
An enhancer sequence from bovine papilloma virus DNA consists of two essential regions

H.Weiber and M.R.Botchan

Department of Molecular Biology, University of California, Berkeley, CA 94720, USA

Received 9 January 1984; Accepted 30 January 1984

ABSTRACT

A comprehensive structural analysis of an enhancer sequence from bovine papilloma virus DNA is presented based on the construction and functional analysis of 20 mutant derivatives. The results, obtained in CV-1 tissue culture cells, show that this enhancer is a small genetic element -- only 40 bp in length -- that contains two essential regions. The two regions exhibit homology to each other and to DNA fragments from other viral genomes that also act as enhancers in the assay used. However, there is a certain latitude in the sequences that have enhancer activity in CV-1 cells, even within the critical regions. The results are discussed with respect to the model that enhancers are binding sites for tissue specific transcription factors. The formation of Z-DNA might be involved in the enhancement process. However, single base pair transitions in an 8 base pair stretch of alternating purines and pyrimidines within the BPV enhancer which conserve this pattern destroy enhancer function.

INTRODUCTION

Gene expression in eukaryotes can be regulated at the level of initiation of RNA synthesis at promoters, the sites of specific interaction between RNA polymerase and DNA. Detailed mutational analyses of different promoters for RNA polymerase II allowed the definition of a eukaryotic promoter as a tripartite structure spanning approximately 120 base pairs just upstream of the RNA initiation site (1, 2). For efficient expression of many genes in mammalian cells, however, additional cis acting DNA elements, termed enhancer sequences, are required. These elements are signals that enhance the rate of transcription initiation at a given promoter in a given genetic environment (3-5). They are unusual in that although they are cis acting, they occur and function in nature in different positions and orientations relative to the promoter they act upon (6-12). Furthermore, when present on plasmids, they can be used to activate transcription from heterologous promoters (3, 13). Enhancers have been found in viral genomes such as SV40 (3, 6, 8), Polyoma (4), Bovine papilloma virus (10), retroviral LTRs (14, 15) and also in cellular DNA (11, 12,

16). In general, enhancers display cell type specificity; that is, they function more efficiently in some cells than in others (10-12, 15-17).

Enhancer function can be assayed directly or indirectly in a number of ways. First, gene expression can be directly monitored by measuring the levels of RNA, protein or enzymatic activity produced transiently after introduction of suitable indicator genes into cells. Second, the efficiency with which a given cell line is transformed by DNA carrying a selectable marker can be used to measure enhancer activity; if expression and not integration of the marker DNA into the chromosome is the limiting step in this process (18, 19).

Following the discovery of enhancer elements in many different systems, the question arose whether conserved sequences could be identified that account for the enhancing effect. Comparison of the DNA sequences characterized as enhancer elements in different systems revealed little homology, suggesting that the sequence requirements differ from cell type to cell type (15). However, on the basis of both sequence comparisons between different enhancers and a functional analysis of mutants within the SV40 enhancer, a so-called core sequence has been identified as a characteristic feature of an enhancer sequence (20). This octamer sequence, GTGG A/T^A/T^A/T G (or C T/A^T/A^T/A CCAC in the complementary DNA strand), has been found to be quite conserved -- 7 out of 8 nucleotides -- between the viral enhancer elements of Polyoma, SV40, Bovine papilloma virus, Moloney sarcoma virus and the lymphocyte specific cellular enhancer sequence within an immunoglobulin heavy chain gene (11, 12).

The aim of this work was a comprehensive structural analysis of an enhancer sequence that had been previously identified on the 8 Kb genome of Bovine papilloma virus 1 (BPV-1). This element is located on a 60 bp restriction fragment 3' to a gene cluster that is expressed early after infection of bovine epithelial cells as well as in rodent tissue culture cells morphologically transformed by the virus (21, 22). In vivo and in mouse C127 tissue culture cells, the viral DNA is stably maintained as a multicopy nuclear plasmid (23). Expression of the viral early genes is dependent upon the presence of the enhancer sequence and required for transformation. Deletion of the enhancer results in a loss of transforming activity, and reinsertion of the 60 bp fragment in different orientations and sites on the genome fully restores this activity (10). Furthermore, this element, as does an enhancer element from Polyoma DNA, substitutes for the SV40 enhancer function in an indirect assay for SV40 T antigen expression in simian CV-1 cells (see below and ref. 10). Finally, in a stable transformation assay, it enhances the efficiency of

transformation of Rat 2 Tk⁻ to a Tk⁺ phenotype when linked to the Herpes simplex thymidine kinase gene. In order to precisely define the nucleotides essential for its function, we have constructed and tested a series of base substitution mutants in the BPV enhancer sequence.

MATERIALS AND METHODS

Replication Assay

10 ng of plasmid DNA were transfected onto 1-2 x 10⁵ CV-1 cells in a 60 mm petri dish at a confluency of 20% to 50% using a modification of a procedure described by Sompayrac and Danna (24). The cells were sequentially washed once with PBS and Tris-buffered saline before the DNA was added in 200 μ l of 0.5 mg/ml DEAE dextran in Tris buffered saline. After an incubation of 30 min at 37°C the DNA solution was aspirated and the cells were again washed once with Tris buffered saline and once with PBS. Subsequently, DMEM (Dulbecco's modified Eagle medium) containing 10% fetal calf serum was added. 1, 24, and 48 hr after transfection the cells were lysed, and low molecular weight DNA was prepared as described by Hirt (25). The equivalent of 1/4 of a plate was analyzed subsequently by agarose gel electrophoresis and Southern blotting (26).

