases and isolated by gel electrophoresis. The *Hin*dIII-*Bam*HI pSβ-IVS2 and *gpt* fragments were ligated with T4 DNA ligase, and the products were used to transform *E. coli* HB101. Plasmid DNAs obtained from ampicillin-resistant colonies were analyzed by restriction enzyme digestion to verify the appropriate structure for pSβ-*gpt*.

(ii) Construction of pSβX. Plasmid pSβX is a derivative of pSβ in which the *Bgl*I restriction site at SV40 *ori* has been converted to an *Xho*I restriction site by the insertion of a synthetic *Xho*I octanucleotide linker (Fig. 3). Thus, pSβX lacks SV40 *ori* function but retains the early-promoter and T-antigen binding sites (16, 58). To remove the normal early-promoter region, pSβ DNA was digested with *Pvu*II and *Hin*dIII endonucleases, and the large 4,342-bp fragment was purified by gel electrophoresis. A *Xho*I linker-modified *ori* segment was obtained from pX-8 DNA (16) by digestion with *Pvu*II and *Hin*dIII endonucleases, followed by gel electrophoresis. The 4,342-bp pSβ fragment and the 342-bp pX-8 fragment were ligated with T4 DNA ligase, and the products were used to transform *E. coli* HB101. Ampicillin-resistant transformants were picked, and their plasmid DNAs were analyzed by restriction enzyme digestion to identify the modified pSβ recombinants.

(iii) Construction of pSβXH. pSβXH is a derivative of pSβX which lacks the sequences between the *Xho*I restriction site inserted at *ori* and the *Hin*dIII site just proximal to the β-globin coding region (Fig. 3). Therefore, pSβXH lacks both SV40 *ori* function and the sequences of T-antigen binding sites I and II. To construct pSβXH, pSβX DNA was digested with *Hin*dIII and *Xho*I endonucleases, and the large 4,614-bp fragment was isolated by gel electrophoresis. The 5′-protruding ends of the *Hin*dIII-*Xho*I fragment were converted to blunt ends by incubation with T4 DNA polymerase and the four dNTPs at 50 μM in 10 mM Tris (pH 7.5)–10 mM $MgCl_2$ at 37°C for 10 min. The blunt-end fragment was circularized by ligation with T4 DNA ligase, and the products were used to transform *E. coli* HB101. Plasmid DNAs obtained from ampicillin-resistant colonies were analyzed by restriction enzyme digestion to identify those having the indicated deletion in the early-region promoter.

(iv) Construction of pSβO, pSβXO, and pSβXHO. Plasmids pSβO, pSβXO, and pSβXHO are derivatives of pSβ, pSβX, and pSβXH, respectively, which contain an additional *ori* sequence distant from the β-globin transcription unit. To construct these plasmids, it was convenient to utilize the *ori* segment in the previously described plasmid pS-232 (16). pS-232 is an SV40 recombinant plasmid in which SV40 nucleotides 115 to 345 are missing. The deletion removes the 72-bp repeat region but leaves the sequences necessary for SV40 *ori* function.

pS-232 DNA was digested with *Pst*I endonuclease, and a 2,561-bp fragment containing *ori* was isolated by gel electrophoresis. DNA of each of plasmids pSβ, pSβX, and pSβXH was digested with *Pst*I endonuclease and calf alkaline phosphatase. The 3.7-kilobase-pair fragments which contain the various β-globin expression modules were isolated from

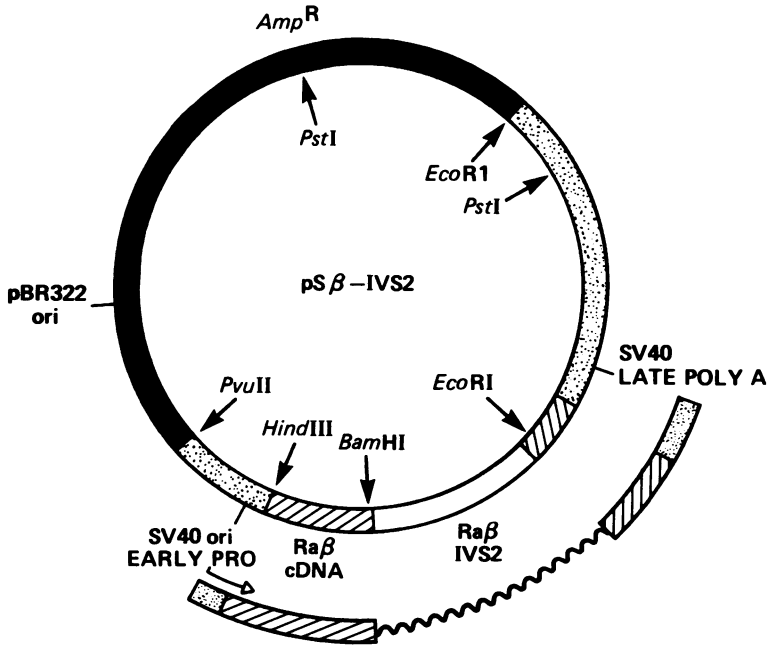


FIG. 2. Structure of plasmid pSβ-IVS2 and plasmid-derived hybrid β-globin RNA. The circular part depicts the general structure of plasmid pSβ-IVS2 (hereafter referred to as pSβ). The shaded segment represents sequences derived from pBR322, and stippled sections are SV40 sequences. The hatched segments are derived from rabbit β-globin cDNA sequences (15, 47), and the unshaded segment represents sequences of the second intron of the rabbit β-globin gene IVS2 (26). The segment containing the SV40 early-region promoter extends from the *Pvu*II site at nucleotide 274 to the *Hin*dIII site at nucleotide 5118. The structure of cytoplasmic RNA produced from the hybrid β-globin transcription unit is shown below the diagram of the plasmid. The wavy region represents the sequences of IVS2 which are removed by RNA splicing. The hatched segments indicate the rabbit β-globin coding region.

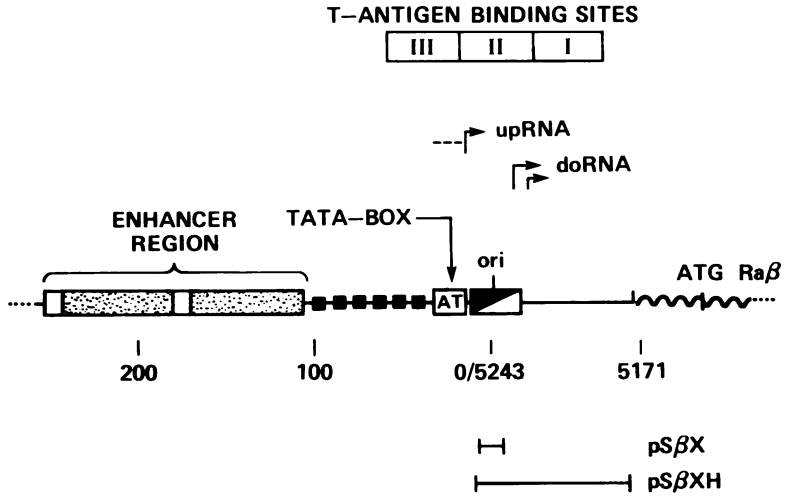


FIG. 3. Structure of the SV40 early-region promoter in plasmids pSβ, pSβX, and pSβXH. The symbols representing the SV40 early-promoter region are described in the legend for Fig. 1. The wavy line represents rabbit β-globin cDNA sequences which are joined to the SV40 early-region promoter at the *Hin*dIII site (nucleotide 5171). The initiator ATG codon for the rabbit β-globin coding region is shown. The solid lines at the bottom of the figure indicate the sequences of the early-region promoter which were deleted in pSβX and pSβXH. In pSβX, 8 bp between nucleotides 5238 and 2 of the SV40 *ori* region are replaced with 8 bp of a synthetic *Xho*I linker. In the construction of pSβXH, the sequences between the inserted *Xho*I linker in pSβX and the *Hin*dIII site were deleted. Effectively, this removes nucleotides 5176 to 2 of the *ori* region, which includes all the sequences of T-antigen binding site I and most of site II.

each sample by agarose gel electrophoresis. Each of the pSβ, pSβX, and pSβXH *Pst*I fragments was ligated individually with the fragment containing *ori* from pS-232 (Fig. 4). The ligated DNAs were used to transform *E. coli* HB101, and the ampicillin-resistant colonies were analyzed by restriction enzyme digestion to identify the plasmids containing the additional SV40 *ori* segment.

RESULTS

Large T antigen regulates the shift in early-region transcription initiation sites. The 5′ ends of doRNAs and upRNAs are located at different positions relative to the known T-antigen binding sites (59, 60) (Fig. 1). The 5′ ends of doRNAs are at the boundaries of sites I and II, whereas the 5′ ends of upRNAs are between sites II and III. Thus, the binding of large T antigen could conceivably regulate both the kind and the amount of these early-region transcripts. To investigate the role of T antigen in the shift in early-region transcription initiation sites, the behavior of the mutant virus *tsA*58, which produces a thermolabile large T-antigen protein, was examined. Because large T antigen is required for the initiation of viral DNA synthesis, viral DNA replication does not occur when cells infected with *tsA*58 are incubated at the nonpermissive temperature (41°C).

