Multiple Point Mutations Affecting the Simian Virus 40 Enhancer

Hans Weiher, Monika König, Peter Gruss

Summary. Enhancers, or activators, dramatically increase the transcriptional activity of certain eukaryotic genes. A series of multiple point mutations affecting the simian virus 40 (SV40) enhancer-activator region were generated in order to define the nucleotide sequence required for this function. Three independent assays provided information leading to the identification of nucleotides essential for enhancer function. One class leads to a decrease in gene expression, while the second completely abolishes functional activity. One critical replacement appears to be the first G (guanine) in a sequence TGGAAAG (T, thymine, A, adenine) located in the 5′ region of the 72 base-pair repeat of SV40. Comparison of this sequence with nucleotide sequences in other known enhancers leads to the identification of potential related core elements.

An understanding of the regulatory events controlling gene expression in higher eukaryotes is important in the study of such complex phenomena as differentiation, transformation, and malignancy. The substantial progress made on this subject has been aided by achievements in molecular biology such as molecular cloning and site directed mutagenesis. Much of this information is derived from studies with animal viruses. Because of their limited genetic capacity, the viruses rely on cellular enzymes, and thus provide a useful model system to study the basic mechanisms of cellular gene control.

One of the best studied animal virus systems is simian virus 40 (SV40) (Fig. 1) (*1*). Its genome is a double-stranded covalently closed circular DNA molecule of 5243 nucleotides (*2–4*). In permissive monkey kidney cells, an early and a late phase characterize its life cycle. The early and late coding sequences are situated on opposite halves of the genome and are transcribed in opposite directions (Fig. 1). Early gene expression has been extensively studied, and the controlling elements for transcriptional initiation do not appear to require virally coded gene products.

Several eukaryotic promoters have been examined both by DNA sequencing and by in vitro and in vivo analysis of mutants. These studies have led to the identification of the so-called Goldberg-Hogness or TATA box (T, thymine; A, adenine), a signal that is involved in the precise positioning of 5′ RNA ends (*5, 6*). Other sequences upstream of the TATA box, which are absolutely required for promoter activity, have been identified in various promoter sequences (*7*). Recently, a detailed analysis of the promoter of the herpes simplex thymidine kinase (TK) gene (*8*) resulted in an identification of three essential regions within 105 base pairs (bp) upstream of the RNA initiation site.

In SV40 a 72-bp repeated unit located approximately 100 to 175 bp upstream of the early cap (initiation site) position appeared to be essential in *cis* for the expression of the SV40 early genes (*9*). The repeated nature of this element is not essential, as indicated by naturally occurring variants of SV40 containing only a single copy of the 72-bp repeat (*4*). Further analyses revealed some surprising characteristics of this particular controlling element. It was found that the tandem repeats retained their activity when inserted at a different position or in inverse orientation relative to the early coding region (*10, 11*). Earlier studies of Capecchi (*12*) had shown that linkage of SV40 sequences to the herpes TK gene enhanced the frequency of transformation of TK^- to TK^+ cells. Similarly, a dramatic increase in rabbit globin gene transcription mediated by the SV40 72-bp repeats was demonstrated (*13*). These experiments showed that this region of the SV40 genome can stimulate transcription from heterologous promoters. The terms "enhancer" (*13*) and "activator" (*14*) have been proposed to describe elements of this kind. Analogous elements have also been found in other animal viruses such as Moloney sarcoma virus (MSV), Rous sarcoma virus (RSV), bovine papilloma virus (BPV), polyoma virus, and in the human genome (*11, 14–16*).

Several of the described activator elements have been tested for their ability to substitute for one another in functional assays. The MSV (*14*) and the BPV (*11*) enhancers can both replace the SV40 72-bp repeats to some extent as assayed by replication in monkey cells. However, quantitative measurements of gene expression from the SV40 early promoter with either the SV40 or MSV enhancer has indicated that these elements may have different host cell specificity (*17*). In addition, alterations within a region of the polyoma virus genome which contains the enhancer function (*16*) enables the mutated virus to grow in nondifferentiated teratocarcinoma stem cells (*18*). These data not only point to the central role of enhancers as *cis*-acting signal sequences, but also indicate their possible regulatory function. Although activator elements often seem to reside within repeated sequences, very little information is available concerning the critical sequences necessary for enhancer function. Identification of these sequences could help in the elucidation of the mechanism of enhancement. One problem in studying deletion mutants is that the removal of essential sequences brings other sequences into an abnormal juxtaposition. In an effort to define critical nucleotides within the SV40 enhancer region, we have constructed and analyzed mutants containing multiple base pair changes within the 72-bp repeat.

Strategy of Generating Mutants in the 72-bp Repeat Region

The approach we selected is based on a base-specific chemical treatment in which sodium bisulfite is used in vitro to convert cytosine residues in single-stranded DNA to uracil, without doing

This article represents work done by Hans Weiher, a postdoctoral fellow at the Institute of Microbiology at the University of Heidelberg and by Monika König, a microbiologist, and Peter Gruss, a visiting scientist, at the Laboratory of Molecular Virology, National Institutes of Health, Bethesda, Maryland 20205. The present address of Dr. Weiher is Department of Molecular Biology, Stanley Hall, University of California, Berkeley 94720. The present address of Dr. Gruss, is Department of Microbiology, University of Heidelberg, Heidelberg, Federal Republic of Germany.