Plasmid Constructions

All molecular cloning procedures were done according to Maniatis et al. (27). *E. coli* strains HBl01, MM294, GM113 (dam⁻, ref. 28) and K58 (ung⁻, ref. 29) were used as host strains for bacterial plasmids.

Mutagenesis Procedure

5 μ g of pLS plasmid DNA prepared from *E. coli* GM113 cells was cleaved with the restriction endonuclease ClaI and subsequently digested with 30 units S1 nuclease (Miles Laboratories) in high salt buffer as described by Berk and Sharp (30) for 15 min at 22°C. The reaction was stopped by addition of EDTA to a final concentration of 25 mM and the sample was then phenol extracted, desalted over Sephadex G100, ethanol precipitated and dried. 2 μ g of this DNA were then mixed with an equimolar amount of StuI cut p1903 DNA (see Fig. 3) in 100 μ l TEN buffer (10 mM Tris pH 8, 1 mM EDTA, 100 mM NaCl). The mixture was denatured by heating it to 100°C for 10 min and subsequently allowed to renature by cooling it down slowly (>4 hr) to room temperature. A sample of this hybrid was then tested for the presence of DNA comigrating with open circular plasmid DNA by agarose gel electrophoresis. If the fraction of open circles characterizing the heteroduplexes between p1903 DNA and the respective pLS DNA was more than 20% of the total DNA, the hybrid was treated with a 4 M

sodium bisulfite reaction mixture for 8 hr following a procedure described by Shortle and Nathans (31). After mutagenesis the DNA was transfected into *E. coli* K58 (*ung*⁻) cells. Since this strain transforms at least 100 times less efficiently than *E. coli* HB101, approximately 500 ng of DNA were necessary in each transfection reaction in order to get a sufficient number of transformants. Single colonies were picked, and plasmid DNA was prepared and tested in the replication assay. To assure the clonal character of the putative enhancer mutants and the intactness of the SV40 early region, the plasmids were recombined *in vitro* with a wild type SV40 early region using the enzymes SphI and BamHI (see Fig. 1). However, original and recombined plasmids of each mutant did not differ in their behavior in the replication assay, indicating not only the clonality of the mutants in the original K58 cells but also the site specificity of the mutagenesis procedure.

DNA Sequence Analysis

The DNA sequence analysis was carried out using either the base specific chemical cleavage method (32) or the chain termination method (33). In the latter case, dideoxynucleotides (PL Biochemicals) were used as chain terminators and a synthetic 13 mer oligonucleotide (kindly provided by Peter Seeburg, Genentech, Inc.) representing the nucleotides 4310 to 4322 of pBR322 (34) served as primer for DNA synthesis on NaOH denatured plasmid templates.

RESULTS

To screen different restriction fragments or mutants thereof for enhancer activity, we used an indirect assay for SV40 T antigen expression which has been previously described (10). Briefly, plasmids that contain the SV40 origin of replication and an intact early region coding for T antigen are able to replicate in monkey kidney CV-1 cells if the sequences required for T antigen expression are provided. The plasmid pJYΔSph -- shown in Fig. 1 -- carries a SphI-BamHI fragment containing the origin of replication, the A gene coding for the T antigens and an inactive early promoter. It lacks a functional enhancer sequence essential for A gene expression and is therefore unable to replicate after transfection into CV-1 cells. In COS7 cells (35), where T antigen is provided in trans at levels sufficient to support replication, this defect can be overcome (10). We have shown in an earlier study that insertion of enhancer sequences from different sources (SV40, BPV) into different sites on the plasmid (SalI, BamHI) activates the early promoter and thus allows for plasmid replication in CV-1 cells (10). This phenomenon is independent of the orientation in which the enhancer sequence is inserted. A

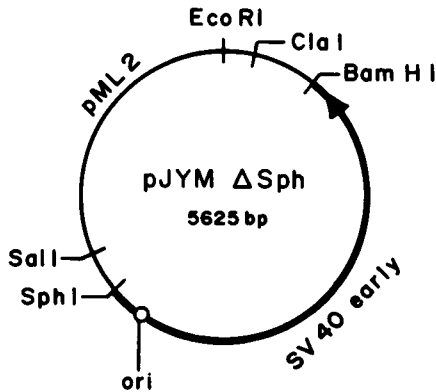


FIGURE 1. Physical map of the plasmid pJYMΔSph (10). It consists of a plasmid component conferring ampicillin resistance and the SV40 early region (thick line) fused at the BamHI and SphI sites shown. The BamHI and the EcoRI site were used to insert the enhancer elements to be tested.

typical assay involves transfecting 10 ng of plasmid DNA onto $1-2 \times 10^5$ CV-1 cells using DEAE Dextran as facilitating agent (24). The presence of plasmid DNA is monitored at 1, 24, and 48 hr after transfection by gel electrophoresis and Southern blotting of Hirt supernatants (25). As illustrated in Fig. 4 for different enhancer mutants described below, Form I DNA is barely detectable after 24 hr, indicating degradation of most of the input. If T antigen is expressed, however, the few persisting molecules begin replicating to levels above input, as detected after 48 hr. Thus, the absence of Form I DNA after 48 hr indicates a defect in T antigen expression.