*tsA*58 early-region RNAs were isolated at various times after infection and annealed to an SV40 DNA fragment that was 5′ ^{32}P labeled at an *Ava*II restriction site in the early region. After digestion of the RNA-DNA hybrids with S1 nuclease, the products were separated by electrophoresis on an 8 M urea–8% polyacrylamide gel (Fig. 5). The numbers at the right in Fig. 5 correspond to the nucleotide position (10) at which the RNA-DNA hybrids were cleaved by S1 nuclease. In cells infected with *tsA*58 at the permissive temperature (31°C), the pattern of 5′ ends for early-region transcripts was similar to that with wild-type virus infection (16). Thus, most of the RNA appearing early in infection has 5′ ends that are clustered at nucleotides 5233 and 5237 (doRNAs), whereas at later times, the majority of the RNAs appear to be initiated upstream of nucleotide 15 (upRNAs) (see above for a discussion of the mapping of upRNA 5′ ends). In a number of experiments, the relative ratio of upRNAs to doRNAs rises from 0.2 at 24 h to 3 at 86 h. However, during infections with *tsA*58 at the nonpermissive temperature (41°C), only low levels of upRNAs were produced at a time (48 h) when substantial quantities of doRNAs were formed. Also, upRNAs were not produced in infections at the permissive temperature (31°C) which were treated continuously with ara-C, an inhibitor of viral DNA replication (38). However, the addition of ara-C late in infection (from 72 to 86 h) did not cause the disappearance of the accumulated upRNAs. These results indicate that the large T antigen and viral DNA replication are needed for the appearance of upRNAs late in infection.

It is possible that upRNA formation depends on large T-antigen binding at the doRNA initiation site, on the ability of T antigen to initiate viral DNA replication, or on both. Alternatively, T antigen might exert an indirect effect, such as causing RNA polymerase to initiate transcription preferentially at the upstream site. The mechanism by which T antigen induces upRNA production is dealt with more directly in experiments utilizing SV40 plasmid recombinants, which are discussed below.

Since T antigen is known to repress early-region transcription, we sought to determine whether doRNA transcription and upRNA transcription are repressed equally or differentially. For this purpose, cells infected with *tsA*58 were incubated at the permissive temperature until late in infection (72 or 86 h) and then shifted to the nonpermissive temperature for 7 or 14 h. In both experiments there was a 10-fold increase in the formation of total early-region RNAs after the temperature shift; however, although upRNA was the predominant species before the temperature shift, doRNA was the more prevalent species afterward. Thus, large T antigen seems to repress doRNA production to a greater extent than upRNA synthesis. This may stem from

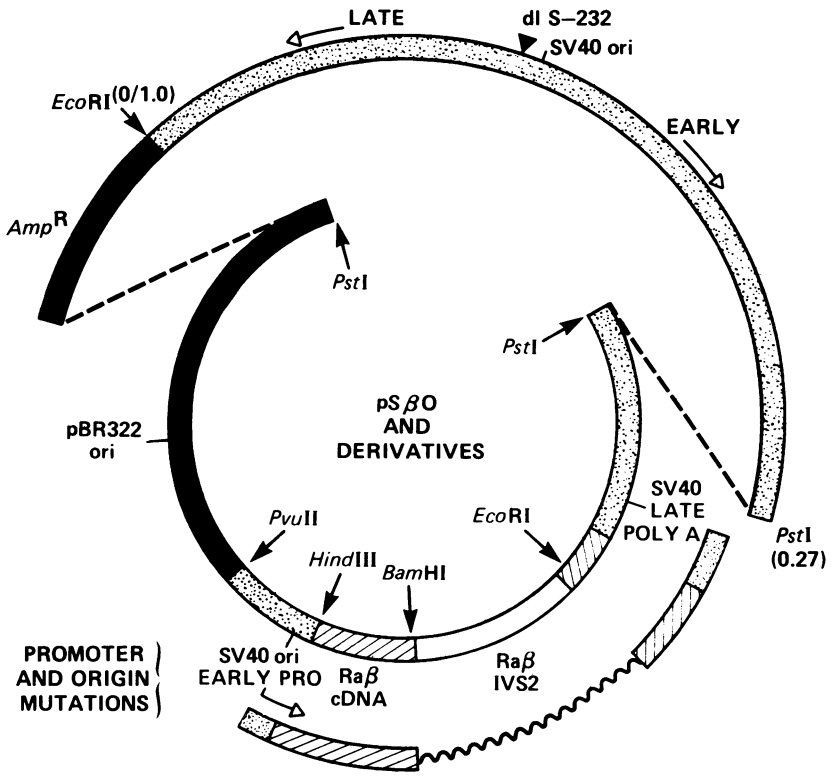


FIG. 4. Structure of plasmid pSβO and its derivatives. The construction of pSβO and its derivatives pSβXO and pSβXHO are described in the text. The symbols used to illustrate the structures of these plasmids are described in the legend of Fig. 2. The two *Pst*I restriction fragments used to construct each of these plasmids are indicated, joined by the dashed lines. The bottom section of the figure depicts the structures of the 3.7-kilobase-pair *Pst*I fragments derived from pSβ, pSβX, or pSβXH. The top section of the figure represents the structure of the 4.3-kilobase-pair *Pst*I fragment derived from pS-232 (16). The locations of SV40 *ori*, the S-232 deletion, and the orientation of the early and late region are indicated on the pS-232 fragment.

the fact that the doRNA initiation sites are closer to site I, the highest-affinity T-antigen binding site (Fig. 1).

Plasmid vector system for studying the shift in initiation sites of early-region transcription. To investigate the mechanism by which large T antigen increases the formation of upRNAs it was important to use a system in which the production of large T antigen is not dependent on transcription from the SV40 early-region promoter that is to be manipulated and studied. For this purpose, we have used the plasmid vector pSβ-IVS2 (Buchman and Berg, manuscript in preparation) (hereafter referred to as pSβ) (Fig. 2). The SV40 early-region promoter in pSβ transcribes a rabbit β-globin gene instead of the normal T-antigen coding sequence. Monkey cells transfected with pSβ plasmid DNA produce substantial quantities of β-globin RNA initiated at the SV40 early-region promoter. The 5′ ends of these RNAs can be located by using homologous DNA probes that are 5′ ^{32}P labeled at restriction sites within the rabbit β-globin sequence.

To assess the role of T antigen in early-region transcription, both CV-1 and COS cells were used as hosts for transfections with plasmid DNA. CV-1 is the normal monkey fibroblast cell line used as host for SV40 infections. COS is a derivative of CV-1 that expresses large T antigen constitutively from an SV40 early region that is stably integrated in the cellular genome (21). Since pSβ contains SV40 *ori*, the plasmid replicates after transfection into COS cells.

The 5′ ends of β-globin RNAs produced in CV-1 and COS cells after transfection with pSβ were mapped with an *Ava*II restriction fragment of pSβ DNA labeled at both 5′ ends with ^{32}P as probe (Fig. 6). One end of the probe is the *Ava*II restriction site within the β-globin region of pSβ, and the other end is an *Ava*II site within the pBR322 portion of the plasmid. Digestion of this probe with *Bgl*I endonuclease generated a convenient marker fragment, 225 bases in length, that extends from the center of the inverted repeat at *ori* to the *Ava*II site in the β-globin gene. The expected sizes for the S1 nuclease protected fragments of upRNA and doRNA from pSβ are 238 and 210 bases, respectively. The poly(A)$^+$ RNA made after transfection of CV-1 cells with pSβ was primarily doRNA. After transfection of COS cells with pSβ, about 20-fold more β-globin RNA was produced. This is probably the result of the increase in template copy number caused by replication of the plasmid in COS cells. Both upRNA and doRNA were detected in roughly equal amounts in COS cells transfected with pSβ. Therefore, as in the infection with *tsA*58, transcription in the absence of T antigen initiates primarily at the downstream site; but with T antigen present, transcription is increased from the upstream site.

In the analyses of RNA produced in COS cells, there was also a minor fragment of 150 bases protected from S1 nuclease digestion. This fragment arises in COS cells because replication of the template promotes transcription from the late-region promoter, thereby producing a transcript which protects the end-labeled late strand in the probe. The presence of the 150-base fragment was useful in

