Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 721–725, February 1983
Biochemistry

Definition of the simian virus 40 early promoter region and demonstration of a host range bias in the enhancement effect of the simian virus 40 72-base-pair repeat

(RNA polymerase II control region/simian virus 40 21-base-pair repeated sequence/sequential deletions/herpes simplex virus 1 *tk* and *Eco-gpt* marker genes/transformation frequencies)

BARRY J. BYRNE, MARK S. DAVIS, JULIE YAMAGUCHI, DERK J. BERGSMA*, AND KIRANUR N. SUBRAMANIAN

Department of Microbiology and Immunology, University of Illinois at Chicago, Health Sciences Center, Chicago, Illinois 60612

Communicated by Leon O. Jacobson, November 12, 1982

ABSTRACT **The simian virus 40 (SV40) origin region includes the viral replication origin and the early and late promoters and consists of a few palindromes, a 17-base-pair (bp) A+T-rich sequence, three copies of a G+C-rich 21-bp repeat, and two copies of a 72-bp repeat. We have made sequential deletions in the SV40 origin region and determined the early promoter efficiencies of these truncated DNA segments by connecting them in the correct orientation with the coding regions of selectable marker genes and assaying the expression of the chimeric marker genes *in vivo* in different host cell lines. A truncated SV40 early promoter segment containing only the TATA box and the major *in vivo* mRNA initiation sites has essentially no promoter efficiency. We have located the major component of the SV40 early promoter within the 21-bp repeated sequences, which consist of an alternating and mutually overlapping array of two C-rich oligonucleotides having the consensus sequences Y-Y-C-C-G-C-C-C (Y = pyrimidine nucleoside) and G-C-C-C-(C)-$\frac{T}{A}$-$\frac{A}{T}$-$\frac{A}{(T)}$-C-T. Between one and two copies of the 21-bp repeat were adequate for gene expression under conditions in which the enhancement effect of the 72-bp repeat was minimal. We also find that the SV40 72-bp repeat exhibits a pronounced host range in its enhancement of gene expression; the enhancement is only 2-fold in the nonpermissive mouse cells but amounts to 10- or 20-fold in the permissive monkey cells or the semipermissive human cells, respectively.**

Comparison of the promoter sequences of various cellular and viral genes transcribed by eukaryotic RNA polymerase II has identified certain consensus sequences located at similar distances upstream from the mRNA initiation sites (cap sites). One of these, called the "TATA box" (1), located about 28 nucleotides upstream from the cap site, is involved in directing the accurate initiation of transcription (2). Another sequence called the CAT box, having the consensus sequence G-G-Y-C-A-A-T-C-T (Y = pyrimidine nucleoside), occurs about 80 nucleotides upstream from the cap site (3). Deletion of the region spanning the CAT box drastically reduces promoter efficiency in a few systems examined so far (4–6). A more recent study on the promoter of the thymidine kinase (*tk*) gene of herpes simplex virus 1 (HSV-1) has shown that mutations, not within the CAT box, but in a sequence C-C-C-C-G-C-C-C located 105 nucleotides upstream from the cap site, reduce promoter efficiency drastically (7).

The simian virus 40 (SV40) early promoter region is present within a broader control region located within nucleotides 5,171† and 270† in the DNA that also includes the viral replication origin and the promoter of the late region. This region contains, sequentially, a 27-base-pair (bp) perfect palindrome (containing the major *in vivo* early mRNA cap sites; see ref. 12), a 17-bp A+T-rich sequence containing the early region TATA box, three copies of a G+C-rich 21-bp repeat, and two copies of a 72-bp repeat present near the most upstream location of this control region.

Removal of the TATA box and downstream sequences including the major cap sites does not affect the SV40 early promoter efficiency (2, 13). Removal of one copy of the 72-bp repeat gives rise to a viable mutant of SV40 (14), but removal of both copies causes nonviability (14) and severely impairs the functioning of the viral early region (15). This impairment can be rectified by reinserting the 72-bp sequence, either at the same location or as far as 4,000 bp from the early coding region, and in either orientation (16). The SV40 72-bp repeat is also capable of enhancing the expression of heterologous genes, again in a *cis* fashion (16, 17). Thus, the major role of the 72-bp repeat is as an enhancer of gene expression, by some unknown mechanism. The orientation-independence and long-distance aspects of its action show that it is not a promoter in the conventional definition of the term. A similar enhancer sequence has also been found in polyoma virus DNA (18, 19).