The assay described here is not a direct measurement of gene expression, and we do not know the precise relationship between the amount of replicated DNA and enhancer strength. This relationship might be influenced by autoregulation of T antigen synthesis at the level of RNA synthesis (36-38) and replication itself might obscure an accurate assessment of gene expression per template. Nevertheless, the assay provides a fast, simple and sensitive measure of enhancer function in CV-1 cells.

Comparison of Different Viral Enhancers

We have used the replication assay to define and compare sequences from different viral genomes for their ability to enhance gene expression in CV-1 cells. Two different SV40 fragments were tested. A fragment spanning nucleotides 186-273 on the SV40 genome (see Fig. 2) was positive in the assay when cloned in either orientation into the BamHI site of pJYMΔSph using BamHI linkers. The fragment contains most of the 72 bp element that is repeated in wt

| | | ENHANCER FUNCTION |
|-------------------------|--|----------------------|
| SV 40 186-273 | TGACTAATTGAGATGCATGCTTTGCATACCTTCGCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACATTCACAGC ACTGATTAACCTCTACGTACGAAACGTATGAAGACGGACGACCCCTCGGACCCCTGAAAGGTTGGGATTGACTGTGTGAAGGTGTCG | YES |
| SV 40 200-273 | GCATGCTTTGCATACCTTCGCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACATTCACAGC CGTACGAAACGTATGAAGACGGACGACCCCTCGGACCCCTGAAAGGTTGGGATTGACTGTGTGAAGGTGTCG | NO |
| POLYOMA A2 5131-5264 | CTGGCCGTGCGACATCCTCTTTTAATTAGTTGCTAGGCAACGCCCTCCAGAGGCGGTGTGGTTTTGCAAG GACCGGCACGCTGTAGGAGAAAATTAATCAACGATCCGTTGCGGGAGGTTCTCCGCACACCAAAAACGTTCT AGGAAGCAAAAGCCTTCCACCCAGGCCCTAGAATGTTTCCACCCAATCATTACTATGACAACAG TCCTTCGTTTTCCGAGAGGTTGGTCCGGATCTTACAAAGGTGGGTTAGTAATGATACTGTTGTC | YES |
| BPV 4391-4454 | GATCACCGAGGTACACACCCTCCGAGCAGCAGGTTCCACATCATCGCTTGCATCAATAGGATC CTAGTGGCTCCATGTGTGGTGGGCTCCGATCTTACAAAGGTGGGTTAGTAATGATACTGTTGTC | YES |

FIGURE 2. DNA sequence and function of fragments tested for enhancer activity in CV-1 cells using the replication assay. Regions of homology between the fragments from three different viral genomes are emphasized by brackets. The nucleotide numbering is according to van Heuverswyn and Piers (39) for SV40, de Villiers and Schaffner (4) for polyoma and Chen *et al.* (22) for BPV.

SV40 and flanking sequences up to a PvuII site (39). If this fragment is shortened by cleavage with SphI at position 200, it loses its enhancer function. A 134 bp fragment from Polyoma A2 DNA (positions 5131-5264; see Fig. 2) that had been identified as carrying enhancer function by de Villiers and Schaffner (4) also supports replication when cloned in either orientation between EcoRI and BamHI of pJYMASph. Finally, a 60 bp Sau3A fragment from Bovine papilloma virus cloned in either the BamHI site or between BamHI and EcoRI (see plasmid p1903 in Fig. 3) also allows for replication via enhancement of early gene expression. Comparison of the sequences of these three functionally analogous fragments revealed two 8 bp stretches of homology separated by a 13 bp to 16 bp long spacer region, as shown by brackets in Fig. 2. These regions of homology overlap with the core sequence, C T/A^T/A^T/A CCAC, that was previously proposed as a characteristic feature of enhancer sequences (20). However, at least in SV40, the homologous sequences alone cannot account for enhancer function, since a deletion of the nonhomologous region from pos. 186-199 destroys enhancer activity (see Fig. 2).