significant damage to other bases (*19*). Thus, the resulting change will eventually be a transition from C (cytosine) to T. This agent has been successfully used to introduce mutations within the SV40 origin of replication (*20*). Recently, a very efficient modification of this method has been reported (*21*); its major advantage is the cloning of the desired DNA fragment into a single-stranded bacteriophage, which renders the DNA readily accessible to the chemical agent (Fig. 2). In brief, SV40 fragments were transferred into a suitable single-stranded vector such as fd 106 (*22*). Single-stranded DNA from two clones (fd-SV40-E_1 and fd-SV40-E_4) carrying SV40 inserts in opposite orientations (*23*) was treated in vitro with sodium bisulfite. This modified DNA was hybridized with fd-specific fragments, which subsequently served as primers for the repair synthesis of the second strand, creating U · A base pairs (U, uracil) in all positions originally affected by the chemical treatment. After the repair synthesis, the molecules were digested with Sph I and Kpn I, which excised a fragment containing approximately 50 nucleotides of the 5′ end of the 72-bp sequence plus 47 nucleotides of the SV40 late (L) leader (see Fig. 1) (*2–4*). The excised fragments, which consisted of a mixture of mutagenized DNA segments, were ligated to Sph I and Kpn I digested pSVTR 1-3. This plasmid contains the viable SV40 genome inserted at its only Taq I site into the Cla I site of the plasmid pAd 190 (Fig. 2). The SV40 insert in pSVTR 1-3 is a viable mutant which has been altered at the Hpa II position by insertion of a Bam HI restriction enzyme linker (*24*), and carries only one copy of the 72-bp repeated in wild-type SV40. The plasmids resulting from this cloning procedure also contain a single copy of the SV40 72-bp repeat and carry base changes exclusively within the Kpn I–Sph I fragment. In order to confirm the physical structure and to determine the individual point mutations, the nucleotide sequence of several mutants in the critical region was determined.

Both strands were sequenced (*25*) beginning at position 346. A compilation of the resulting nucleotide sequence data of seven cloned mutants compared to the SV40 wild-type sequence is presented in Fig. 3A. It is obvious that all mutants carry the expected nucleotide exchanges only between Kpn I and Sph I. Depending on the orientation of the SV40 insert within the fd vector used for initial mutagenesis (Fig. 2), the expected C → T or G → A (G, guanine) transitions were observed.

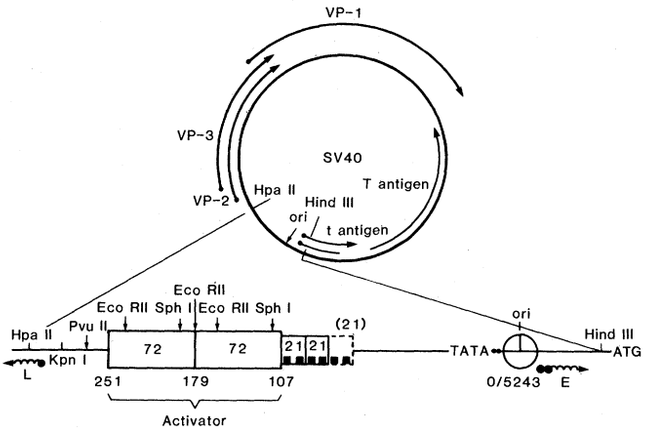


Fig. 1. Schematic representation of a functional map of SV40. SV40 has a circular covalently closed double-stranded DNA of 5243 bp (*1–4*). Taking the origin of replication (*ori*) as a start point, the early gene products (small t antigen and large T antigen) are transcribed counterclockwise. After the onset of DNA replication mediated by the origin of replication, abundant late gene products, VP-1, VP-2, and VP-3, are produced. The origin region is enlarged, demonstrating the presence of a 72-bp repeat and a 21-bp repeat. The filled square boxes within the 21-bp repeat represent a CCGCCC unit present at least six times. TATA is representing the Goldberg-Hogness box for early transcripts. The nucleotide numbers used are those of Van Heuverswyn and Fiers (*4*).