Recently, we have precisely mapped the boundaries of the SV40 replication origin by making sequential deletions of the origin region starting from the end near the 72-bp repeat and assaying the replication efficiencies of chimeric plasmids containing these truncated origin region segments *in vivo* in COS-1 monkey cells (20). These strategic deletions have been exploited in the present study to map the viral early promoter. We have cut out the truncated SV40 origin region segments from the chimeric plasmids referred to above and connected them (in the orientation of the viral early transcription) to the promoterless coding regions of selectable marker genes. We have determined the promoter efficiencies of the SV40 origin region segments by measuring the transformation frequencies of the respective chimeric marker genes *in vivo*, using different cell types.

MATERIALS AND METHODS

Materials. The HSV-1 *tk* gene and the *Escherichia coli* xanthine/guanine phosphoribosyltransferase (*Eco-gpt*) gene were the marker genes used for assaying the promoter efficiencies of the SV40 origin region fragments in this study. The HSV-1

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: SV40, simian virus 40; HSV-1, herpes simplex virus 1; bp, base pair(s); T antigen, tumor antigen.

* Present address: Dept. of Cell Biology, Baylor College of Medicine, Houston, TX 77030.

† The SV40 numbering system of Buchman *et al.* (8), derived from the sequence of SV40 DNA determined by Reddy *et al.* (9), Fiers *et al.* (10), and Van Heuverswyn and Fiers (11), is used in this paper.

tk plasmid pRB103, identical with the plasmid pFG5 described in the literature (21), and the *Eco-gpt* plasmid pSV2-gpt (22), were kindly provided by B. Roizman (University of Chicago) and P. Berg (Stanford University), respectively. Human 143B tk^- and mouse L tk^- $aprt^-$ cell lines (*aprt* encodes adenine phosphoribosyltransferase), used for the *tk* transformation assays, were obtained from E. Mocarski (University of Chicago) and E. Kaufman (University of Illinois at Chicago), respectively. Monkey cell line CV-1P and mouse L cell line were used for the *Eco-gpt* transformation assays. The mycophenolic acid used in the *Eco-gpt* selective media was a gift from Lilly Research Laboratory (Indianapolis, IN). Enzymes and radioactive materials were purchased from commercial sources. Phage T4 DNA ligase was a generous gift from P. Matsumura (University of Illinois at Chicago).

Directional Joining of SV40 Early Control Region Fragments to the Marker Genes. The construction of a series of pSVori recombinants, whose SV40 inserts have sequential deletions in the origin region, has been described (20). The chimeric marker genes, in which the coding regions of the marker genes are driven by the truncated SV40 early control region segments derived from the pSVori recombinants, were constructed as shown schematically in Fig. 1. The *tk* plasmid, pRB103, was cut with *Hin*dIII and *Bgl* II to remove the smaller fragment containing the *tk* promoter by gel electrophoretic fractionation. The larger fragment, containing the *tk* coding region, has different 5′ cohesive termini and is suitable for the directional joining of the SV40 fragments. The pSVori recombinants were cut first with *Hin*dIII at nucleotide 29 of the plasmid (23), followed by cleavage at nucleotide 5,226 of the SV40 insert (just downstream from the major early cap sites; see ref. 12) with *Alu* I, to produce fragments containing the SV40 origin region and flanking plasmid sequences. The *Alu* I ends of these fragments were modified to *Bgl* II ends by using *Bgl* II linkers. These *Hin*dIII/*Bgl* II-ended SV40 fragments were then ligated with the *Hin*dIII/*Bgl* II-ended *tk* vector fragment with the result that the *tk* coding region was just downstream from the SV40 early cap sites. The *Eco-gpt* constructs were made similarly as shown in Fig. 1. Promoterless HSV-1 *tk* and *Eco-gpt* constructs were made by removal of the promoters, conversion of the ends to a compatible type by using the same linkers, and cyclization by ligation *in vitro*.