Mutagenesis of the BPV Enhancer

After we located the enhancer activity on the 60 bp Sau3A restriction fragment of BPV DNA described above, we were interested in identifying the nucleotides critical for the enhancer function within this fragment. Following an experimental design developed by McKnight and Kingsbury (1), a set of linker scanning mutants was constructed. As a first step, two sets of exon-

| | Eco RI | Sau 3A | Bam HI |
|-----------|-------------------|---|--|
| p 1903 wt | AATTCTGCAGGTGGATC | ACCGAGGTACACACC | ACTCCGAGCAGCAGGGTCCACATCATCGCTTGCAATAGGATC |
| p LS 835 | AATTCTGCAGGTGGATC | ACCGAGG <u>ATCGATCG</u> CACTCCGAGCAGCAGGGTCCACATCATCGCTTGCAATAGGATC | |
| p LS 1943 | AATTCTGCAGGTGGATC | ACCGAGGTACACACC <u>ATCGATGG</u> CAGCAGGGTCCACATCATCGCTTGCAATAGGATC | |
| p LS 947 | AATTCTGCAGGTGGATC | ACCGAGGTACACACC <u>ACTCCGAGCAGCAGGGTCCACATCATCGCTTGCAATAGGATC</u> | |
| p LS 2339 | AATTCTGCAGGTGGATC | ACCGAGGTACACACC <u>ACTCCGAGCAGGGTCCATCGATGG</u> CTTGCAATAGGATC | |

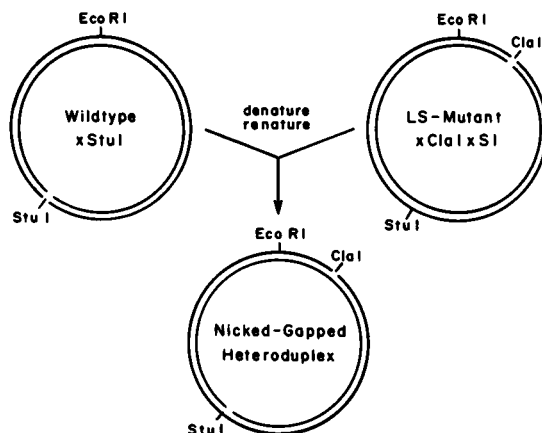


FIGURE 3. Schematic outline of the site-specific mutagenesis procedure. Top: Nucleotide sequence of the EcoRI-BamHI fragment from p1903 and the four linker scanning (LS) mutants used. The enhancer is localized between Sau3A and the BamHI site in p1903. The linker sequence in the LS mutants is underlined.

Bottom: Scheme for the generation of heteroduplex molecules specifically single stranded in preselected sites as required for the single strand specific bisulfite mutagenesis protocol described in experimental procedures. The site selection is done by the choice of the appropriate LS mutant used for heteroduplex formation with p1903.

lease deletion mutants were generated from either the EcoRI site or the BamHI site of plasmid p1903 (see Fig. 3), and ClaI linkers were added at the end-points of the deletions. In order to map the size of the deletions, the resulting mutant plasmids were subjected to sequence analysis. In a second step, deletion mutants from both ends were recombined *in vitro* using the ClaI sites. In this process, mutants were chosen that complement each other so that the resulting plasmids carry the ClaI linker as an internal substitution within the original sequence (for example see plasmids pLS835 and pLS1943 in Fig. 3). If within a collection of deletion mutants such complementary mutants are not available, imperfect combinations can be used to produce mutants that carry small deletions or insertions in addition to the linker substitution (see pLS947 and pLS2339 in Fig. 3). The deletion mutants, as

well as the linker scanning mutants, were subsequently analyzed for enhancer activity in the replication assay. A total of 24 deletions were tested, and the two smallest functional deletion mutants from either side, pD10 and pD37 (see Fig. 5), allowed for the further minimalization of the region essential for enhancer function. However, all four linker scanning mutants located in this region (pLS835, pLS1943, pLS947, pLS2339; see Fig. 3) were negative in the replication assay. The replication defect was due to insufficient T antigen expression because in COS7 cells the plasmids replicated to levels equivalent to the wild type plasmid (data not shown). These results indicated either that the entire sequence covered by the linker scanning mutants is essential for enhancer function or that there is an inhibitory effect of the linker sequence itself. The latter seemed to be more likely because a polymer of the ClaI linker sequence inserted into the ClaI site of functional deletion derivatives such as pD10 inactivated such mutants (data not shown). Furthermore, it was noted that the tetranucleotide CATC, which is part of the linker sequence, occurs three times within the Sau3A fragment. This suggested the possibility of an abortive interaction between a putative protein and the linker sequence due to its resemblance to essential parts of the enhancer sequence. In conclusion, these considerations indicated the need for a more subtle mutagenesis method, which would allow for the introduction of a few or even single base changes within the enhancer element.

This was accomplished by using the linker scanning mutants to direct chemical bisulfite mutagenesis to specific sites in vitro. The procedure, an extension of heteroduplex mutagenesis protocols developed by Peden and Nathans (40) and Kalderon et al. (41) is schematically outlined in Fig. 3 and described in detail in the Experimental Procedures. Briefly, a linker scanning mutant is cut with the linker enzyme ClaI and then treated mildly with nuclease S₁ to delete the 5' overhanging ends at the cleavage site. This linear molecule is hybridized to a wild type molecule that has been linearized by cleavage with another single cut enzyme (in this case StuI). The resulting heteroduplex carries a small single-stranded region of wt DNA where the linker scanning mutant was opened and deleted. This region is the preferred target for the single-strand specific mutagen sodium bisulfite, which very efficiently converts C residues to U residues in vitro (31, 42). Mutagen treated heteroduplexes establish C/G to T/A transition mutations after being transfected into uracil repair deficient E. coli strains (ung⁻). Linear homoduplexes, also formed in this procedure, transform with much lower efficiency than the heteroduplex circles. Therefore, given extensive mutagenic treatment, the

mutant yield can be as high as 100%. Another advantage of this protocol is that, as shown below, it does not require precise linker scanning mutants in order to direct mutagenesis to a specific region, but allows the use of imprecise linker insertion derivatives such as pLS947 and pLS2339 (see Fig. 3).