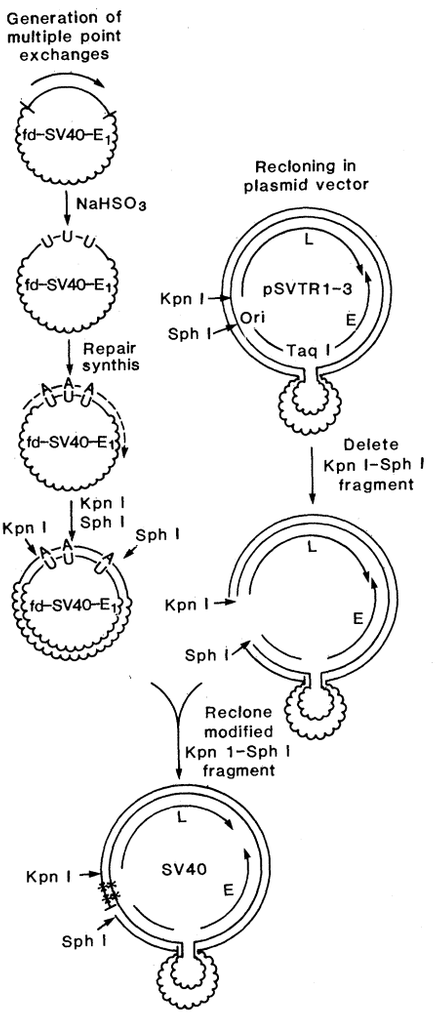


Fig. 2. Construction of SV40 point mutants in the enhancer or activator region. *Escherichia coli* C600 and fd recombinant phages (fd-SV40-E_1 and fd-SV40-E_4) were used (*23*). Phage fd-SV40-E_4 contains the entire late strand of SV40; fd-SV40-E_1 is the early strand with a 1500-bp deletion within the early region (*42*). Plasmid pSVTR 1-3 consists of a 2.3-kb plasmid (pAd 190) (*43*) conferring an ampicillin resistance marker and the plasmid origin of replication and of a complete SV40 genome. The SV40 is opened at its only Taq I site and fused to the only Cla I site of the plasmid vector. SV40 can be released by Taq I digestion. The SV40 portion of pSVTR 1-3 is different from wild-type SV40 in two regions. It contains only a single 72-bp unit and an additional Bam HI site, which has been introduced at the Hpa II site of SV40 in position 346 (*24*). The point mutants described are derivatives of pSVTR 1-3. For mutagenesis, single-stranded DNA (5 μg and 1 pmole) from phage fd-SV40-E_4 or fd-SV40-E_1 was dissolved in 17 μl of TE buffer (10 m*M* tris-HCl, *p*H 8, 0.1 m*M* EDTA) and 80 μl of 4*M* sodium bisulfite reaction mixture (156 mg of $NaHSO_3$, 64 mg of Na_2SO_3, and 0.43 ml of H_2O) and 3 μl of 50 m*M* hydroquinone were added. After 3 hours at 37°C in the dark, the samples were chromatographed on Sephadex G-100 columns (2 ml) previously equilibrated with 10 m*M* potassium phosphate buffer, *p*H 6, to remove bisulfite. The void volume containing the DNA was collected and adjusted to 0.2*M* tris-HCl, *p*H 9.2, 50 m*M* NaCl, 2 m*M* EDTA. After 12 hours at 37°C, the samples were desalted by filtration through Sephadex G-100 and lyophilized. The treated phage DNA was then hybridized for 10 minutes at 45°C with 0.1 μg of a mixture of fd phage–specific oligonucleotides (*44*) in 20 μl of DS buffer (40 m*M* tris-HCl, *p*H 7.2, 10 m*M* $MgCl_2$, 0.2 m*M* dithioerythritol, and 100 m*M* KCl). Subsequently, α-^{32}P-labeled deoxynucleotides dNTP (40 mCi/μmole) were added to a final concentration of 0.2 m*M* each in a final volume of 30 μl. DNA polymerase I (10 units) was added, and the sample was incubated for 25 minutes at 20°C and 10 minutes at 37°C. The reaction was stopped by ethanol precipitation of the DNA. After cleavage with Sph I and Kpn I, a 94-bp mutagenized DNA fragment was isolated after separation on a 6 percent polyacrylamide gel. This fragment was subsequently recloned after ligation to the Kpn I and Sph I site of pSVTR 1-3. From the mutagenesis of fd-SV40-E_4, four mutant clones were further analyzed and from fd-SV40-E_1 mutagenesis, three clones were analyzed.

Biological Activity of Point Mutants in the Activator-Enhancer Region

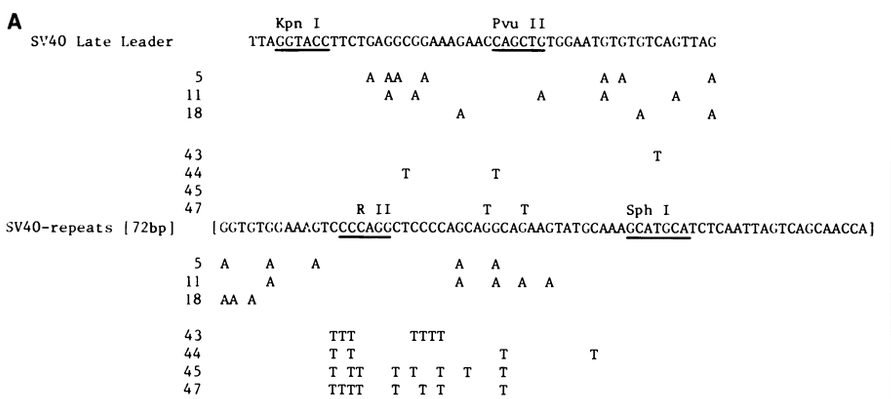
In an attempt to determine the biological effect of the multiple point mutations in individually cloned mutants, we first tested their ability to form plaques. This analysis was greatly facilitated by the available constructs, since cleavage with Taq I releases a functional SV40 genome (Fig. 2). These data are summarized in Table 1.

We found that two of the seven mutants (5 and 11) were unable to yield plaques, suggesting that these mutants must have nucleotide exchanges at crucial positions. It was further noted that mutants 18, 44, 45, and 47 seemed to produce plaques observable 1 to 2 days later than the wild-type control, and that mutant 43 gave rise to very small plaques 3 to 4 days later than did wild-type SV40. Since the region approximately 47 nucleotides toward the late side of the 72-bp repeat was also a target for mutagenesis, it was possible that the defect in mutants 5 and 11 were either in the enhancer element or in some late function (for example, promoter). In order to discern between these possibilities, we attempted to measure directly the amount of early gene product.