Competent *E. coli* MC1061 cells were transformed with the respective ligation products (24). Ampicillin-resistant colonies were picked and probed suitably, using known colony hybridization protocols (25). Positive colonies were analyzed by the miniscreen procedure (26), and those with the correct restriction analysis were chosen for large-scale DNA production.

Transformation of the Cell Lines. The *tk* and *Eco-gpt* transformations were done according to Wigler *et al.* (27) and Mulligan and Berg (22), respectively, using between 1 and 10 μg of each construct per 5×10^5 cells. Only those cells that took up the marker gene and expressed it would be able to survive and proliferate to form clones in the selective medium. The healthy colonies of greater than 50 cells produced in each case were counted.

The parent plasmids pRB103 and pSV2-gpt were used as the positive controls in these experiments. The transformation frequencies were expressed as number of colonies produced per 1×10^6 cells per μg of *tk* or *gpt* gene. Plasmid pRB103 had a

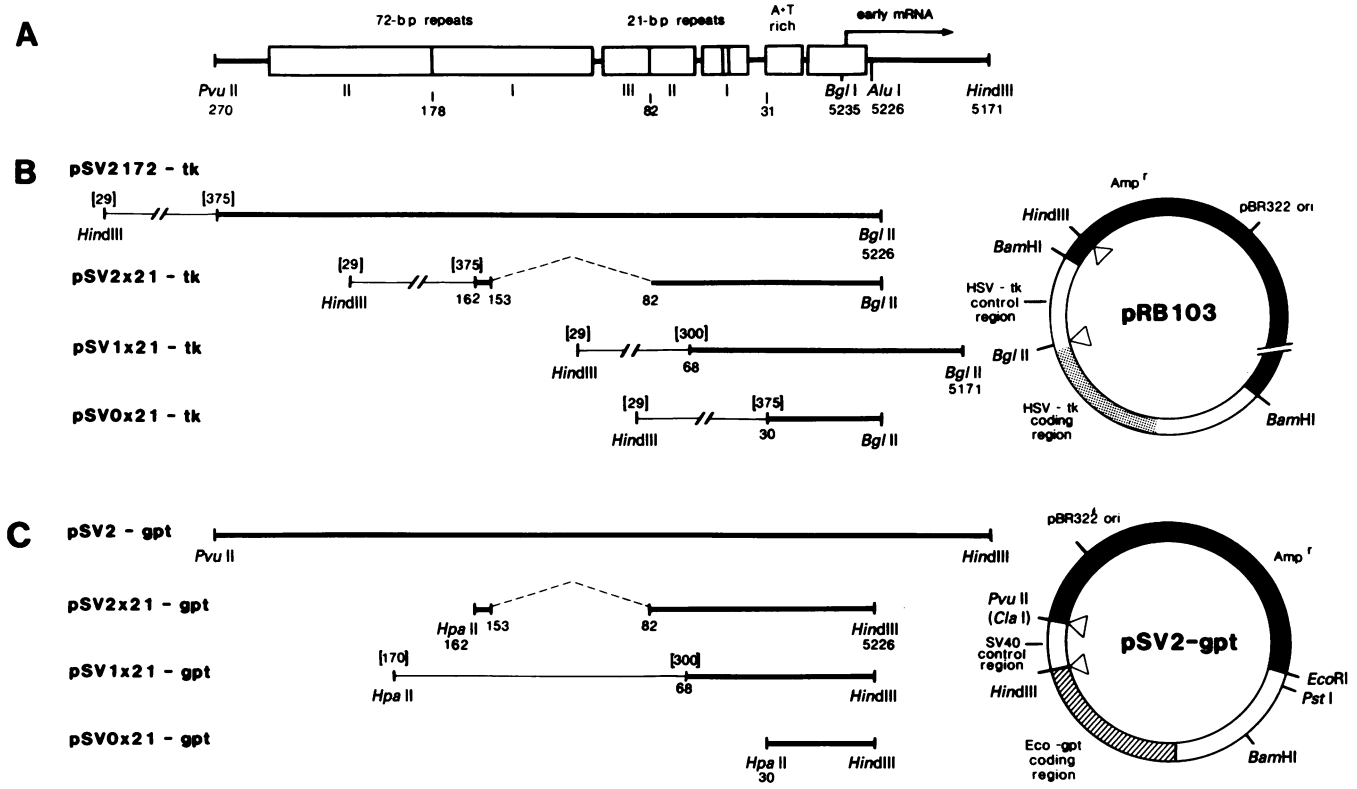

FIG. 1. Construction of the pSV-tk and pSV-gpt fusion plasmids. The procedure is described in detail in *Materials and Methods*. (*A*) SV40 early control region. The end points of the bold lines below, in *B* and *C*, show the extents of this control region contained in the fragments derived from the respective pSVori recombinants (20). Thin lines connected to the bold lines show the adjoining plasmid sequences, if any, in the fragments. The nucleotide numbers of the plasmids are shown within brackets and the numbers for SV40 without brackets. Empty triangles within the parent *tk* (pRB103) and *Eco-gpt* (pSV2-gpt) vectors, shown in the right-hand side of *B* and *C*, indicate the points where the vectors were cut to remove the intervening promoter regions and replace them with the SV40 control region fragments shown to their left. The broken lines in the pSV2x21 fragments indicate the extent of an internal deletion. Amp^r, ampicillin resistance.