Applying this mutagenesis procedure to hybrids between plasmid p1903 and the pLS plasmids 835, 1943, 947, and 2339 shown in Fig. 3, a total of 96 putative mutant plasmids were isolated from four analogous mutagenesis reactions. Sequence analysis of the enhancer region in 19 plasmid isolates revealed the 18 different mutant sequences shown in Fig. 5. Five mutants carried single point mutations, seven had double mutations, three mutants were changed in three, two in four and one in five positions, respectively. As expected, all mutagenic events occurred in the target region rendered single stranded in the respective hybrid subjected to bisulfite treatment. For example, mutagenesis of the hybrid p1903xpLS835 gave rise to mutants p1 to p6, which carry base changes in and around the region that is replaced by linker sequences in pLS835. Furthermore -- with the exception of mutant p6 -- multiple base changes occur in clusters. That is, under the given experimental conditions, adjacent target sites -- as compared to scattered ones -- were mutated preferentially. In summary, as demonstrated in Fig. 5 the 18 mutant plasmids cover most of the C/G pairs (21/24) in the essential region of the enhancer defined by the deletion mutants pD10 and pD37. The clonal character of the mutant plasmids and the intactness of the SV40 sequences were confirmed by recloning the DNA fragments containing the variant enhancer elements into an unmutated parent plasmid as described in the Experimental Procedures.

Biological Activity of the Point Mutants in the Replication Assay

The mutant plasmids described above were subsequently tested in the replication assay. Fig. 4 shows the result of this comparative analysis. As previously mentioned, the absence of supercoiled plasmid at 48 hr after transfection indicates a defect in the enhancer sequence. This is the case for 12 out of the 18 mutant plasmids tested. The remaining 6 plasmids, p6, p25, p30, p52, p74 and p75 replicated and therefore must have an altered but nonetheless functional enhancer. Note that all six mutants replicated to approximately the same levels after 48 hr. Fig. 5 summarizes the structure and function of the enhancer mutants tested. There are evidently two domains within the enhancer element where mutations lead to a loss of function. Both of these regions display sequence homology to functionally equivalent DNA fragments from other viruses (see above and Fig. 2) and to the core sequence proposed by Weiher *et al.* (20). In particular, four out of five single point mutations

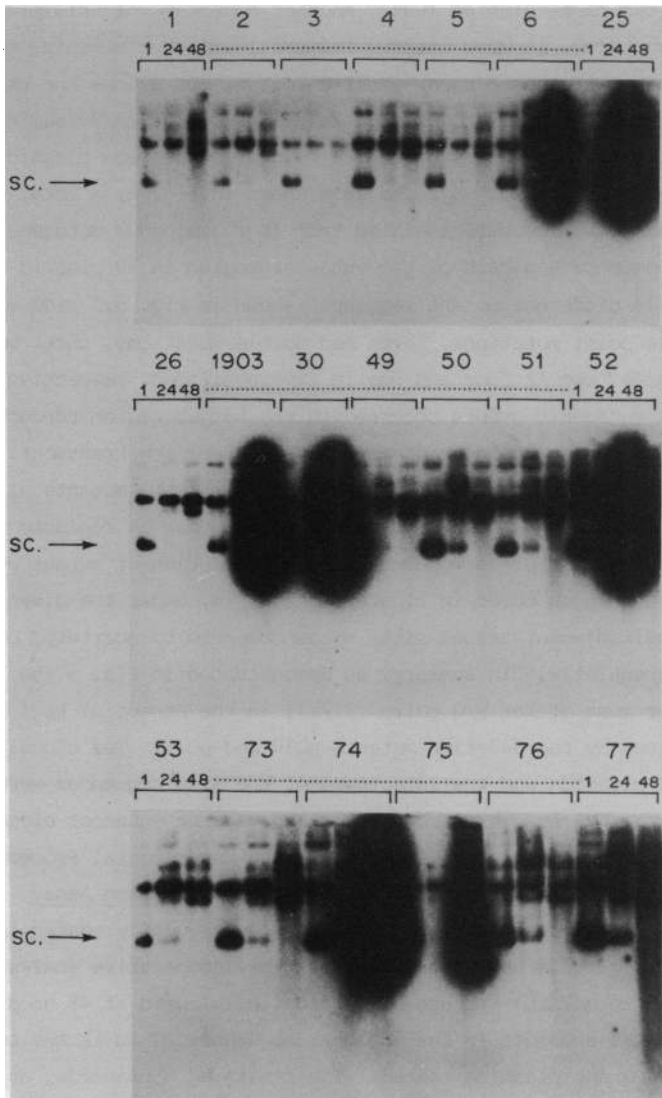


FIGURE 4. Assay for replication competence in CV-1 cells of plasmid p1903 and mutant derivatives of this plasmid. A Southern blot of low molecular weight DNA is shown, isolated from cells at different timepoints after transfection and hybridized to p1903 DNA. Samples were loaded in groups of three lanes for each plasmid, representing three time points (1, 24 and 48 hr after transfection from left to right). S.C. marks the position of the supercoiled form I DNA. Replicating plasmids are easy to discriminate by the presence of form I DNA at 48 hr after transfection (plasmid 1903 and mutants 6, 25, 30, 52, 74 75). The autoradiogram is overexposed in order to make the input DNA (time-points 1 hr after transfection) visible.