Replication of Mutant Viral DNA in Monkey Kidney Cells

The early gene product T antigen is essential for the replication of the SV40 genome (*20, 26*). We therefore used an assay in which we indirectly determined the amount of SV40 T antigen by measuring the amount of DNA replication in permissive monkey kidney (CV-1) cells (*27, 28*). This assay should enable us to determine the relative strength of the early promoter of individual mutants. It has been shown (*11*) with this assay that sequences to the late side of the Sph I site are not required for efficient early gene expression if an intact enhancer is present elsewhere in the molecule. In our experiments, parallel samples of CV-1 cells were transfected with individually cloned mutant DNA's released from plasmid sequences with the restriction endonuclease Taq I in the presence of DEAE-Dextran. At various times after transfection, cells were harvested and DNA was extracted (*29*). The appearance of supercoiled DNA was taken as an indication that replication had occurred. The first time point (6 hours after transfection) (Fig. 4), demonstrates that approximately the same amount of linear DNA from each mutant DNA preparation was taken up by CV-1 cells.

Subsequent time points (24, 48, 72, and 120 hours) show differences among the mutant DNA preparations. Wild-type controls (pA11-2352 carrying two 72-bp tandem repeats, and pSVTR 1-3,



B

Plaque Formation	DNA Replication	CAT Assay (arbitrary units)
+	+	100.0
-	-	5.2
-	-	8.8
+	+	27.0
(+)	+	15.4
+	+	24.7
+	+	75.3
+	+	n.d.

Fig. 3. Sequence analysis of multiple point mutants of SV40. (A) The positions and base changes found in the seven mutants are shown below the wild-type SV40 sequence. (B) A summary of the data obtained with the mutants in three different assays is shown. For plaque formation see Table 1, for DNA replication see Fig. 4, and for CAT assay see Fig. 6. The CAT assay calculation was done with the use of the data from the 30-minute time point in Fig. 6. The amount of acetylated chloramphenicol seen with plasmid 1-3 was assigned 100 arbitrary units.

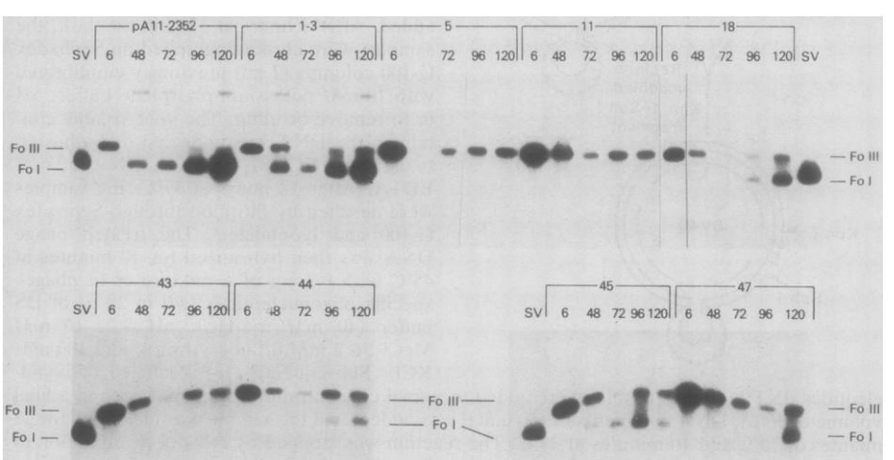


Fig. 4. Replication assay. Parallel dishes (60 mm) of a continuous line of African green monkey kidney cells (CV-1) were incubated for 70 minutes with 0.2 ml of minimal essential medium containing linear mutant DNA (previously cleaved with Taq I) (150 ng/ml) and DEAE-Dextran as facilitating agent. Low molecular weight DNA was isolated as described by Hirt (*26*), at 6, 48, 72, 96, and 120 hours after DNA transfection. A reflection of the linear input DNA was obtained by harvesting DNA prior to the onset of replication at 6 hours after transfection. Samples of each mutant at each time point were taken and subjected to electrophoresis in a 1.4 percent agarose gel. The DNA was transferred to a nitrocellulose filter (Schleicher and Schuell, 0.45 μm) basically as described (*45*), except that the gel was treated before denaturing the DNA with 0.25*M* HCl for 15 minutes. This transferred DNA was hybridized to a nick-translated SV40 DNA with a specific activity of approximately 2×10^8 cpm/μg. ^{32}P-Labeled bands were detected after exposure to an SB-5 x-ray film (Kodak). The numbers in the top lane indicate the mutants used (see Fig. 3). Plasmid 1-3 represents a wild type carrying only a single 72-bp repeat; pAll-2352 represents a 72-bp tandem repeat wild type. The bottom lane indicates the time of DNA harvest. SV stands for SV40 superhelical DNA (FoI). FoIII symbolizes linearized SV40 DNA.

carrying only a single 72-bp copy) and mutants 18, 44, 45, and 47 seem to generate increasing amounts of superhelical DNA, although to various extents. Mutant 43 produces only a minor amount of superhelical DNA and mutants 5 and 11 do not generate detectable circular duplex DNA. These data confirm the results of the plaque analysis. Although these experiments suggest that the mutations affect early gene expression, mutations interfering with the ability of the DNA molecule to replicate for other reasons could lead to spurious conclusions. We therefore wanted to determine directly the relative enhancer strength of each individual mutant independent of viral DNA replication. For this purpose, we have used a sensitive and quantitative assay to measure the effect of the various mutants on gene expression.