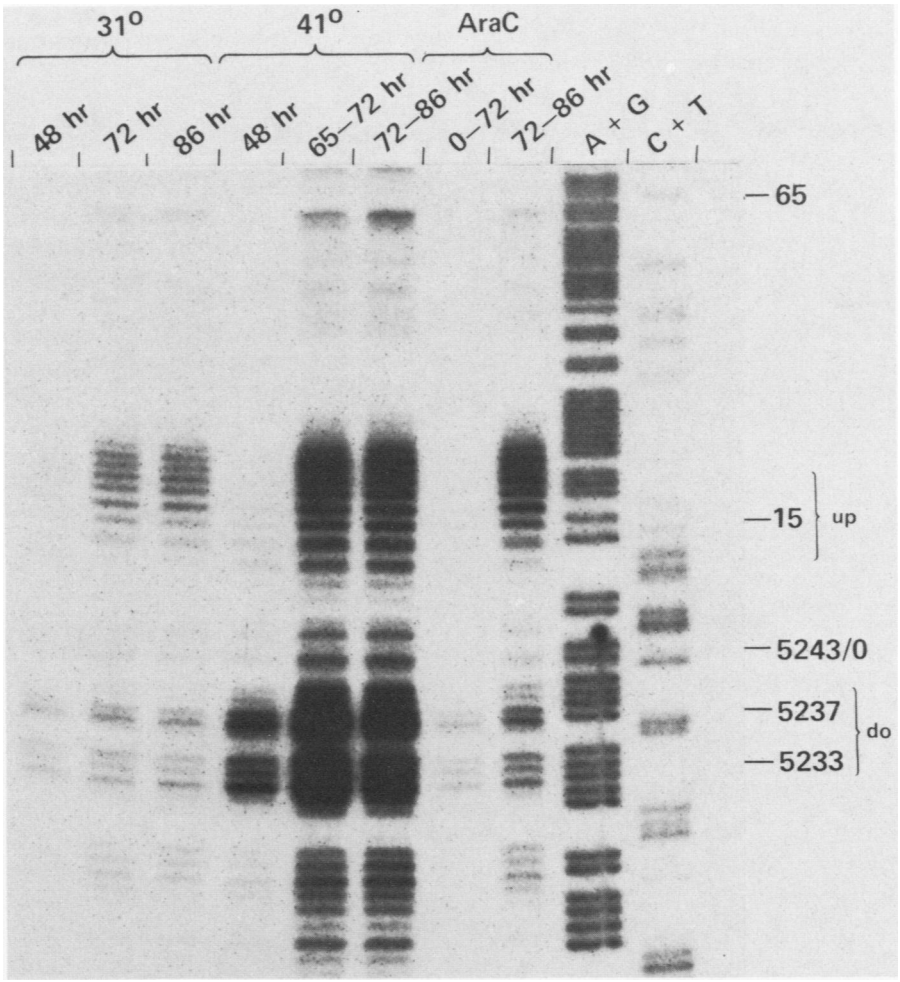
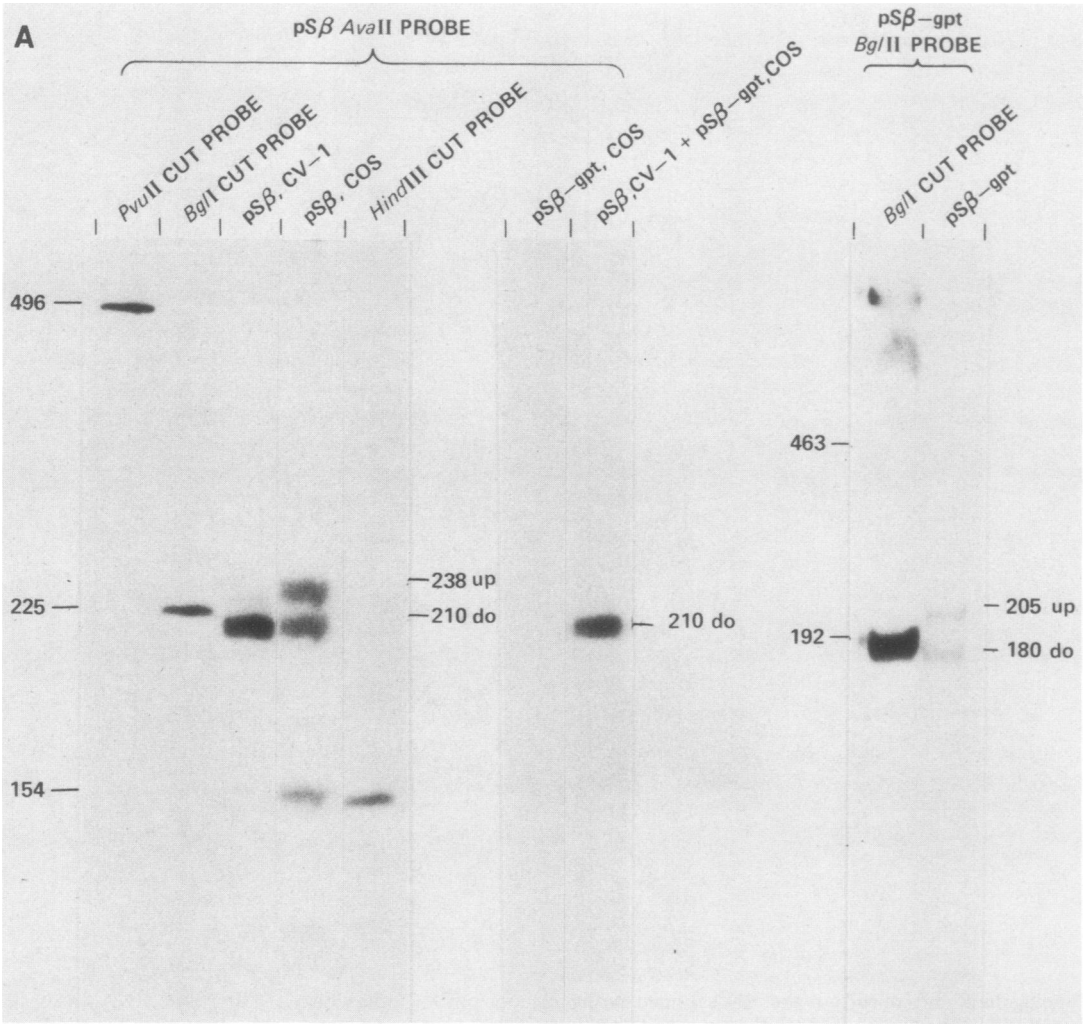


FIG. 5. Mapping the 5′ ends of early-region RNA formed during infection of CV-1 cells with *tsA*58. Poly(A)$^+$ RNA was isolated from CV-1 cells infected at 10 PFU per cell with *tsA*58 at the times and conditions indicated in the headings. Hybridization of the RNA to an SV40 DNA fragment labeled with ^{32}P at the 5′ end of an *Ava*II restriction site (nucleotide 5118) was followed by S1 nuclease digestion, electrophoresis of the protected fragments on an 8% sequencing gel, and autoradiography as described in the text. The adenine-plus-guanine and cytosine-plus-thymine sequencing ladders, obtained by applying the Maxam-Gilbert sequencing procedure (40) to the same labeled fragment, provide the size markers. The numbers shown in the right margin are in the SV numbering system (10) as diagrammed in Fig. 1. The lanes below the 31°C heading are analyses of RNA isolated at the indicated times from infections incubated continuously at 31°C. The 48-h lane under the 41°C heading is with RNA isolated at 48 h after continuous incubation at 41°C. The 65–72 hr and 72–86 hr lanes under the 41°C heading are with RNA in which the infections were incubated first at 31°C and then shifted to 41°C for the indicated time span before the extraction of RNA. The lanes under the AraC heading are analyses of RNA isolated from *tsA*58 infections incubated at 31°C and exposed to ara-C (30 μg/ml) for the time periods indicated. The up and do labels in the right margin indicate the positions of the protected fragments generated by the upstream and downstream early-region RNAs, respectively.

certain experiments for evaluating relative transfection efficiencies, but it has no bearing on transcription from the SV40 early-region promoter and is not discussed further.

We have considered the possibility that the fragment corresponding to upRNA produced by pSβ transfection of COS cells arises by hybridization of doRNA from pSβ and upRNA transcribed from the integrated SV40 early region to the same strand of the end-labeled probe. This possibility was eliminated in a control experiment in which plasmid pSβ-*gpt*, which contains an *Hin*dIII-*Bam*HI fragment of the *gpt* gene (44) instead of the *Hin*dIII-*Bam*HI β-globin region of pSβ, was used to transfect COS cells. The RNAs transcribed from pSβ-*gpt* were measured specifically with a homologous DNA fragment that was 5′ ^{32}P labeled at the *Bgl*II restriction site in the *gpt* segment of pSβ-*gpt*. As expected, fragments corresponding to upRNA and doRNA were detected in COS cells transfected with pSβ-*gpt* (right panel of Fig. 6). However, a mixture of RNA from pSβ-*gpt*-transfected COS cells and pSβ-transfected CV-1 cells did not produce protected fragments corresponding to upRNA with the β-globin-specific *Ava*II probe. Therefore, the fragments corresponding to upRNA from plasmid transfections of COS cells are the result of authentic initiation at the upstream site of the plasmid-derived early-region promoter.

Since the experiments with *tsA*58-infected cells and ara-C treatment indicated that DNA replication is important for the production of upRNA, this dependence was reexamined in COS cells transfected with pSβ. The S1 nuclease mapping of RNA isolated from pSβ-transfected COS cells, both in the presence and absence of ara-C, is shown in Fig. 7. The DNA fragment used as probe for these experiments was 5′ ^{32}P labeled at the *Bam*HI site in the β-globin cDNA segment.



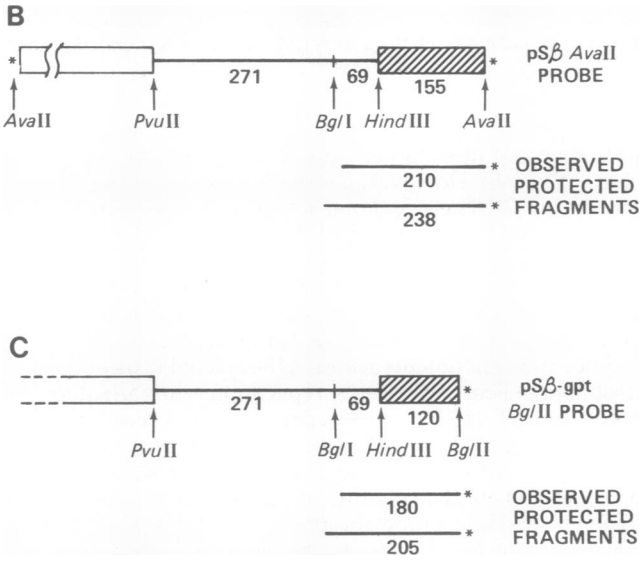


FIG. 6. Mapping the 5′ ends of RNA formed in transfections of CV-1 and COS cells with plasmids pSβ and pSβ-*gpt*. (A) Poly(A)$^+$ RNA was isolated from CV-1 and COS cells transfected with pSβ or pSβ-*gpt* plasmid DNA and subjected to S1 nuclease analysis. The DNA probes for these analyses are indicated in the headings, and the structures of the probe fragments are depicted in B and C. The products of the S1 nuclease digestions were separated by electrophoresis in 8 M urea–4% polyacrylamide gels. RNA from 5×10^5 transfected COS cells and 5×10^6 transfected CV-1 cells were used. The marker fragments labeled *Pvu*II, *Bgl*I, and *Hin*dIII are derived from the appropriate probe by digestion with these restriction endonucleases. The numbers on the left of each panel show the calculated sizes of the corresponding marker fragments. Marker 463 indicates the position of a fragment of 463 bases (data not shown). The numbers to the right of each panel indicate the sizes of the corresponding protected fragments. The up and do labels indicate the protected fragments corresponding to upRNA and doRNA, respectively. (B) The pSβ DNA probe was prepared from a 1,936-bp *Ava*II restriction fragment. Both 5′ ends of the probe contain ^{32}P label, as indicated with asterisks. The hatched segment represents rabbit β-globin cDNA sequences. The open box region represents pBR322 sequences, and the section indicated by the solid line represents the SV40 early-promoter sequences. (C) The pSβ-*gpt* *Bgl*II probe was a DNA fragment from pSβ-*gpt* labeled with ^{32}P at the 5′ end at the *Bgl*II restriction site within the *gpt* sequence. The probe fragment extends from this *Bgl*II site to an *Eco*RI restriction site within the pBR322 section of the plasmid, which is unlabeled. The hatched region represents *gpt* sequences. The unshaded section indicates pBR322 sequences, and the solid line represents the promoter region.