transformation frequency of 1×10^{-3} in mouse L tk^- $aprt^-$ cells in our hands, in good agreement with the value reported by Colbere-Garapin *et al.* (21), who used an identical plasmid. The transformation efficiency of human 143B tk^- cells by pRB103 was about 1/200th as high. Interestingly, the parent *gpt* vector, pSV2-gpt, which contains the complete SV40 early promoter and enhancer sequences, transformed monkey cells with a 4- to 5-fold higher frequency compared to mouse L cells, whereas pSV2x21-gpt, lacking the 72-bp enhancer sequence, transformed both monkey and mouse cells with about the same frequency. The promoterless *tk* and *Eco-gpt* vectors, used as negative controls, did not produce any colonies in these experiments, and the same result was obtained with mock transfections.

RESULTS

Structural Analyses of the Chimeric Plasmids. The extent of the SV40 origin region sequences present in the *tk* and *Eco-gpt* constructs, confirmed by digestions with the appropriate combinations of restriction endonucleases (data not presented), are as shown in Fig. 1. All constructs made in this study (except the promoterless *tk* and *Eco-gpt* vectors) contain the 27-bp perfect palindrome (having the major early *in vivo* cap sites) and the 17-bp A+T-rich sequence (having the early TATA box). The pSV0x21 constructs do not have any of the copies of the upstream repeated sequences. The pSV1x- and 2x21 constructs have one or two copies of the 21-bp repeat, respectively. The pSV2172-tk construct (and the equivalent pSV2-gpt vector) contain three copies of the 21-bp repeat and two copies of the 72-bp repeat.

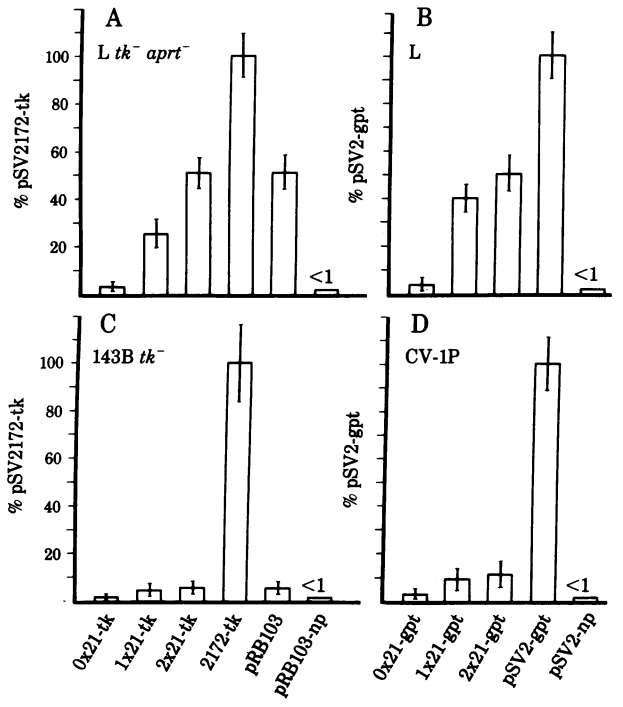

FIG. 2. The relative transformation frequencies of the pSV-tk and pSV-gpt chimeric genes obtained with three different cell lines. (*A*) Mouse cells and *tk* constructs; (*B*) Mouse cells and *gpt* constructs; (*C*) human cells and *tk* constructs; and (*D*) monkey cells and *gpt* constructs. The transformation frequencies of the constructs are expressed relative to those of pSV2172-tk and pSV2-gpt taken as 100%. The terms pRB103-np and pSV2-np refer to the promoterless constructs of the *tk* and *Eco-gpt* vectors, respectively. The results shown are the means and SEMs of six different experiments with five plates for each condition in each experiment.