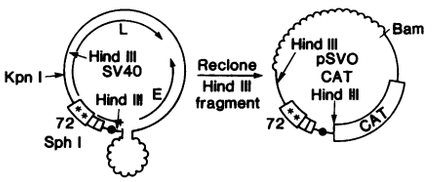


Fig. 5. Construction of mutants with Tn 9 (CAT) as indicator gene. The prokaryotic gene coding for the enzyme CAT has been cloned in a pSV-2 background (*30*) and was utilized for expression in different eukaryotic cells (*17*). A plasmid construct lacking only the eukaryotic promoter element carrying a suitable Hind III cloning site was also reported (*17, 30*). This plasmid (*46*) was used as a recipient for a Hind III fragment isolated from each SV40 point mutant (see also Fig. 2). Each Hind III fragment carried the mutated 72-bp region plus unaltered 21-bp repeats and Goldberg-Hogness box from SV40. After the individual fragments were cloned, their orientation was determined with the use of the restriction enzyme Bgl I, and only the sense orientation relative to the CAT coding region was selected.

Table 1. Infectious center-plaque assay of SV40 wild-type and point-mutant DNA. A continuous line of African green monkey kidney cells (CV-1) (4×10^4) was transfected with approximately 50 ng of viral specific DNA released from plasmid sequences by cutting each recombinant DNA with Taq I. The assay was performed as an infectious center assay (*48*). Cells were stained with neutral red 14 days after DNA transfection. Each (+) represents an average of approximately five to ten plaques per dish. The numbers used describe individual mutants (see Fig. 3). Number 1–3 represents a SV40 equivalent carrying a single 72-bp unit as control.

DNA	Plaques after transfection on day				
	14	15	16	18	20
1–3 (wt)	+	+	+	+	+
5	–	–	–	–	–
11	–	–	–	–	–
18	–	+	+	+	+
43	–	–	–	+	+
44	–	+	+	+	+
45	–	+	+	+	+
47	–	–	+	+	+

Quantitative Assay Measuring Gene Expression

The assay we selected has been described (*30*). In this assay, a bacterial gene from the transposable element Tn9 was cloned between the SV40 early promoter, and the SV40 small t intron and transcriptional termination signal. With these eukaryotic regulatory signals, this gene can be expressed in eukaryotic cells. It encodes the bacterial enzyme chloramphenicol-3-*O*-acetyltransferase (CAT) (E.C. 2.3.1.28), which can be sensitively and accurately quantified (*30, 31*). The functional analysis of the enzyme is based on the conversion of chloramphenicol to an acetylated form. Since the enzyme is normally absent from eukaryotes, the amount of enzyme produced after transfection of an appropriate plasmid construct into eukaryotic cells reflects the strength of a particular promoter or enhancer. It has also been shown that the extent of CAT activity in cell extracts correlates with gene activity at the level of RNA production. Therefore, the amount of enzyme produced should be proportional to the strength of an enhancer. With the use of this assay, it was recently demonstrated that full enhancer activity is exhibited by a permuted 72-bp repeat (*32*). Thus, in this assay no detectable activity is added by sequences upstream (5′) of the 72-bp repeat.

In order to determine the strength of the mutated enhancer elements, we transferred the SV40 Hind III-C fragment of individual mutants (5, 11, 18, 43, 44, 45) and of a control carrying a single copy of the 72-bp repeat into a pSVO recipient (see Fig. 5). The pSVO plasmid (*17, 30, 33*) effectively lacks all promoter elements, but provides the CAT coding region, the SV40 t intron, and the SV40 transcriptional termination signal. The cloning of the individual Hind III fragments in the sense orientation restores promoter elements, such as the TATA box and the 21-bp repeats in an unaltered form. Thus, a possible defect in the 72-bp repeat will be directly reflected by a reduction of CAT activity. In our experiment, each plasmid was transfected with Ca^{2+} as a facilitating agent onto semiconfluent CV-1 cells. Approximately 44 to 48 hours after transfection, the cells were harvested and the extract was prepared (*30*). A portion (50 microliters) of the extract was used for conversion of ^{14}C-labeled chloramphenicol to [^{14}C]acetylchloramphenicol in the presence of acetyl-coenzyme A as the acetyl donor, and the resulting products were separated by chromatography. In Fig. 6A, the exposed autoradiograms are presented. Depending on the mutant used, varying levels of enzyme activity were detected, as indicated by the amount of acetyl-