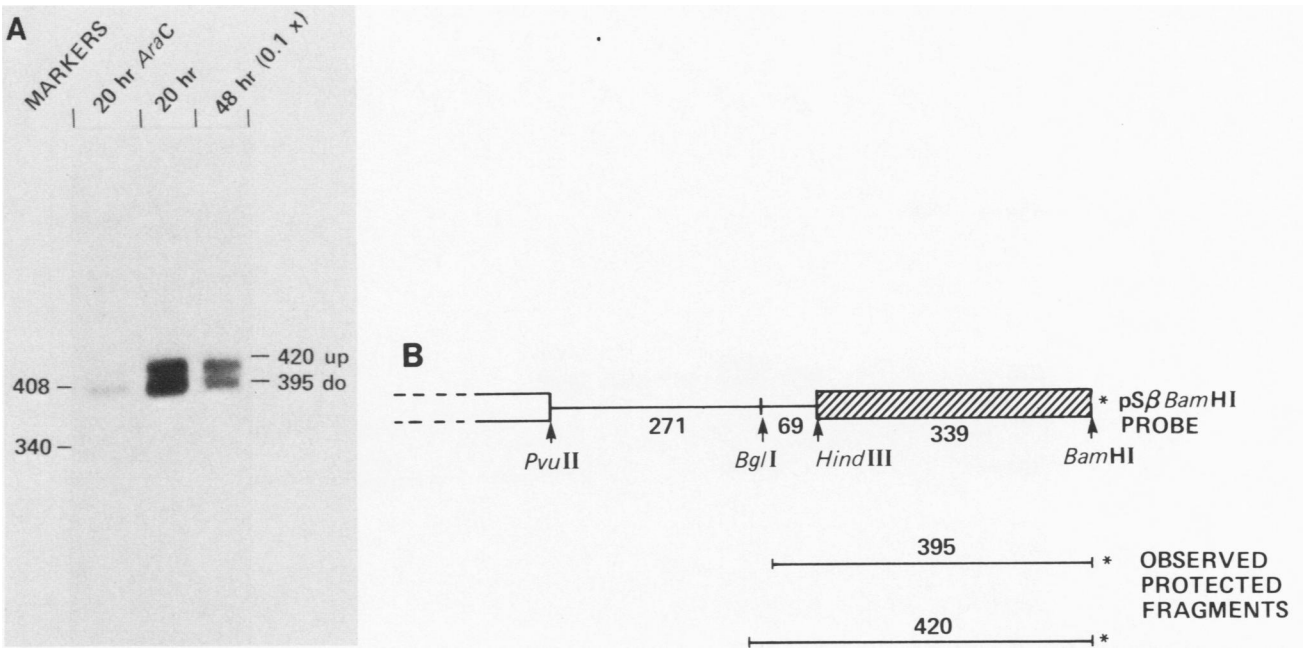


FIG. 7. Mapping the 5′ ends of RNA formed in transfections of COS cells with pSβ in the presence of ara-C. (A) Cytoplasmic poly(A)$^+$ RNA was isolated from COS cells transfected with pSβ DNA and subjected to S1 nuclease analysis with the pSβ *Bam*HI probe described in Fig. 7B. The products were separated by electrophoresis on a 1.5% alkaline agarose gel. The analyses were performed with RNA from 5×10^6 COS cells transfected for 20 h with and without ara-C treatment (30 μg/ml) or from 5×10^5 COS cells transfected for 48 h (indicated by the 0.1 × label). Coordinates 408 and 340 in the left margin indicate the positions of *Bgl*I and *Hin*dIII endonuclease-cut probe, respectively (data not shown). The coordinates in the right margin indicate the calculated sizes of the observed protected fragments. The up and do labels indicate the protected fragments derived from upRNA and doRNA, respectively. (B) The pSβ *Bam*HI probe was prepared by ^{32}P labeling a pSβ DNA fragment at the 5′ end of a *Bam*HI restriction site within the rabbit β-globin sequence. The probe fragment extends from this *Bam*HI site to an *Eco*RI restriction site within the pBR322 section of the plasmid, which is unlabeled. The hatched box represents rabbit β-globin cDNA sequences. The open box indicates pBR322 sequences, and the solid line represents the promoter region.

Equal amounts of both upRNA and doRNA were present 20 and 48 h after transfection in the absence of ara-C. In contrast, RNA isolated 20 h after transfection in the presence of ara-C was predominantly doRNA. Therefore, as with the virus infections, ara-C inhibited the formation of upRNA. However, ara-C treatment also reduced the total production of β-globin RNA about 15-fold. In other experiments (data not shown), ara-C did not inhibit transcription in COS cells transfected with pSβ derivatives which lack a functional *ori*. Thus, it appears that transcription templates are inactivated when they replicate in the presence of ara-C. The fact that ara-C is incorporated into replicating DNA and also causes premature chain termination during DNA replication (23) could explain this result. In any case, because ara-C treatment has this effect, it is not possible to define the role of large T antigen in upRNA production from these experiments alone.

Effects of *ori* and T-antigen binding site sequences on transcription from the early-region promoter. To determine whether T-antigen binding at the promoter region or T antigen-induced viral DNA replication is responsible for the increase in upRNA production, two derivatives of pSβ, which contain alterations at *ori* and the T-antigen binding sites, were examined. One derivative, pSβX, lacks a functional SV40 *ori* but still contains the T-antigen binding sites (Fig. 3). The other plasmid, pSβXH, lacks both *ori* and the binding sites for T antigen. In pSβX, 8 bp at the *Bgl*I restriction site (nucleotide 5235) at *ori* are replaced by 8 bp of a synthetic DNA linker that contains an *Xho*I restriction site. The insertion of the *Xho*I linker sequence inactivates *ori* function (16) but does not prevent T-antigen binding to the site II region in vitro (D. C. Rio and R. Tjian, personal communication). Similar alterations within the inverted repeat region have been shown to inactivate *ori* without affecting the ability of T antigen to bind in vitro to the altered site II (58). Thus, pSβX retains the high-affinity T-antigen binding sequences at site I as well as the lower-affinity binding at site II. Plasmid pSβXH lacks the sequences between the inserted *Xho*I restriction site in pSβX and the *Hin*dIII site at the junction of SV40 early-region and rabbit β-globin cDNA sequences. pSβXH, therefore, lacks all of the nucleotides comprising T-antigen binding site I and most of T-antigen binding site II.

The consequences of these deletions on the sites of transcription initiation were determined by mapping the 5′ ends of RNA produced in transfections of CV-1 or COS cells with plasmids pSβ, pSβX, or pSβXH (Fig. 8). The probes used in these assays were 5′-^{32}P-labeled *Ava*II restriction fragments derived from the homologous plasmids. There was only a 3-fold difference in the amount of β-globin RNA produced in COS versus CV-1 cells by both pSβX and pSβXH, in contrast with the 20-fold difference observed with pSβ. This very likely reflects the fact that plasmids pSβX and pSβXH do not replicate in COS cells for lack of a functional *ori*. More important, transfections with pSβX and pSβXH produced predominantly doRNA in both CV-1 and COS cells, and there is no difference in the ratio of doRNA to upRNA in COS cells compared with that in CV-1 cells.