Mapping of the Major Component of the SV40 Early Promoter Within the G+C-Rich 21-bp Repeats. The promoter efficiencies of the SV40 early control region fragments in driving the expression of the marker genes were assayed by determining the transformation frequencies of the constructs relative to those of the positive controls used in the experiments. In mouse L tk^- $aprt^-$ cells, plasmid pSV2x21-tk exhibited a transformation frequency that was identical with that of pRB103 (Fig. 2*A*). Plasmid pSV1x21-tk exhibited about a half of the transformation frequency of pSV2x21-tk, and that of pSV0x21-tk was extremely low. Because the pSV0x21-tk control region contains the SV40 major early mRNA cap sites and TATA box, this result indicates that these sequences alone [whose deletions do not affect the expression of the viral early region as shown by earlier results (2, 13)] are not sufficient for promotion of gene expression *in vivo*. The addition of one copy of the 21-bp repeat to the control region (as in pSV1x21-tk) causes a significant increase in transformation frequency, and the addition of two copies of the 21-bp repeat (as in pSV2x21-tk) raises the transformation frequency to the level achieved with the homologous HSV-1 *tk* promoter (Fig. 2*A*). From these results, we conclude that the major component of the SV40 early promoter activity resides within the G+C-rich 21-bp repeat.

Very similar results were obtained in mouse L cells by using the corresponding *Eco-gpt* constructs (Fig. 2*B*), supporting the conclusion reached above on the importance of the 21-bp repeat in the SV40 early promoter region.

SV40 72-bp Repeat Exhibits a Pronounced Host Range Bias in Its Enhancement Effect. The major structural difference between pSV2172-tk (and its equivalent, pSV2-gpt, in the *Eco-gpt* system) on the one hand, and the corresponding pSV2x21 constructs on the other hand, is the presence of the 72-bp enhancer sequence in the former, but not in the latter. In the nonpermissive‡ mouse L cells, the enhancement effect of the expression of the marker genes by the 72-bp repeat was found to be approximately 2-fold (Fig. 2 *A* and *B*). However, in permissive‡ (monkey) or semipermissive‡ (human) cells, the constructs containing the 72-bp repeat exhibit a severalfold higher transformation frequency, compared to those not having the 72-bp repeat (Fig. 2 *C* and *D*). In monkey cells, the difference between the promoter efficiencies of the pSV2172 and pSV2x21 constructs is approximately 10-fold (Fig. 2*D*), whereas it is approximately 20-fold in human cells (Fig. 2*C*). These results show that, in permissive and semipermissive cells, the 72-bp repeat is required for efficient gene expression *in vivo*.

In monkey cells, the 21-bp repeats exhibit a basal level of promoter efficiency, which is further reduced when the number of copies of the 21-bp repeat is decreased, with the construct containing none of the copies of this sequence exhibiting the lowest level of promoter efficiency (Fig. 2*D*). This finding confirms our previous finding obtained by using the nonpermissive mouse cells (Fig. 2 *A* and *B*) that the 21-bp repeat is an important component of the SV40 early promoter.