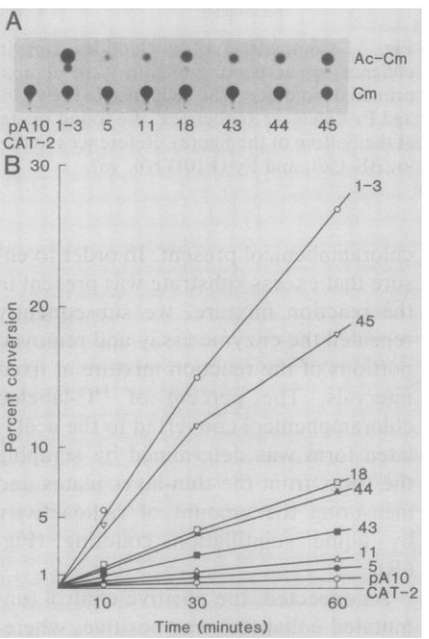


Fig. 6. Determination of CAT activities of individual point mutants. Each mutant plasmid and plasmid 1-3 carrying a single unmutated 72-bp unit was transfected onto CV-1 cells (70 percent confluent) by a modification of a Ca^{2+} precipitation method (*30, 47*). For each 100 mm dish, 25 μg of plasmid DNA was used. After incubation for 44 to 46 hours, cells were harvested and cell extracts were prepared (*17, 30*). The enzyme reactions were performed with 50 μl out of a total of 180 μl extract per dish and incubated with ^{14}C-labeled chloramphenicol in presence of 4 m*M* acetyl-coenzyme A in 250 m*M* tris, *p*H 7.8. The reactions were terminated by extraction with ethyl acetate either at 60 minutes (A), or portions were taken and extracted at the indicated time (B). The ^{14}C-labeled chloramphenicol was subsequently separated from the acetylated ^{14}C-labeled chloramphenicol by ascending thin-layer chormatography (silica gel 1B, Baker-flex) with a chloroform:methanol solvent (95:5 by volume). (A) An autoradiogram of the amount of CAT present in 50 μl extract, incubated for 60 minutes. In (B), portions of the enzyme reaction mixture were removed at intervals, and the percentage of acetylated chloramphenicol was determined by scraping the spots from the thin-layer plates and measuring the amount of radioactivity by liquid-scintillation counting.

Fig. 7. Comparison of nucleotides carrying enhancer activity. Nucleotides with known enhancer or activator function were compared. The crucial core nucleotide in the SV40 is printed in boldface. The sequence in the vicinity of this essential G was compared to MSV, BK, and Py enhancer sequences. We found related sets of nucleotides (underlined and emphasized at the bottom of the figure). References to the sequences are as follows: MSV (*17, 49*); SV40 (*1, 9*); BK (*50*); and Py (F101) (*16, 18*).

chloramphenicol present. In order to ensure that excess substrate was present in the reaction mixture, we subsequently repeated the enzyme assay and removed portions of the reaction mixture at fixed intervals. The percent of ^{14}C-labeled chloramphenicol converted to the acetylated form was determined by scraping the spots from the thin-layer plates and measuring the amount of radioactivity by liquid scintillation counting (Fig. 6B).

As expected, the positive control (unmutated enhancer) was positive, whereas the negative control, missing 50 nucleotides of the enhancer 5′ end (pA10-CAT-2) (*17*) yielded the predicted low conversion rates. It is interesting that neither of the mutants that were negative in plaque and replication assays (5 and 11) convert significantly more chloramphenicol to acetylchloramphenicol than does the repeat minus construct pA10-CAT-2. Thus, we conclude that both mutants have a defect in a vital function, most likely within the SV40 activator. Mutants 18, 43, 44, and 45 all yield CAT activity, although to a variable extent. It should be noted that the low activity of mutant 43 is also reflected in the plaque assay and in the replication assay.

We subsequently compared the data from the replication and the CAT assays. The amount of superhelical DNA present in CV-1 cells 120 hours after transfection with each mutant DNA was quantified by densitometry (see Fig. 4). We assigned 100 arbitrary units to the amount of superhelical DNA (form I, FoI) present in CV-1 cells transfected with the single-repeat control viral DNA and calculated the amount of mutant FoI DNA accordingly.

Similarly, 100 arbitrary units were assigned to the amount of conversion of chloramphenicol to acetylchloramphenicol with the single-repeat plasmid 1-3 (Fig. 5, right), and again the relative CAT conversion rate of each mutant was calculated. We found an excellent correlation between the amount of superhelical DNA and the amount of CAT enzyme (data not presented; Fig. 3B) in these assays.

Identification of Critical Nucleotides

It is of obvious interest to define the role of individual nucleotides that appear to be involved in the function of the SV40 enhancer. Close inspection of the exchanged nucleotides (see Fig. 3) of mutants 43, 44, 45, and 47 indicates that most C → T transitions occurred around the Eco RII site and downstream towards the Sph I site. None of these mutants completely abolished the functional activity of the 72-bp repeat. However, to varying degrees, all mutants showed a decrease in overall gene expression. We therefore conclude that the region containing most of the C → T transitions seems to contribute to the total activity, although none of the exchanged nucleotides completely eliminates the function. A number of recent studies are significant in relation to this observation. Using the pA10-CAT-2 construct as a control (Fig. 6A), we have shown that the 23 nucleotides at the 3′ terminus of the repeat do not provide enhancer function. Even the 56 nucleotides from the 3′ end of the 72-bp repeat do not mediate functional enhancer activity (*13*). In contrast, Contreras and coworkers (*34*) found that deletion mutants (dependent on the mutant) containing only 13 to 26 nucleotides at the 5′ end of the 72-bp repeat still retain viability, although it is greatly reduced. Thus, the region between the Eco RII site and the Sph I site probably does not include an absolutely essential element of the enhancer.

Close inspection, however, of the nucleotide sequence exchanges of the remaining mutants 5, 11, and 18 can help to define essential nucleotides. Mutants 5 and 11 are completely defective in gene expression, indicating that in these cases, the enhancer function is abolished. If the deletion mutants (*34, 35*) are taken into consideration, only three nucleotides within the 72-bp repeat to the left of the Eco RII site (see Fig. 3A) are exchanged in the case of 5 and only one nucleotide in the case of 11. The G → A transition in mutant 11 is identical to one exchange in mutant 5. As demonstrated by mutant 18, the first three G nucleotides within the 72-bp repeat may add to the overall activity but are not absolutely essential. Therefore, a crucial core element could include the nucleotides T*G*GAAAGT, where the first G (italics) seems to be absolutely required.