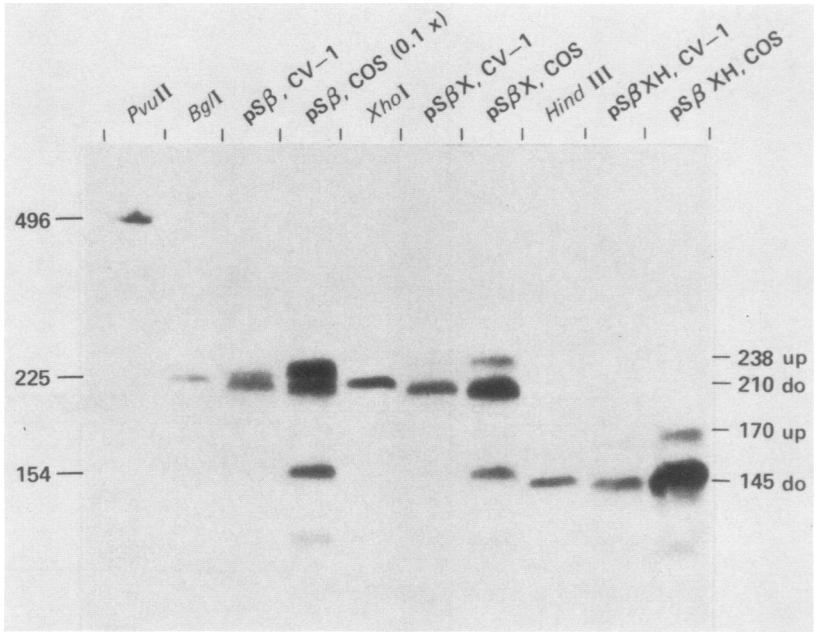


FIG. 8. Mapping the 5′ ends of RNA formed in transfections of CV-1 and COS cells with plasmids pSβ, pSβX, and pSβXH. Cytoplasmic poly(A)$^+$. RNA was isolated from CV-1 and COS cells after transfection with pSβ, pSβX, or pSβXH plasmid DNA and subjected to S1 nuclease analysis with homologous 5′-^{32}P-labeled *Ava*II restriction fragments as probe (described in Fig. 6B). The products of the S1 nuclease reactions were separated by electrophoresis in an 8 M urea–4% polyacrylamide gel. The markers indicated as *Pvu*II, *Bgl*I, and *Hin*dIII were derived from the pSβ *Ava*II probe by digestion with these restriction enzymes; also shown is a fragment produced by *Xho*I endonuclease digestion of the pSβX *Ava*II probe. The numbers in the left margin indicate the sizes of the corresponding marker fragments. All the analyses were with RNA from 5×10^6 transfected cells, except the pSβ-COS sample which contained RNA from 5×10^5 transfected cells (indicated by the 0.1 × label). The numbers in the right margin indicate the calculated sizes of the observed protected fragments. The up and do labels indicate the protected fragments derived from upRNA and doRNA, respectively.

Recall that in transfections with pSβ, upRNA transcription was specifically increased in COS cells. These findings indicate that the mere presence of T antigen is not sufficient to induce transcription from the upstream site. Furthermore, since plasmid pSβX showed no increased upRNA production in COS cells, we infer from the studies of T-antigen binding that the association of T antigen at site I is not sufficient to induce transcription of upRNA.

Replication of template DNA is sufficient to increase the formation or upRNA. To test whether template DNA replication is sufficient to increase upRNA formation, functional SV40 *ori* segments were inserted outside the β-globin transcription units of pSβ, pSβX, and pSβXH to generate the derivatives designated pSβO, pSβXO, pSβXHO, respectively (Fig. 4). These three plasmids contain functional *ori* sequences, and, therefore, they replicate in COS cells, despite the alterations at the early-region promoter transcribing the β-globin gene. COS cells were transfected with pSβ, pSβO, pSβXO, or pSβXHO, and the RNAs were analyzed by S1 nuclease mapping with 5′-^{32}P-labeled *Ava*II restriction fragments as probes (Fig. 9). Note that pSβO produced the same amount of both doRNA and upRNA as did pSβ, demonstrating that the additional *ori* segment did not affect transcription of the β-globin region from a normal early-region promoter. The significant result was that COS cells transfected with either pSβXO or pSβXHO produced the same amount of upRNA as did those transfected with pSβ or pSβO. By contrast, comparable transfections with pSβX and pSβXH, which lack a functional *ori*, showed no increase in transcription at the upstream site. Thus, the presence of a functional *ori* is sufficient to induce normal levels of upRNA even in the absence of T-antigen binding sites at the early-region promoter. These experiments indicate that the functional *ori* need not coincide with the site of transcription initiation. They also suggest that the effect of replication may be a consequence of the ensuing amplification of the DNA.

T antigen represses the production of doRNA from plasmids containing the SV40 early-region promoter. In the experiments described thus far, we assumed that large T antigen represses transcription from the SV40 early-region promoter on plasmids in the same manner that it represses transcription during normal virus infection. There is considerable evidence that T antigen mediates the repression of early-region transcription by binding to DNA at the site where transcription is initiated (25, 49). Since transcription initiation at the downstream site is the most sensitive to repression by T antigen, we anticipated that plasmids which lack the T-antigen binding sites (pSβXH and pSβXHO) would produce more doRNA in transfected COS cells than would the pairs of plasmids which retain the T-antigen binding sites (pSβ and pSβO as well as pSβX and pSβXO). However, plasmid pSβXHO produced only marginally more doRNA in transfected COS cells than did pSβ, pSβO, or pSβXO. This may be a consequence of the lower level of large T antigen produced in COS cells compared with the amount present late in SV40 lytic infection (29).

To validate the conclusions drawn in the preceding sections, the question of T-antigen-mediated repression was examined by using coinfections of CV-1 cells with plasmid DNA and either wild-type (SVS) or *tsA*58 virus. Temperature-shift experiments were performed as described above to

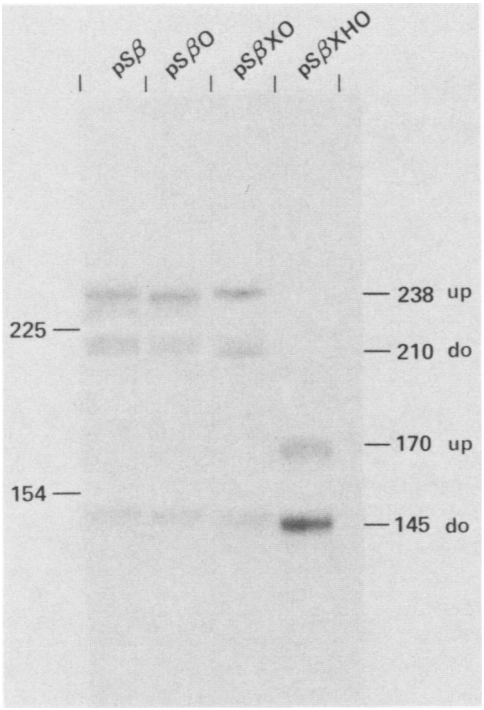


FIG. 9. Mapping the 5′ ends of RNA formed after the transfection of COS cells with plasmids pSβ, pSβO, pSβXO, and pSβXHO. Cytoplasmic poly(A)$^+$ RNA was isolated from COS cells transfected with pSβ, pSβO, pSβXO, or pSβXHO DNA and subjected to S1 nuclease analysis with homologous 5′-^{32}P-labeled *Ava*II restriction fragments as probe (shown in Fig. 6B). The products of the S1 nuclease reactions were separated by electrophoresis in an 8 M urea–4% polyacrylamide gel. The analyses were performed with RNA from 5×10^5 transfected cells. Numbers in the left margin indicate the sizes and positions of markers obtained by *Bgl*I and *Hin*dIII endonuclease cleavage of the pSβ *Ava*II probe (data not shown; Fig. 6). Numbers in the right margin indicate the calculated sizes of the corresponding protected fragments. The up and do labels indicate the protected fragments derived from the upRNA and doRNA, respectively.

determine whether the production of plasmid-derived doRNA is repressed by T antigen. CV-1 cells were infected first with either SVS or *tsA*58 virus and subsequently transfected with pSβO, pSβXO, or pSβXHO plasmid DNA. The cells were incubated at 31°C for 72 h and then either maintained at 31°C or shifted to 41°C for an additional 7 h. Cytoplasmic poly(A)$^+$ RNA was isolated from each coinfection and analyzed by S1 nuclease mapping by using homologous DNA probes ^{32}P labeled at the 5′ terminus of the *Bam*HI restriction site in the β-globin sequence (Fig. 10). The results show that T antigen represses the transcription of doRNA from the unmodified early-region promoter in pSβO. After the shift in temperature fourfold more doRNA was present in coinfections of pSβO and *tsA*58 than in those without the temperature shift, but there was no overproduction of doRNA in the coinfections of wild-type SV40 (SVS) with pSβO with or without a shift in temperature. Similarly, T antigen also represses the production of doRNA from pSβXO, despite the insertion of the *Xho*I linker sequence at site II. However, transcription from the downstream site in pSβXHO is not repressed by T antigen. doRNA was overproduced compared with upRNA in the coinfections of pSβXHO and wild-type SV40, and there was no greater overproduction of doRNA in pSβXHO-*tsA*58 coinfections after the shift in temperature. These results, therefore, are consistent with the prevailing model that repression of early-region transcription is mediated by T-antigen binding at the early promoter. Furthermore, since coinfections of pSβXHO with SVS or *tsA*58 produced the same amount of upRNA as did coinfections of pSβO and pSβXO, the high-affinity T-antigen binding sites required to repress the production of doRNA are not required for the induction of upRNA.

The 72-bp repeat segment is required for the production of doRNA but not upRNA in transfected COS cells. Deletion analysis of the SV40 early-region promoter has revealed that the 72-bp repeat region contains an enhancer of early-region transcription (3, 16). The enhancer function is manifested even when the 72-bp repeat segment is in different positions and orientations relative to the promoter (17). Moreover, the 72-bp repeat region increases transcription from other eucaryotic promoters present in the same DNA (3, 43). Considering the unusual properties exhibited by the enhancer, it was of interest to investigate its role in regulating both doRNA and upRNA production. This was examined by using plasmids pSV2-*gpt* (44–46) and pSV2-*gpt* S-232 (Sklar and Berg, unpublished data) shown in Fig. 11. Although the eucaryotic transcription units of pSβ and pSV2-*gpt* are slightly different, the functional organization of these plasmids is analogous. Most important, pSβ and pSV2-*gpt* contain the same SV40 early-region promoter and *ori* segment. pSV2-*gpt* S-232 is a derivative of pSV2-*gpt* which lacks the 72-bp repeat sequence that supplies enhancer function (16).