| Fragment | 4400 | 4420 | 4440 | Function |
|------------------|--|------|---------|----------|
| WT | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| D37 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 1 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 4 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 6 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 2 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 3 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 5 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 50 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 52 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 51 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 53 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 26,49 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 25 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 30 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 75 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 77 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 74 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | + |
| 76 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| 73 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATCGCTTGCATCAATAGGATC | | | - |
| D10 | GATCACCAGGTTACACACCACTCCGAGCAGCAGGGTCCACATCATAAGCTTTAATGCGTAACT | | | + |
| Enhancer Domains | GGTACACACC | CCG | CCACATC | |

FIGURE 5. Nucleotide sequence and function of 19 point mutants and two deletion derivatives of the BPV enhancer. Deviations from the original sequence in p1903 (top line) are indicated by boxes. On the bottom line the two enhancer domains, as derived from the analysis of the mutation sites and the function of the different mutants, are pointed out. D10 and D37 are deletion derivatives of p1903; D37 was deleted with exonuclease from the EcoRI site of p1903 (Figure 3) and a ClaI linker (CATCGATG) was added onto pos. 4395. D10 was deleted analogously, but from the BamHI site of p1903, and the linker was added on to pos. 4428 in this case. Plasmid or SV40 specific sequences that had also been deleted in the exonuclease reaction were subsequently restored by combining the deleted plasmid with pJYMSph sequences using the enzymes ClaI and SphI (see Fig. 1). Note that the wt. sequence in p1903 differs from the BPV sequence published by Chen *et al.* (22) in two positions (4398 and 4417, G vs. A in both cases).

(in p2, p3, p4 and p51) inactivate the function and so define essential nucleotides within the enhancer region. In addition, the nonfunctional single and multiple point mutants represented by the plasmids p1, p2, p3, p4, p5, p26, p49, p50, p51 and p53 define a functionally important region of a minimum of 16 bp (pos. 4400-pos. 4415) at the left end of the enhancer element, as

depicted in Fig. 5. This region is interrupted by a nonessential G/C pair in pos. 4411, which is indicated by the function of mutant p52. The second functional domain within the enhancer sequence (pos. 4427-pos. 4433) is defined by the mutations in the enhancer deficient plasmids p73, p77, and p76. The region between these two domains does not appear to require a defined nucleotide sequence since all G/C pairs between pos. 4418 and pos. 4425 can be mutated to A/T pairs without loss of enhancer function (see p25, p30, p75).

The simple model suggested by the data described so far is confused by two variants of the original sequence, namely p6 and p74. Both are functionally active and carry multiple point mutations. The enhancer element in p6 contains the same mutation as p4 and in addition a double mutation that characterizes p53. However, since p4 and p53 are both enhancer deficient, the p4 mutation must compensate for the p53 mutations and vice versa. Similarly, p74 carries an active but mutated enhancer element. It has a double base change in common with the enhancer deficient mutant p77 and an additional base change that is also found in p73 and p76. The latter base change, in pos. 4430, appears to compensate for the double mutation in p77 but not for the base change in pos. 4433 that is present in p73 and p76. These data, particularly the properties of p6 and p74, indicate a remarkable latitude in the sequences that function as enhancers in CV-1 cells.

DISCUSSION

Enhancers are genetically defined as cis acting DNA sequences that activate gene expression in a cell type-specific manner from remote positions with respect to the promoter they act upon. However, biochemically they are still poorly characterized, and the mechanism by which they exert their function is essentially unknown. A first step towards understanding the molecular basis of enhancement is to define the DNA sequences that are essential for the function. Comparisons of nucleotide sequences that act as enhancers in different systems do not provide a satisfactory approach to the problem, since the sequence requirements for enhancer function may differ from cell type to cell type. We have therefore addressed this problem by systematically mutating a particular enhancer sequence from BPV DNA and testing the mutants for function in CV-1 tissue culture cells.

Employing a new, very simple and versatile technique for site directed mutagenesis, we have isolated and analyzed a set of 18 mutants containing single or multiple base substitutions within a 40 bp enhancer element (pos. 4394 to pos. 4433) defined by deletion mutants. This analysis revealed that

the BPV enhancer consists of two functional domains of different sizes (minimal 16 bp and 6 bp) which are separated by a stretch of at least 8 bp that can be changed without affecting the enhancer function. Both domains are similar in sequence and homologous to sequences from SV40 and polyoma that exhibit activity in the enhancement assay used here. They are also homologous to the core sequence proposed as a characteristic feature of enhancer elements by Weiher *et al.* (20). Furthermore, the centers of the two domains are approximately 25 bp apart from one another, indicating that they face opposite directions on the DNA helix, if the DNA is in the B configuration.