DISCUSSION

In this study, SV40 early control region fragments, containing strategic deletions placed sequentially starting from the most upstream location, have been excised from the SV40 early cod-

‡ These terms are used strictly in the context of accepted convention on the ability of a cell line to support an infection by SV40. In nonpermissive cells, the early region of SV40 is expressed, but viral DNA replication and late region expression do not occur. In semipermissive cells, early region expression and DNA replication occur, but there is a block in synthesis of the late proteins. In permissive cells, all of these phenomena occur normally.

ing region, connected to the coding regions of the HSV-1 *tk* and *Eco-gpt* marker genes, and assayed *in vivo* in different cell types for their efficiencies in driving the expression of the marker genes. This approach has the following advantages: (*i*) the experiments can be performed across species barriers, using different mammalian host cell lines; (*ii*) the transformation assay used is very sensitive, enabling one to measure both small and large differences in promoter efficiencies; and (*iii*) one of the products of the SV40 early region, the large tumor (T) antigen, is known to repress the expression of the early region by binding to the control region and acting at the transcriptional level (28, 29); using marker genes for the assay, we have eliminated this complication due to autoregulation by the T antigen.

One of the interesting observations made in this study is the host range bias exerted by the SV40 72-bp repeat in its enhancement effect. The enhancement effect is much more pronounced in the permissive monkey cells and semipermissive human cells than in the nonpermissive mouse cells. The reason for this host range bias is not presently understood, and awaits clarification from studies on the mechanism of the enhancing effect of the SV40 72-bp repeat. Our finding on the host range bias of the SV40 72-bp repeat is similar, but in the reverse phylogenetic direction, to the observation made by Levinson *et al.* (30) on the substitution of the SV40 72-bp repeat with a 72-bp repeat derived form the long terminal repeat of a murine sarcoma virus. This murine viral sequence did substitute for the SV40 72-bp repeat, but the effectiveness of the former was only a fraction of that of the latter in monkey cells.

Another observation that we have made in this study is significant, but again awaits clarification from studies on the mechanism of action of the SV40 72-bp repeat. We find that the SV40 72-bp repeat is a more stringent requirement for gene expression in the semipermissive human cells than in the permissive monkey cells (see Fig. 2 *C* and *D*). This finding is unexpected because one would expect the maximal enhancing effect in the homologous cell type.

Last, but not least, the observations made in this study show that the G+C-rich 21-bp repeat is an important component of the SV40 early promoter. In a situation in which the complication due to the enhancement effect of the 72-bp repeat is minimal—namely, in mouse L cells—the 21-bp repeat is found to function as an efficient promoter, comparable in efficiency to the homologous promoter of the HSV-1 *tk* gene (Fig. 2 *A* and *B*). The complete removal of the 21-bp repeat (in addition to the 72-bp repeat) causes an almost total abolition of promoter efficiency, though the TATA box sequence and the major early mRNA cap sites are still present in the defective control region. This finding points to the 21-bp repeat as the main recognition signal for the eukaryotic RNA polymerase II for initiation of transcription downstream from this sequence. This is the second function that we ascribe to the SV40 21-bp repeat, in addition to our earlier finding that the 21-bp repeat, though not a part of the core region of the SV40 replication origin, exerts a pronounced enhancement on the replication efficiency of the DNA (20).

The sequence of the pertinent portion of the SV40 early promoter region is shown in Fig. 3. The three copies of the 21-bp repeat consist of six copies each of two alternating and mutually overlapping C-rich oligonucleotides having the consensus sequences Y-Y-C-C-G-C-C-C and G-C-C-C-(C)-$\frac{T}{A}$-$\frac{A}{T}$-$\frac{A}{(T)}$-C-T as shown in Fig. 3. The former oligonucleotide, whose presence as a repeated sequence was noted by Subramanian *et al.* (31) during their sequence analysis of the SV40 *Eco*RII-G fragment containing this region, is strongly reminiscent of the sequence C-C-C-C-G-C-C-C located at position −105 of the upstream control region of the HSV-1 *tk* gene and found to be an important element of the HSV-1 *tk* promoter (7). The latter oligonucleotide, G-C-C-C-(C)-$\frac{T}{A}$-$\frac{A}{T}$-$\frac{A}{(T)}$-C-T, resembles the CAT box consensus sequence, G-G-Y-C-A-A-T-C-T, believed to be a component of eukaryotic RNA polymerase II promoters (3).

From the foregoing discussion it is evident that the SV40 early promoter region is unique in containing an array of multiple alternating and mutually overlapping copies of two putative upstream control elements that are present as single copies in the other known cellular and viral promoters interacting with eukaryotic RNA polymerase II. This multiplicity of control elements explains why only a portion of this region is required for efficient promotion of gene expression, as in the case of our pSV1x21 and pSV2x21 constructs. The multiplicity of these control elements also explains the dosage effect that we observe in comparing the promoter efficiencies of the SV40 early control region segments present in the pSV0x21, pSV1x21, and pSV2x21 constructs.