CV-1 and COS cells were transfected with plasmids pSV2-*gpt* and pSV2-*gpt* S-232. Cytoplasmic poly(A)$^+$ RNA was isolated from each transfection and analyzed with S1 nuclease by using a DNA probe that was ^{32}P labeled at the 5′ end of the *Bgl*II restriction site in the *gpt* region (Fig. 12). As anticipated, primarily doRNA was produced in transfections of CV-1 cells with pSV2-*gpt*, and both doRNA and upRNA were produced in transfected COS cells. However, in transfections of either CV-1 or COS cells with pSV2-*gpt* S-232, there was a marked reduction in the amount of doRNA, whereas the formation of upRNA in COS cells was unaffected. Thus, although transcription from the downstream site is greatly dependent on the enhancer function of the 72-bp repeat region, these sequences are dispensable for transcription at the upstream site.

DISCUSSION

Mutant virus and specially designed plasmid DNAs have been used to examine the complex regulatory system which governs early-region expression during infection of monkey cells with SV40. In particular, our experiments have focused on the temporal shift in the initiation sites of early-region transcription. Early in the infection, transcription is initiated almost exclusively from doRNAs, whereas transcripts produced later are initiated mainly from upRNAs (16, 18, 25) (Fig. 1).

Large T antigen has already been implicated in the regulation of the amount of early-region transcription during SV40 infection (33, 53, 57). The generally accepted model (16, 25, 49) is that binding of T antigen to site I and probably site II prevents RNA polymerase binding or initiation of transcription or both, thereby repressing early RNA formation. But T-antigen binding at site II also initiates DNA replication at *ori* and, as a consequence, causes increased transcription from further upstream.

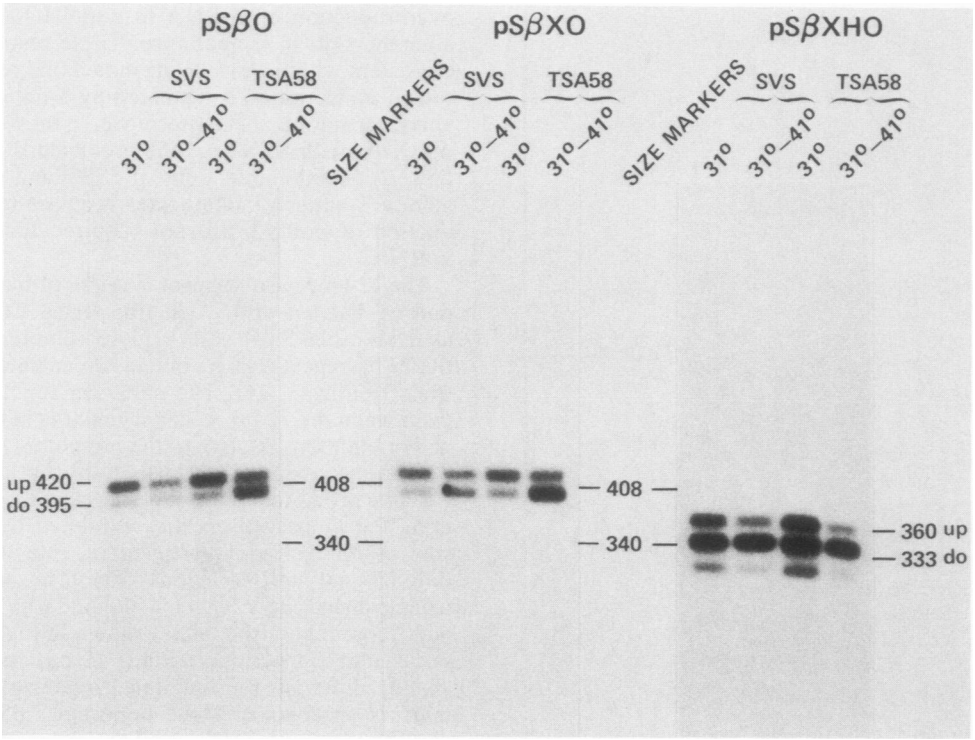


FIG. 10. Mapping the 5′ ends of RNA produced after coinfection of CV1 cells with SVS or *tsA58* virus and plasmids pSβO, pSβXO, or pSβXHO. CV-1 cells were coinfected with either SVS or *tsA58* virus and pSβ, pSβO, or pSβXHO plasmid DNA as indicated in the headings. Infections were either incubated continuously at 31°C for 79 h or incubated at 31°C for 72 h and then shifted to 41°C for an additional 7 h as shown in the labels above each lane. Poly(A)$^+$ cytoplasmic RNA was isolated from each coinfection and subjected to S1 nuclease analysis with homologous 5′-^{32}P-labeled *Bam*HI restriction fragments as probe (Fig. 7B). The products were separated by electrophoresis on a 1.5% alkaline agarose gel. Each analysis was performed with RNA from 5×10^5 cells. Numbers between the panels indicate the positions of markers produced by *Bgl*I (408 bases) and *Hin*dIII (340 bases) endonuclease digestion of the pSβ probe (data not shown). Numbers in the left and right margins indicate the calculated sizes of the corresponding protected fragments. The up and do labels indicate the fragments referred to in the text as upRNA and doRNA, respectively.

How does T antigen induce the formation of upRNA and regulate the relative amounts of doRNA and upRNA? Our experiments establish that transcription of upRNA is not induced by the presence of T antigen alone or by binding of T antigen to sites I and II in the absence of DNA replication; only amplification of the template, a condition that requires functional T antigen and an intact *ori* sequence, appears to be needed for production of upRNA; moreover, replication does not have to be initiated at the promoter itself to increase upRNA production. Another distinctive difference in the formation of the two species of early-region RNA is that transcription of doRNA is strongly dependent upon a function supplied by the enhancer region, whereas upRNA formation does not require enhancer function. Overall, our results indicate that the SV40 early region is transcribed from overlapping, yet distinctly different, promoters: the downstream and upstream promoters.

Downstream promoter and enhancer function. Transcription from the downstream promoter occurs early in infection before viral DNA replication begins and is dependent on functions supplied by enhancer sequences within the 72-bp repeat region. Quite possibly, the function of the downstream promoter is to facilitate efficient production of doRNA when the viral DNA copy number is low. The accumulation of doRNA leads to the generation of enough early-region proteins, particularly large T antigen, to allow viral DNA replication to begin. As the level of T antigen rises, the binding of T antigen to sites I and II (*ori*) represses transcription of doRNA and concommitantly induces viral DNA replication.

It is not known how the enhancer mediates its effect on transcription from the downstream site. An unusual feature of its function is that the enhancer segment can augment early-region transcription from various distances and in different orientations relative to the promoter (17). Furthermore, the enhancer element is not specific for the SV40 early region since it increases transcription when it is placed *cis* to other promoter sequences, including the herpes simplex thymidine kinase (11; M. Fromm, Ph.D. thesis, Stanford University, Stanford, Calif., 1982), human β-globin (30; Sklar and Berg, unpublished data), rabbit β-globin (3), maize *adh* (L. Dennis and P. Berg, unpublished data), chicken conalbumin and ovalbumin (43), and the adenovirus type 2 late promoters (43). It has been proposed (16) that the enhancer activates transcription at promoters by controlling or opening the chromatin structure; support for this model stems from the finding that the enhancer segment confers DNase I hypersensitivity to that sequence in chromatin (17). Another idea is that the enhancer provides an RNA polymerase binding or entry site for the promoter (3, 43). However the enhancer acts, it is notable that transcription from the downstream site is strongly dependent on enhancer function, whereas transcription from the upstream site is not, at least once DNA replication has begun. Wasylyk et al. (62) have

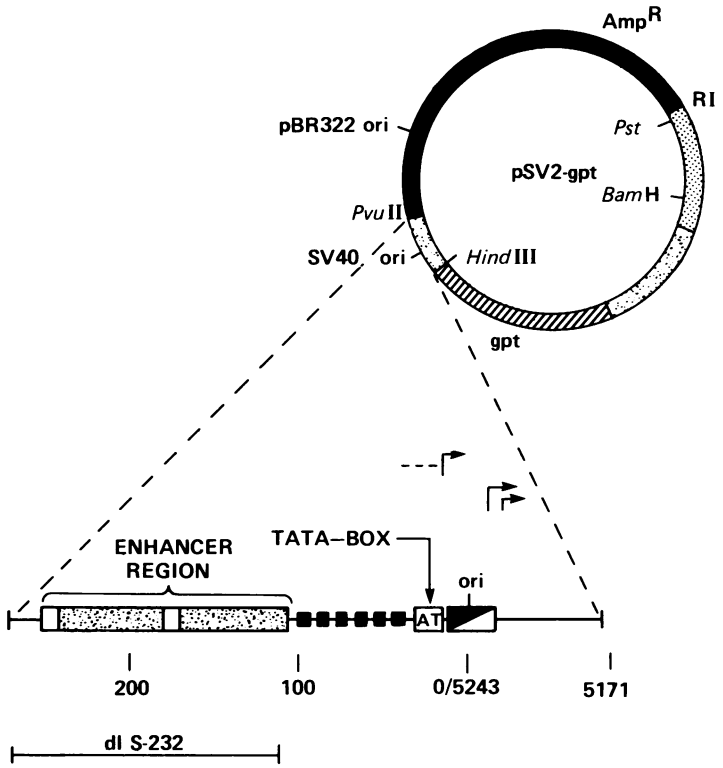


FIG. 11. Diagram of the structures of plasmid pSV2-*gpt* and the derivative pSV2-*gpt* S-232. The circular part of the figure depicts the general structure of plasmid pSV2-*gpt* (44). The shaded portion of this diagram indicates the pBR322-derived portion of the plasmid, stippled segments show those areas derived from SV40 sequences, and the hatched section indicates the sequences derived from *gpt*. The bottom part of the figure depicts an enlarged view of the early-promoter region of pSV2-*gpt*, with symbols as those described in Fig. 1. The solid line below the promoter region diagram shows the sequences of the 72-bp repeat that are deleted in the derivative plasmid pSV2-*gpt* S-232.

reported that the 72-bp repeat is important for the production of upRNA from nonreplicating templates in HeLa cells. This suggests the possibility that the enhancer requirement for the transcription of upRNA changes after the templates have replicated or become amplified.