In an attempt to find similar sets of nucleotides in known enhancer elements, we compared this core sequence with other activator sequences (Fig. 7). The comparison revealed that each enhancer element harbors certain nucleotides also characteristic of the SV40 core element (see Fig. 7). The underlined nucleotides demonstrate short stretches of sequences which could potentially be shown to have similar properties. A summary of the related sequences is shown at the bottom of Fig. 7. The fact that no extensive homologies have been detected (*36*) may indicate that the sequence requirements for enhancer function are different in different cells (*11, 17*).

Conclusion

Enhancer or activator sequences are transcriptional controlling elements that exhibit some remarkable features. They seem to be absolutely essential for the expression of some viral genes (*9, 16*) and furthermore drastically increase the transcriptional activity of certain cellular

genes (*11–13*). They mediate their activity relatively independent of distance and orientation with respect to a given promoter (*10, 11*) and in a host-specific manner (*17*). These elements provide an extremely active promoter function that can be utilized for the expression of interesting genes in eukaryotes and can help to decipher central questions concerning the control of transcriptional initiation. We constructed a series of multiple point mutants in an attempt to define the core nucleotides required for the functional activity. This led to a spectrum of mutants with nucleotide transitions in different regions of the 72-bp repeat. We were able to correlate the particular functional change with changes of specific nucleotides. From the analysis of seven independently isolated mutants, we arrived at the following conclusions: All mutants carrying either C → T or G → A transitions are, to different degrees, "down mutations." No "up mutation" was isolated when the nonselective mutagenesis procedure was followed (*21*). Two groups of mutants can be distinguished. The first class has major exchanges around the Eco RII site and toward the Sph I site (see Fig. 3). This group includes only viable mutants, suggesting that an element essential for gene expression does not reside in this region. Interestingly, the second group with major exchanges at the 5′ end of the 72-bp repeat completely abolishes enhancer activity. The essential nucleotide (italics) is most likely located within a TGGAAAGT sequence. This block of nucleotides is flanked at the 5′ end by mostly guanosines (GGTG) and at the 3′ end by four cytidines, which can be exchanged without complete elimination of functional activity (Fig. 3). The comparison of this stretch of nucleotide sequences to other known enhancer elements reveals a certain degree of homology (see Fig. 7). In addition, sequences comprising a bovine papilloma enhancer region (*37*), and a potential enhancer located within the human genome (*38*) also show homology with these core nucleotides.

Speculation on the process by which the individual mutants decrease functional activity is limited by our lack of understanding of the mechanism of enhancement. However, we recently demonstrated an involvement of host-specific factors in the functional activity of individual enhancers (*17*). Thus, one possible hypothesis is that the activator provides binding sites for cell specific regulatory proteins (for example, σ-factor-like). If crucial core sequences, such as the ones in mutants 5 and 11, are altered, no binding of this putative protein would occur, whereas interaction to a lesser extent could still be possible if nonessential nucleotides are replaced. Alternatively, specific secondary or tertiary structures could be required for the interaction of the enhancer with cellular factors; this alteration of nucleotides essential for the formation of a particular structure may eliminate both the structure and its function.

It is remarkable that sequences around −60 to −70 in the major late promoter of adenovirus 2 (*39*) and at −75 to −85 in the promoter of rabbit β-globin (*40*) show some surprising sequence similarity to the SV40 enhancer core region. In case of rabbit β-globin, it has been shown that certain nucleotide exchanges in this region drastically reduces the in vivo transcriptional activity (*41*). It remains to be determined whether these critical upstream regions of both genes carry a functional activity similar to the SV40 enhancer sequence.