The results of the present study and other earlier results on

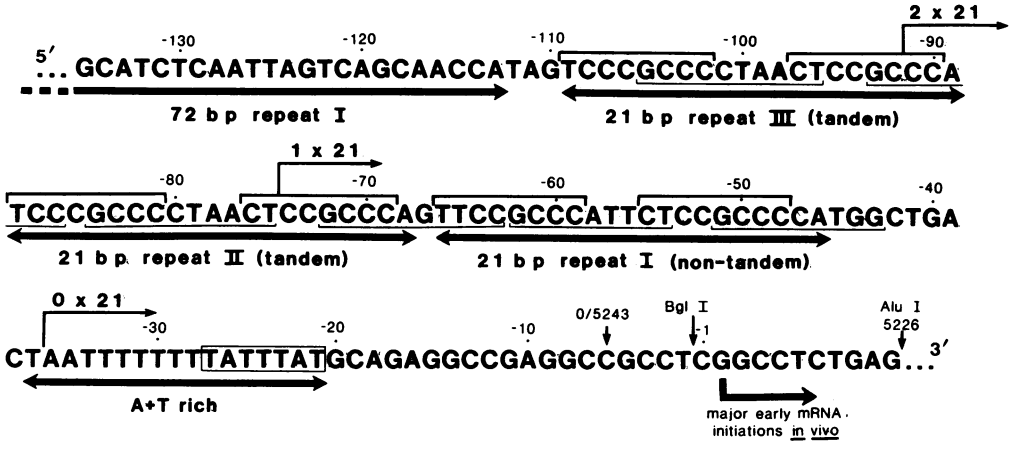

FIG. 3. Partial sequence of the SV40 early control region. Arrows above the sequence show the upstream boundaries of the SV40 control region segments of pSV0x-, 1x-, and 2x21 constructs. The major *in vivo* early mRNA initiation site shown at bottom right is the most upstream one of a set of three sites located within a spacing of five nucleotides at this location (12). The negative numbers above the sequence are upstream nucleotides relative to this mRNA initiation site. The early TATA box is shown boxed within the A+T-rich sequence. Brackets above the 21-bp repeats show the six copies of the C-rich octanucleotide having the consensus sequence Y-Y-C-C-G-C-C-C; brackets below show six copies of a sequence analogous to the CAT box consensus sequence (3). Note the alternating and partially overlapping arrangement of these two C-rich oligonucleotides within the 21-bp repeats.

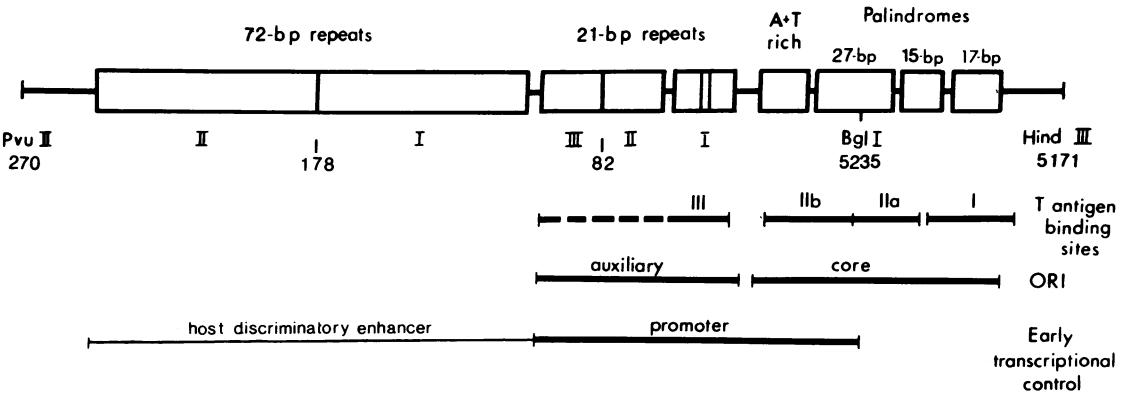

FIG. 4. Sequences at the SV40 control region and their functions determined so far. The T antigen binding sites are based on Myers *et al.* (29), modified subsequently. The term ORI refers to the viral replication origin whose limits were determined by Bergsma *et al.* (20). The G+C-rich 21-bp repeats increase the viral replication efficiency (20) and function as the major component of the viral early promoter (the present study). The TATA box and the major early cap sites are also included within the early promoter because of their roles in the accurate initiation of transcription (2). The 72-bp repeat is an enhancer sequence (16, 17), required for efficient gene expression in permissive or semipermissive cells but not in non-permissive cells (the present study).