Enhancer activity has been detected in DNA segments of other animal virus genomes, for example, Moloney murine sarcoma virus (35), adenovirus (28), polyoma (13), hepatitis B (S. Subramani and P. Berg, unpublished data), and bovine papilloma virus (36). Nonetheless, enhancer elements are apparently not specific for viral genomes, since cellular enhancers of promoter function have recently been found in association with light- (52) and heavy- (2, 20, 41) chain immunoglobulin genes. However, the issue of whether enhancer sequences have a general or more specialized role in regulating mammalian gene expression remains to be established.

Upstream promoter and the function of upRNA. It had been proposed that the binding of T antigen to the high-affinity sites I and II blocks transcription from the downstream site and diverts RNA polymerase to the upstream sites (18, 25). But this explanation is apparently incorrect, since normal quantities of upRNA are produced even when the adjacent T-antigen binding sites are eliminated. Our experiments show clearly that template DNA replication is required to increase upRNA transcription, but the basis of this dependence is unknown. The replication of viral DNA could cause a change in chromatin structure (34) that specifically activates transcription from the upstream site. Or, replication could simply increase the copy number of transcription templates and increase the association of RNA polymerase with the upstream promoter. The latter explanation gains some support from the finding that upRNA is produced at low levels even in the absence of DNA replication (Fig. 8).

T antigen appears to repress early-region transcription by blocking the initiation of transcription at the T-antigen binding site (16, 25, 49). However, T antigen binding to sites I and II does not prevent the elongation of transcripts that are initiated upstream (16, 25, 49). This probably explains why T antigen represses doRNA formation to a greater extent than upRNA formation. Although the temperature shift experiments with *tsA58* indicated that upRNA was repressed somewhat by T antigen, at least very late in infection, it appears that transcription from the upstream site is less sensitive to T-antigen expression than is transcription from the downstream site. Our data suggest that the 5′ termini of upRNA occur just upstream of T-antigen binding site II. Hansen et al. (25) have deduced the same location for the 5′ termini of upRNA from transcription studies in vitro; moreover, they have correlated repression of doRNA transcription with T antigen binding to site I and repression of upRNA formation with T antigen binding to site II.

There are several striking similarities between transcription from the SV40 early-region upstream promoter and the late-region promoter. To a large extent, transcription from both is dependent upon T antigen for viral DNA replication. Studies with deletion mutants have shown that expression of the late region also occurs in the absence of the enhancer

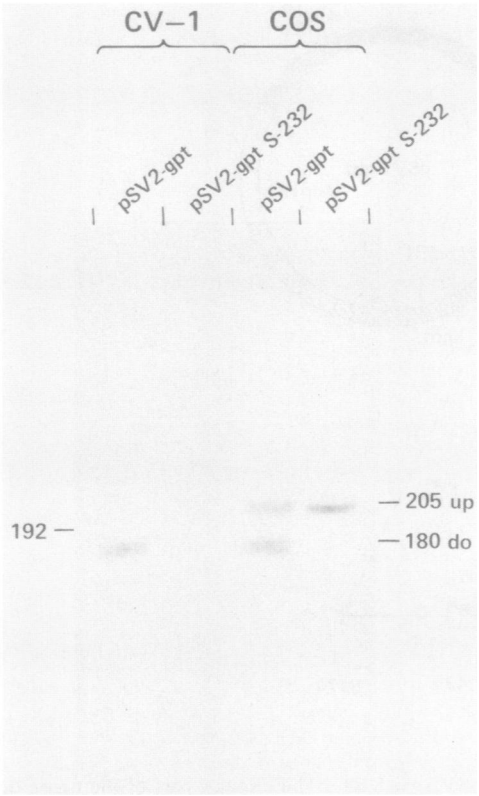


FIG. 12. Mapping the 5′ ends of RNA formed in transfections of COS and CV-1 cells with plasmids pSV2-*gpt* and pSV2-*gpt* S-232. Cytoplasmic poly(A)$^+$ RNA was isolated from CV-1 or COS cells transfected with pSV2-*gpt* or pSV2-*gpt* S-232 DNA as indicated in the headings and labels above each lane. The RNA was subjected to S1 nuclease analysis with a probe derived from pSV2-*gpt* DNA which was 5′ ^{32}P labeled at the *Bgl*II restriction site within the *gpt* segment (Fig. 6C). The products were separated by electrophoresis on an 8 M urea–4% polyacrylamide gel. The number in the left margin indicates the position of a marker (192 bases) produced by *Bgl*I endonuclease cleavage of the pSV2-*gpt* *Bgl*II probe (data not shown). Numbers in the right margin indicate the calculated sizes of the corresponding protected fragments. The up and do labels indicate the protected fragments referred to in the text as upstream and downstream, respectively. The analyses were performed with RNA from 2×10^7 transfected CV-1 cells or with 2×10^6 transfected COS cells.

sequence (16). There is no obvious TATA box sequence for positioning the 5′ termini of either upRNAs or late-region RNAs; moreover, although there are major 5′ termini for the late-region RNA and upRNA, minor species, which initiate throughout an overlapping region, can be detected for both. For example, upRNA species were noted with 5′ termini at about nucleotide 65 (16), and minor species of late-region RNAs have been found with 5′ termini located between nucleotides 28 and 264. The observations that the guanosine-cytosine-rich repeat region is involved in both late-region transcription (16, 17) and the formation of early-region upRNA (4, 9) hint at a possible connection between these divergently transcribed promoters.

Questions abound regarding the physiological function of upRNA and the elaborate regulation of the two early-region promoters. Since transcription from the upstream site proceeds through *ori* (27), it is conceivable that the upstream promoter plays a role in activating viral DNA replication, as has been suggested for phage lambda (14). This could occur by opening the DNA helix at *ori* to permit initiation by a DNA primase; or, the upRNA itself could be cleaved at *ori* to provide a primer RNA (31). Another possible function is suggested by the existence in upRNA, but not in doRNA, of an open reading frame 5′ to the start of the T-antigen coding sequences; this additional reading frame could code for a protein of 23 amino acids. Although no protein of this size has been detected during the late phase of SV40 infection, the possibility of its synthesis has yet to be thoroughly investigated. Perhaps this small protein provides a function needed late in lytic infection. A search for this leader protein seems justified considering the recent discovery of the agnogene protein which is encoded within the leader segment of the SV40 16S late RNA (32).

Both the small t-and large T-antigen-type splices occur in upRNA (Fromm, Ph.D. thesis), and it appears that the upRNA is translated to large T and small t proteins (9). However, preliminary findings indicate that upRNA is only 10 to 20% as efficient as doRNA in the translation of T antigens. Therefore, a plausible function for the upRNA is to provide another level of regulation for the production of early-region proteins late in infection.

Possible reasons for the transcriptional shift. There are numerous examples in both procaryotic and eucaryotic systems in which the same protein coding seqences are under the control of two different promoters: the *E. coli* galactose operon (48), the bacteriophage lambda *c*I gene (54), the yeast invertase (*SUC2*) gene (12), the *Drosophila* alcohol dehydrogenase gene (6), and the mouse α-amylase gene (64). In general, this kind of regulation permits the synthesis of the same protein to be regulated at separate levels under different metabolic conditions or at different stages of differentiation. Transcription from the downstream and upstream promoters is coordinated with the two stages of virus infection, before and after the onset of viral DNA replication. But what advantage could there be to using different promoters at these two stages of infection? Although the function supplied by small t antigen in lytic infection in unknown, large T antigen appears to have several functions. In addition to initiating viral DNA replication at *ori*, large T antigen induces a variety of host cell metabolic functions, including the synthesis of ribosomal RNA, histones, and enzymes involved in DNA metabolism (1). Transcription from the downstream site is needed not only to supply enough T antigen for the initiation of viral DNA replication but also to push the host cell into a metabolic state that is primed for virus multiplication. Perhaps the high efficiency of the downstream promoter, utilizing the enhancer and TATA box elements, is geared toward priming the host cell quickly and effectively. Once viral DNA replication has begun, the level of early-region expression generated by the downstream promoter may be wasteful or even deleterious, particularly as the template copy number increases. A shift to the less efficient upstream promoter and the production of the less efficiently translated upRNA could coordinate a continued production of T antigen needed for replication of the increased quantity of viral DNA. Thus, a form of transcriptional gear shifting, from downstream promoter to upstream promoter, may be operating to adjust the rate of early-region expression to appropriate levels at the two different stages of virus growth. Furthermore, the autoregulation of both downstream and upstream promoters by T antigen could fine tune the amount of early-region expression to optimal levels during the course of lytic infection.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM-13235 from the National Institute of General Medical Sciences and by grant CA 15513 from the National Cancer Institute. A.R.B. and M.F. were National Science Foundation predoctoral fellows and trainees under grant 5-T32GM-07599 from the National Institute of General Medical Sciences.

We thank Jeffrey Sklar for helpful discussions and gifts of DNA and Richard Young for comments on the manuscript.