References and Notes

1. J. Tooze, Ed., *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980).
2. W. Fiers *et al.*, *Nature (London)* **273**, 113 (1978).
3. V. B. Reddy *et al.*, *Science* **200**, 494 (1978).
4. H. Van Heuverswyn and W. Fiers, *Eur. J. Biochem.* **100**, 51 (1979).
5. T. Shenk, *Curr. Top. Microbiol. Immunol.* **93**, 25 (1981).
6. P. K. Ghosh, P. Lebowitz, R. J. Frisque, Y. Gluzman, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 100 (1981); C. Benoist and P. Chambon, *Nature (London)* **290**, 304 (1981); D. J. Mathis and P. Chambon, *ibid.*, p. 310.
7. R. Grosschedl and M. I. Birnstiel, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1432 (1980); S. L. McKnight, E. R. Gavis, R. Kingsbury, R. Axel, *Cell* **25**, 385 (1981); D. Benoist and P. Chambon, *Nature (London)* **290**, 304 (1981); P. Mellon, V. Parker, Y. Gluzman, T. Maniatis, *Cell* **27**, 279 (1981); J. A. Hassell *et al.*, in *Eucaryotic Viral Vectors* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), pp. 71–77; M. Fried and E. Ruley, in *ibid.*, pp. 67–70.
8. S. L. McKnight and R. Kingsbury, *Science* **217**, 316 (1982).
9. P. Gruss, R. Dhar, G. Khoury, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 943 (1981); C. Benoist and P. Chambon, *Nature (London)* **290**, 304 (1981).
10. P. Moreau, R. Hen, R. Everett, M. P. Gaub, P. Chambon, *Nucleic Acids Res.* **9**, 6047 (1981).
11. M. Lusky, L. Berg, M. Botchan, in *Eucaryotic Viral Vectors* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), pp. 99–107; S. E. Conrad and M. Botchan, *Mol. Cell. Biol.* **2**, 969 (1982).
12. M. R. Capecchi, *Cell* **22**, 479 (1980).
13. J. Banerji, S. Rusconi, W. Schaffner, *ibid.* **27**, 299 (1981).
14. B. Levinson, G. Khoury, G. Vande Woude, P. Gruss, *Nature (London)* **295**, 568 (1982).
15. E. H. Chang, R. W. Ellis, E. M. Scolnick, D. R. Lowy, *Science* **210**, 1249 (1980); D. G. Blair, M. Oskarsson, T. G. Wood, W. L. McClements, P. J. Fischinger, G. F. Vande Woude, *ibid.* **212**, 941 (1981); W. S. Hayward, B. G. Neel, S. M. Astrin, *Nature (London)* **290**, 475 (1981); G. S. Payne, J. M. Bishop, H. E. Varmus, *ibid.* **295**, 209 (1982).
16. J. deVilliers and W. Schaffner, *Nucleic Acids Res.* **9**, 6251 (1981); C. Tyndall, G. LaMantia, C. M. Thacker, J. Favaloro, R. Kamen, *ibid.*, p. 6231.
17. L. A. Laimins, G. Khoury, C. Gorman, B. Howard, P. Gruss, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6453 (1982).
18. M. Katinka, M. Yaniv, M. Vasseur, D. Blangy, *Cell* **20**, 393 (1980); M. Katinka, M. Vasseur, N. Montreau, M. Yaniv, D. Blangy, *Nature (London)* **290**, 720 (1981); F. K. Fujimura, P. L. Deininger, T. Friedmann, E. Linney, *Cell* **23**, 809 (1981); K. Sekidawa and A. J. Levine, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6556 (1980).
19. R. Shapiro, B. Bravermann, J. B. Louis, R. E. Servis, *J. Biol. Chem.* **248**, 4060 (1973); R. Shapiro, B. J. Cohen, R. E. Servis, *Nature (London)* **227**, 1047 (1970); K. Kai, T. Tsuro, H. Hayatsu, *Nucleic Acids Res.* **1**, 889 (1974).
20. D. Shortle and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2170 (1978).
21. H. Weiher and H. Schaller, *ibid.* **79**, 1408 (1982).
22. R. Herrmann, K. Neugebauer, E. Pirkl, H. Zentgraf, H. Schaller, *Mol. Gen. Genet.* **177**, 231 (1980).
23. These clones were provided by R. Herrmann.
24. P. Gruss, N. Rosenthal, M. König, R. W. Ellis, T. Y. Shih, E. M. Scolnick, G. Khoury, in *Eucaryotic Viral Vectors* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), pp. 13–17.
25. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
26. R. Tjian, *Cell* **13**, 165 (1978); D. Shalloway, T. Kleinberger, D. M. Livingston, *ibid.* **20**, 411 (1980); R. McKay and D. DiMaio, *Nature (London)* **289**, 810 (1981); K. W. C. Peden, J. M. Pipas, S. Pearson-White, D. Nathans, *Science* **209**, 1392 (1980).
27. M. Lusky and M. Botchan, *Nature (London)* **293**, 79 (1981).
28. R. M. Myers and R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6491 (1980).
29. B. Hirt, *J. Mol. Biol.* **26**, 365 (1967).
30. C. Gorman, L. Moffat, B. Howard, *Mol. Cell. Biol.* **2**, 1044 (1982).
31. W. Shaw, *Methods Enzymol.* **53**, 737 (1975); L. Robinson, R. Seligohn, S. Lerner, *Antimicrob. Agents Chemother.* **13**, 25 (1978).
32. L. Laimins and P. Gruss, unpublished data.
33. C. Gorman provided the pSVO plasmid.
34. R. Contreras, D. Gheysen, J. Knowland, A. van de Voorde, W. Fiers, *Nature (London)* **300**, 500 (1982).
35. J. Banerji, S. Rusconi, W. Schaffner, *Cell* **27**, 299 (1981).
36. P. Gruss, R. Dhar, J. Maizel, G. Khoury, in *Genes and Tumor Genes* (Raven, New York, 1982), pp. 39–47.
37. H. Weiher and M. Botchan, personal communication.
38. P. Steele, personal communication.
39. C. Benoist, K. O'Hare, R. Breathnach, P. Chambon, *Nucleic Acids Res.* **8**, 127 (1980).
40. A. Efstratiadis *et al.*, *Cell* **21**, 653 (1980).
41. C. Weissmann, personal communication.
42. R. Herrmann, unpublished observations.
43. B. Zink and H. Schaller, unpublished observations.
44. C. P. Gray, R. Sommer, C. Polke, E. Beck, H. Schaller, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 50 (1978).
45. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
46. This plasmid was provided by C. Gorman.
47. F. Graham and A. van der Eb, *Virology* **52**, 456 (1973).
48. V. W. Brockman and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 942 (1974).
49. R. Dhar, W. L. McClements, L. W. G. Enquist, G. Vande Woude, *ibid.* **77**, 3937 (1980).
50. I. Seif, G. Khoury, R. Dhar, *Cell* **18**, 963 (1979); N. Rosenthal, personal communication.

51. We thank H. Schaller in whose laboratory the construction of the mutants was done, R. Herrmann for fd SV40 recombinant phages; G. Khoury and C. Liu for helpful comments; and M. Kessel, L. Laimins, R. Muschel, and S. Conrad for discussions. Supported in part by a grant from the Deutsche Forschungsgemeinschaft to H. Schaller (Forschergruppe Genexpression, Heidelberg, F.R.G.).