the functions of the sequences located in the SV40 origin region are summarized in Fig. 4.

We thank Velma Kuykendall for typing this manuscript. This work was supported by American Cancer Society Research Grant NP248C (to K.N.S.). K.N.S. is the recipient of American Cancer Society Faculty Research Award FRA-202.

1. Goldberg, M. (1979) Dissertation (Stanford Univ., Stanford, CA).
2. Gluzman, Y., Sambrook, J. & Frisque, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3898–3902.
3. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8**, 127–142.
4. Grosveld, G., deBoer, E., Shewmaker, C. & Flavell, R. (1982) *Nature (London)* **295**, 120–126.
5. Dierks, P., van Ooyen, A., Mantei, N. & Weissmann, C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1411–1415.
6. Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. (1981) *Cell* **27**, 279–288.
7. McKnight, S. & Kingsbury, R. (1982) *Science* **217**, 316–324.
8. Buchman, A., Burnett, L. & Berg, P. (1981) in *DNA Tumor Viruses*, Molecular Biology of Tumor Viruses, ed. Tooze, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed., pp. 799–841.
9. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. & Weissman, S. M. (1978) *Science* **200**, 494–502.
10. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. & Ysebaert, M. (1978) *Nature (London)* **273**, 113–120.
11. Van Heuverswyn, H. & Fiers, W. (1979) *Eur. J. Biochem.* **100**, 51–60.
12. Ghosh, P. & Lebowitz, P. (1981) *J. Virol.* **40**, 224–240.
13. Benoist, C. & Chambon, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3865–3869.
14. Gruss, P., Dhar, R. & Khoury, G. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 943–947.
15. Benoist, C. & Chambon, P. (1981) *Nature (London)* **290**, 304–310.
16. Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. & Chambon, P. (1981) *Nucleic Acids Res.* **9**, 6047–6068.
17. Banerji, J., Rusconi, S. & Schaffner, W. (1981) *Cell* **27**, 299–308.
18. de Villiers, J. & Schaffner, W. (1981) *Nucleic Acids Res.* **9**, 6251–6264.
19. Tyndall, C., LaMantia, G., Thacker, C., Favaloro, J. & Kamen, R. (1981) *Nucleic Acids Res.* **9**, 6231–6250.
20. Bergsma, D. J., Olive, D. M., Hartzell, S. W. & Subramanian, K. N. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 381–385.
21. Colbere-Garapin, F., Chousterman, S., Horodniceanu, F., Kourilsky, P. & Garapin, A.-C. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3755–3759.
22. Mulligan, R. & Berg, P. (1980) *Science* **209**, 1422–1427.
23. Sutcliffe, J. G. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 77–90.
24. Lederberg, E. M. & Cohen, S. N. (1974) *J. Bacteriol.* **119**, 1072–1074.
25. Grunstein, M. & Wallis, J. (1979) *Methods Enzymol.* **68**, 379–389.
26. Birnboim, H. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
27. Wigler, M., Silverstein, S., Lee, L., Pellicer, A., Chang, Y. & Axel, R. (1977) *Cell* **11**, 223–232.
28. Tegtmeyer, P., Schwartz, M., Collins, J. & Rundell, K. (1975) *J. Virol.* **16**, 168–178.
29. Myers, R., Rio, D., Robbins, A. & Tjian, R. (1981) *Cell* **25**, 373–384.
30. Levinson, B., Khoury, G., Vande Woude, G. & Gruss, P. (1982) *Nature (London)* **295**, 568–572.
31. Subramanian, K. N., Dhar, R. & Weissman, S. M. (1977) *J. Biol. Chem.* **252**, 355–367.