Nordheim and Rich (43) have proposed an involvement of Z-DNA (lefthanded DNA) in the process of enhancement. In support of this hypothesis, they find that two 8 base pair stretches of alternating purines and pyrimidines (pos. 198 to pos. 205 and pos. 258 to pos. 265) close to or within the 72 base pair repeat of SV40 can bind a Z-DNA specific antibody *in vitro*. The results presented in Fig. 2 seem to support the idea that these regions might be essential for enhancer activity. Two different fragments from SV40 DNA were tested for enhancer function in the replication assay. One of them, spanning pos. 186 to pos. 273, contains the two regions of potential Z-DNA formation referred to above and carries a functional enhancer: the second fragment, extending from pos. 200 to pos. 273, is defective in enhancer function and has deleted two base pairs of one of the two Z-DNA regions (pos. 198 and pos. 199). In addition, within the BPV enhancer, we also find an 8 base pair long stretch of alternating purines and pyrimidines (pos. 4401 to pos. 4408) superimposed on one of the functional domains defined by mutagenesis. However, since the mutagenesis procedure used does not allow for the introduction of transversion mutations, the purine-pyrimidine pattern in all bisulfite mutants remains unchanged. Therefore, we conclude that a potential for Z-DNA formation is not sufficient for the functionality of this enhancer domain, since the analysis clearly reveals the requirement of a specific sequence in this region (see p2, p3, p4, p5). Finally, it should be pointed out that some fragments which provide enhancer function do not contain significant stretches of alternating purines and pyrimidines. These include the polyoma enhancer fragment analyzed in this study (see Fig. 2) as well as the lymphocyte specific enhancer sequence described by Banerji *et al.* (11) and Gillies *et al.* (12).

It has been proposed that enhancers are binding sites for tissue-specific transcription factors (5, 13). This model is supported by experiments that show that enhancer sequences compete for limiting "factors" after transfection *in vivo* (44). Based on the observation that the preferred site of transcrip-

tional activation is the promoter closest to a given enhancer sequence, it was proposed by Wasylyk *et al.* (5) that these factors bind the enhancer sequence and subsequently translocate to their site of action, the promoter, by linear diffusion along the DNA. Our data define the BPV enhancer as a DNA sequence of approximately 40 base pairs, thus similar in size to other eukaryotic regulatory signals, such as the binding sites for T antigen on the SV40 genome (45), that have been shown to be involved in protein nucleic acid interactions. In this sense the data are consistent with the model discussed above. We furthermore find the reiterated nature of the two functional domains intriguing with respect to how a putative protein might recognize this sequence.

The apparent latitude in the sequences that can function as enhancers in CV-1 cells is remarkable. As pointed out in the Results section, p6 carries an altered but functional enhancer sequence. It can be viewed as a compensating revertant of either p4 or p53, both of which are enhancer deficient (see Fig. 5). In particular, the single base change in p4 and the double mutation in p53 destroy the function, but all three together, as present in p6, result in a completely functional molecule. Similarly, the deficient enhancer function in the double mutant p77 can be restored by the introduction of an additional single base change as present in p74. In order to explain these data on a molecular level, one can speculate that these compensating effects reflect a resemblance of the revertants to the original sequence in their three dimensional structure, regardless of their differences in nucleotide sequence. An alternative hypothesis would postulate that, assuming that enhancer sequences are recognition sites for specific transcription factors, different transcription factors coexisting in the same cell can recognize different but similar enhancer sequences as represented by p1903, p6 or p74. To further examine the phenomenon, however, it will be necessary to test the mutants in a more direct assay. Furthermore, addressing the question of cell type specificity, it will be of interest to know whether the putative transcription factors which may exist in different cells in which this enhancer can function recognize the same structures. This can be tested in future experiments by assaying the collection of mutants presented here in different cells for enhancer activity.

ACKNOWLEDGEMENTS

H. W. thanks Dr. S. Conrad for most valuable discussions, encouragement and substantial contributions to the manuscript. We further thank Drs. H.

Schaller and E. Beck for suggestions on the mutagenesis procedure and Dr. Peter Seeburg for providing us with a sequencing primer as well as with an unpublished modification of the dideoxynucleotide sequencing protocol. Plasmids containing a SV40 enhancer fragment (pos. 114-pos. 273) and recombinant phage containing a polyoma enhancer fragment (pos. 5131-5264) were kindly provided prior to publication by Drs. S. Saragosti and J. de Villiers respectively. This work was supported by grant CA 30490 from the National Cancer Institute and grant MV-91 from the American Cancer Institute to M. B. H. W. is supported by a fellowship from the Deutsche Forschungsgemeinschaft.