LITERATURE CITED

1. **Acheson, N. H.** 1980. Lytic cycle of SV40 and polyoma virus, p. 125–204. *In* J. Tooze (ed.), DNA tumor viruses, part 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
3. **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.
4. **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 promoters. Nucleic Acids Res. **12**:915–932.
5. **Benoist, C., and P. Chambon.** 1981. *In vivo* sequence requirements for the SV40 early promoter region. Nature (London) **290**:304–309.
6. **Benyajati, C., N. Spoerel, H. Haymerie, and M. Ashburner.** 1983. The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5′ end in different developmental stages. Cell **33**:125–133.
7. **Berk, A. J., and P. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:1274–1278.
8. **Boyer, H. W., and D. Roulland-Dossoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
9. **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.
10. **Buchman, A. R., L. Burnett, and P. Berg.** 1980. The SV40 nucleotide sequence, p. 799–829. *In* J. Tooze (ed.), DNA tumor viruses, part II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Capecchi, M. R.** 1980. High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. Cell **22**:479–488.
12. **Carlson, M., and D. Botstein.** 1982. Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. Cell **28**:145–154.
13. **deVilliers, J., L. Olson, C. Tyndall, and W. Schaffner.** 1982. Transcriptional 'enhancers' from SV40 and polyoma show a cell type preference. Nucleic Acids Res. **10**:7965–7976.
14. **Dove, W. F., H. Inokuchi, and W. F. Stevens.** 1971. Replication control in phage lambda, p. 747–771. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Efstratiadis, A., F. C. Kafatos, and T. Maniatis.** 1977. The primary structure of rabbit β-goblin mRNA as determined from cloned cDNA. Cell **10**:571–585.
16. **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.
17. **Fromm, M., and P. Berg.** 1983. SV40 early- and late-region promoter functions are enhanced by the 72-base-pair repeat inserted at distant locations and inverted orientations. Mol. Cell Biol. **3**:991–999.
18. **Ghosh, P. K., and P. Lebowitz.** 1981. Simian virus 40 early mRNA's contain multiple 5′ termini upstream and downstream from a Hogness-Goldberg sequence; a shift in 5′ termini during the lytic cycle is mediated by large T antigen. J. Virol. **40**:224–240.
19. **Ghosh, P. K., P. Lebowitz, R. J. Frisque, and Y. Gluzman.** 1981. Identification of a promoter component involved in positioning the 5′ termini of the SV40 early mRNAs. Proc. Natl. Acad. Sci. U.S.A. **78**:100–104.
20. **Gillies, S. D., S. L. Morrison, V. T. Ori, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33**:717–728.
21. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
22. **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.
23. **Graham, F. L., and G. F. Whitmore.** 1970. Incorporation of 1-α-D-arabinofuranosylcytosine into DNA and on the inhibition of DNA polymerase by 1-α-D-arabinofuranosylcytosine 5′-triphosphate. Cancer Res. **35**:225–236.
24. **Griffin, B. E.** 1980. Structure and genomic organization of SV40 and polyoma virus, p. 61–124. *In* J. Tooze (ed.), DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Hansen, U., D. G. Tenen, D. M. Livingston, and P. Sharp.** 1981. T antigen repression of SV40 early transcription from two promoters. Cell **27**:603–612.
26. **Hardison, R. C., E. T. Butler, E. Lacy, T. Maniatis, N. Rosenthal, and A. Efstratiadis.** 1979. The structure and transcription of four linked rabbit β-like globin genes. Cell **18**:1285–1297.
27. **Hay, R. T., and M. L. DePamphilis.** 1982. Initiation of SV40 DNA replication *in vivo*: location and structure of 5′ ends of DNA synthesized in the ori region. Cell **28**:767–779.
28. **Hearing, P., and T. Shenk.** 1983. The adenovirus type 5 E1A transcription control region contains a duplicated enhancer segment. Cell **33**:695–703.
29. **Horowitz, M., C. L. Cepko, and P. A. Sharp.** 1983. Expression of chimeric genes in the early region of SV40. J. Mol. Appl. Genet. **2**:147–159.
30. **Humphries, R. K., T. Ley, P. Turner, A. D. Moulton, and A. W. Nienhuis.** 1982. Differences in human α-, β-, and δ-globin expression in monkey kidney cells. Cell **30**:173–183.
31. **Itoh, T., and J. Tomizawa.** 1978. Initiation of replication of plasmid ColE1 DNA by RNA polymerase, ribonuclease H, and DNA polymerase I. Cold Spring Harbor Symp. Quant. Biol. **43**:409–417.
32. **Jay, G., S. Nomura, C. W. Anderson, and G. Khoury.** 1981. Identification of the SV40 agnogene products: a DNA binding protein. Nature (London) **291**:346–349.
33. **Khoury, G., and E. May.** 1977. Regulation of early and late simian virus 40 transcription: overproduction of early viral RNA in the absence of functional T-antigen. J. Virol. **23**:167–176.
34. **Klemonauer, K. H., E. Fanning, B. Otto, and R. Kuppers.** 1980. Maturation of newly replicated chromatin of simian virus 40 and its host cell. J. Mol. Biol. **136**:359–374.
35. **Laimanns, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **79**:6453–6457.
36. **Lusky, M., L. Berg., H. Weiher, and M. Botchan.** 1983. Bovine papilloma virus contains an activator of gene expression at the distal end of the early transcription unit. Mol. Cell. Biol. **3**:1108–1122.
37. **McKay, R., and D. DiMaio.** 1981. Binding of a T antigen-related protein to the DNA of SV40 regulatory mutants. Nature (London) **289**:810–813.
38. **Manteuil, S., and M. Girard.** 1974. Inhibitors of DNA synthesis: their influence on replication and transcription of simian virus 40 DNA. Virology **60**:438–454.
39. **Martin, F. M., and I. Tinoco.** 1980. DNA-RNA hybrids containing oligo (dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. Nucleic Acids Res. **8**:2295–2299.
40. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific cleavages. Methods Enzymol. **65**:499–580.
41. **Mercola, M., X. Wang, J. Olsen, and K. Calame.** 1983. Tran-

scriptional enhancer elements in the mouse immunoglobulin heavy chain locus. Science **221**:663–665.
42. **Mertz, J. E., and P. Berg.** 1974. Defective simian virus genomes: isolation and growth of individual clones. Virology **62**:112–124.
43. **Moreau, P., R. Hen, B. Wasylyk, R. Everrett, M. P. Gaub, and P. Chambon.** 1981. The SV40 72 bp repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res. **9**:6047–6069.
44. **Mulligan, R. C., and P. Berg.** 1980. Expression of a bacterial gene in mammalian cells. Science **209**:1422–1427.
45. **Mulligan, R. C., and P. Berg.** 1981. Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phosphoribosyl transferase. Proc. Natl. Acad. Sci. U.S.A. **78**:2072–2076.
46. **Mulligan, R. C., and P. Berg.** 1981. Factors governing the expression of a bacterial gene in mammalian cells. Mol. Cell. Biol. **1**:449–459.
47. **Mulligan, R. C., B. H. Howard, and P. Berg.** 1979. Synthesis of rabbit β-globin in cultured monkey kidney cells following infection with a SV40-β-globin recombinant genome. Nature (London) **277**:108–114.
48. **Musso, R. E., R. DiLauro, S. Adhya, and B. deCrombrugghe.** 1977. Dual control for transcription of the galactose operon by cyclic AMP and its receptor protein at two interspersed promoters. Cell **12**:847–854.
49. **Myers, R. M., D. C. Rio, A. K. Robbins, and R. Tjian.** 1981. SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell **25**:373–384.
50. **Myers, R., and R. Tjian.** 1977. Construction and analysis of simian virus 40 origins defective in tumor antigen binding and DNA replication. Proc. Natl. Acad. Sci. U.S.A. **77**:6491–6495.
51. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. J. Virol. **31**:360–369.
52. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33**:741–748.
53. **Reed, S., G. R. Stark, and J. C. Alwine.** 1976. Autoregulation of simian virus 40 gene A by T antigen. Proc. Natl. Acad. Sci. U.S.A. **73**:3083–3087.
54. **Reichardt, L. F., and A. D. Kaiser.** 1971. Control of λ repressor synthesis. Proc. Natl. Acad. Sci. U.S.A. **68**:2185–2189.
55. **Shortle, D. R., R. F. Margolskee, and D. Nathans.** 1979. Mutational analysis of the simian virus 40 replicon: pseudorevertants of mutants with a defective replication origin. Proc. Natl. Acad. Sci. U.S.A. **76**:6128–6131.
56. **Tegtmeyer, P.** 1980. Genetics of SV40 and polyoma virus, p. 297–337. *In* J. Tooze (ed.), DNA tumor viruses, part II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
57. **Tegtmeyer, P., M. Schwartz, J. K. Collins, and K. Rundell.** 1975. Regulation of tumor antigen synthesis by simian virus gene *A*. J. Virol. **16**:168–178.
58. **Tenen, D. G., L. L. Haines, and D. L. Livingston.** 1982. Binding of an analog of simian virus 40 T antigen to wild type and mutant viral replication origins. J. Mol. Biol. **157**:473–492.
59. **Tjian, R.** 1978. Protein-DNA interaction at the origin of simian virus 40 DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:655–662.
60. **Tjian, R.** 1978. The binding site on SV40 DNA for a T antigen related protein. Cell **13**:165–179.
61. **Villarreal, L. P., R. T. White, and P. Berg.** 1979. Mutational alterations within the simian virus 40 leader segment generate altered 16S and 19S mRNA's. J. Virol. **29**:209–219.
62. **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.
63. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure. Nucleic Acids Res. **75**:1175–1193.
64. **Young, R. A., O. Hagenbuchle, and U. Schibler.** 1981. A single mouse α-amylase gene specifies two different tissue specific mRNAs. Cell **23**:451–458.