REFERENCES

1. McKnight, S. L. and Kingsbury, R. (1982) *Science* 217, 316-324.
2. Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. and Weissmann, C. (1983) *Cell* 32, 695-706.
3. Banerji, J., Rusconi, S. and Schaffner, W. (1981) *Cell* 27, 299-308.
4. de Villiers, J. and Schaffner, W. (1981) *Nucl. Acids Res.* 9, 6251-6264.
5. Wasylyk, B., Wasylyk, C., Augereau, P. and Chambon, P. (1983) *Cell* 32, 503-514.
6. Benoist, C. and Chambon, P. (1981) *Nature* 290, 304-310.
7. Gruss, P., Dhar, R. and Khoury, G. (1981) *Proc. Nat. Acad. Sci. USA* 78, 943-947.
8. Fromm, M. and Berg, P. (1982) *J. Mol. Appl. Genet.* 1, 457-481.
9. Payne, G. S., Bishop, J. M. and Varmus, H. E. (1982) *Nature* 295, 209-214.
10. Lusky, M., Berg, L., Weiher, H. and Botchan, M. (1983) *Mol. Cell. Biol.* 3, 1108-1122.
11. Banerji, J., Olson, L. and Schaffner, W. (1983) *Cell* 33, 729-740.
12. Gillies, S. D., Morrison, S. L., Oi, V. T. and Tonegawa, S. (1983) *Cell* 33, 717-728.
13. Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. P. and Chambon, P. (1981) *Nucl. Acids Res.* 9, 6047-6068.
14. Blair, D. G., McClements, W. L., Oskarsson, M. K., Fischinger, P. J. and Van de Woude, G. (1980) *Proc. Nat. Acad. Sci. USA* 77, 3504-3508.
15. Laimins, L. A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982) *Proc. Nat. Acad. Sci. USA* 79, 6453-6457.
16. Conrad, S. E. and Botchan, M. (1982) *Mol. Cell. Biol.* 2, 949-965.
17. de Villiers, J., Olson, L., Tyndall, C. and Schaffner, W. (1982) *Nucl. Acids Res.* 10, 7965-7976.
18. Luciw, P. A., Bishop, J. M., Varmus, H. E. and Capecchi, M. R. (1983) *Cell* 33, 705-716.
19. Krieglner, M., Perez, C. and Botchan, M. (1983) in *Gene Expression, UCLA Symposium on Molecular and Cellular Biology, New Series, Vol. 8*, Hamer, D. and Rosenberg, M. Eds., Alan R. Liss, Inc., New York, in press.
20. Weiher, H., Konig, M. and Gruss, P. (1983) *Science* 219, 626-631.
21. Heilman, C. A., Engel, L., Lowy, D. R. and Howley, P. (1982) *Virology* 119, 22-34.
22. Chen, E. Y., Howley, P. M., Levinson, A. D. and Seeburg, P. H. (1982) *Nature* 299, 529-534.
23. Law, M.-F., Lowy, D. R., Dvoretzky, J. and Howley, P. M. (1981) *Proc. Nat. Acad. Sci. USA* 78, 2727-2731.
24. Sompayrac, L. M. and Danna, K. J. (1981) *Proc. Nat. Acad. Sci. USA* 78, 7575-7578.
25. Hirt, B. (1967) *J. Mol. Biol.* 26, 365-369.

26. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503-517.
27. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
28. Marinus, M. G. and Konrad, E. B. (1976) *Mol. Gen. Genet.* 149, 273-277.
29. Baas, P. D., van Teeffelen, H.A.A.M., Teertstra, W. R., Jansy, H. S., Veeneman, G. H., van der Marel, G. A. and van Boom, J. H. (1980) *FEBS Letters* 110, 15-20.
30. Berk, A. J. and Sharp, P. A. (1977) *Cell* 12, 721-732.
31. Shortle, D. and Nathans, D. (1978) *Proc. Nat. Acad. Sci. USA* 75, 2170-2174.
32. Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
33. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Nat. Acad. Sci. USA* 74, 5463-5467.
34. Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol.* 43, 77-90.
35. Gluzman, Y. (1981) *Cell* 23, 175-182.
36. Tegtmeyer, P. (1974) *Cold Spring Harbor Symp. Quant Biol.* 39, 9-16.
37. Alwine, J. C., Reed, S. I. and Stark, G. R. (1977) *J. Virol.* 24, 22-27.
38. Rio, D., Robbins, A., Myers, R. and Tjian, R. (1980) *Proc. Nat. Acad. Sci. USA* 77, 5706-5710.
39. van Heuverswyn, H. and Fiers, W. (1979) *Eur. J. Biochem.* 100, 51-60.
40. Peden, K.W.C. and Nathans, D. (1982) *Proc. Nat. Acad. Sci. USA* 79, 7214-7217.
41. Kalderon, D., Oostra, B. A., Ely, B. K. and Smith, A. E. (1982) *Nucl. Acids Res.* 10, 5161-5171.
42. Weiher, H. and Schaller, H. (1982) *Proc. Nat. Acad. Sci. USA* 79, 1408-1412.
43. Nordheim, A. and Rich, A. (1983) *Nature* 303, 674-679.
44. Linney, E., Donerly, S., Olinger, B., Bender, M. and Fujimura, F. K. (1983) in *Teratocarcinoma Stem Cells*, Silver, L. M., Martin, G. R. and Strickland, S. Eds., Cold Spring Harbor Conf. Cell Proliferation 10, 271-284, Cold Spring Harbor, New York.
45. Tjian, R. (1978) *Cell* 13, 165-179.