JOURNAL OF VIROLOGY, May 1990, p. 2411–2415
0022-538X/90/052411-05$02.00/0

Vol. 64, No. 5

Promoter Evolution in BK Virus: Functional Elements Are Created at Sequence Junctions

RHEA-BETH MARKOWITZ,* SARAH TOLBERT, AND WILLIAM S. DYNAN

Department of Chemistry and Biochemistry, Campus Box 215, University of Colorado, Boulder, Colorado 80309-0215

Received 13 October 1989/Accepted 8 January 1990

The archetypal strain of BK virus directed very little early gene expression compared with rearranged laboratory strains of the virus. One laboratory strain that was analyzed in detail contained newly created AP-1 binding enhancer modules spanning the junction between adjacent sequence repeats. Introduction of these sequences into the archetype activated the previously quiescent early promoter.

The genome of the human polyomavirus BK (BKV) consists of a noncoding promoter-and-enhancer region located between the early and late transcription units (1–6, 10, 13, 20, 23, 26, 30, 32). Recent biochemical and genetic studies have identified an assortment of individual genetic elements that contributes to the biological activity of this region (1–5, 10, 15, 30). One of the interesting aspects of BKV is that different strains of the virus vary considerably in the number and boundaries of direct DNA sequence repeats present in the regulatory region. A new perspective on this variability recently emerged with the suggestion that the different strains of BKV evolved from a common archetype (24, 33). This hypothesis fits nicely with available sequence data, as it is possible to envision the derivation of all known strains from the archetype as resulting from small numbers of duplication and deletion events.

The sequence of the BK(WW) strain corresponds closely to the theoretical archetype (24). It has been proposed that BK(WW) may represent the predominant form of the virus in infected individuals and that many of the other known strains of BKV represent the outgrowth of population variants during isolation in culture (24). BK(WW) was isolated by molecular cloning of viral DNA from urine and reportedly will not grow in culture (24). In this study, we show that the BK(WW) early promoter is virtually quiescent in transfection assays, whereas sequence rearrangements in other strains increase the level of early gene expression by up to 100-fold. The specific mechanism of activation was determined by a detailed analysis of one rearranged strain, BK(Dun), and was shown to reflect the formation of novel genetic elements at the junctions of direct sequence repeats. The formation of new elements at repeat junctions may provide a facile, general mechanism for the rapid evolution of new characteristics in BKV and other polyomaviruses.

Transcription experiments used DNA from the archetypal strain, BK(WW), and two rearranged strains, BK(MM) and BK(Dun). Maps of the regulatory regions and transcription factor binding sites of these different strains are shown in Fig. 1. All three strains are identical from the early mRNA start site through the first P sequence block, which ends 113 base pairs (bp) upstream of the early start site. Beyond this point, the sequence arrangements differ. BK(MM) and BK(Dun) have imperfect triplications of P and Q block sequences and deletions of Q and R block sequences (Fig. 1). Despite the rearrangements, the regulatory regions of all three strains are roughly the same size (170 to 271 bp). Each strain contains multiple binding sites for the known transcription factor NF-1 (3, 10, 15, 16, 18), and some strains contain binding sites for the transcription factors Sp1 and AP-1 (3, 15). Because all strains contain multiple binding sites for known positively acting transcription factors, it is not possible to make a priori judgments about relative transcriptional efficiency.

To determine the transcriptional activity of the different strains, we used a DNA transfection assay. The BKV regulatory region was fused in the early orientation to the chloramphenicol acetyltransferase (CAT) gene, and levels of CAT enzyme activity were measured as indicators of early RNA accumulation (9, 17). Table 1 shows a comparison of CAT expression from different BKV strains after transfection into WI-38 human embryonic lung fibroblasts. The BK(WW) regulatory region was nearly quiescent for early gene expression, whereas the regulatory regions of the laboratory strains, BK(Dun) and BK(MM), showed approximately 100-fold-greater activity. The striking difference in the level of CAT expression in the different BKV strains suggests that the upstream deletions and duplications contribute to the formation of an effective promoter-enhancer unit.

There are two possible explanations for the strain-specific differences in gene expression. Sequences inhibitory for transcription might be present in BK(WW) but deleted from the evolved strains. For example, such inhibitory sequences might be present in the R sequence block at the late side of the regulatory region. Alternatively, positively acting sequences might be duplicated or formed de novo in the rearranged strains. To discriminate between these two possibilities, reconstruction experiments were carried out in which DNA fragments isolated from the high-expression strain, BK(Dun), were inserted into the low-expression strain, BK(WW). Figure 1 shows the fragments that were used. BK(Dun) regulatory DNA was digested with the restriction enzyme *Bsu* 36I, which releases DNA fragments of 50 and 68 bp. We have previously shown, by DNase 1 footprint analysis, that each of these fragments contains single binding sites for the transcription factors AP-1 and NF-1 (15). These fragments were inserted in their natural orientation into the unique *Bsu* 36I site of pWWCAT. A 118-bp fragment containing all four protein-binding sites, obtained by partial *Bsu* 36I digestion of BK(Dun), was also inserted into pWWCAT.

Recombinants containing single and multiple copies of the

* Corresponding author.

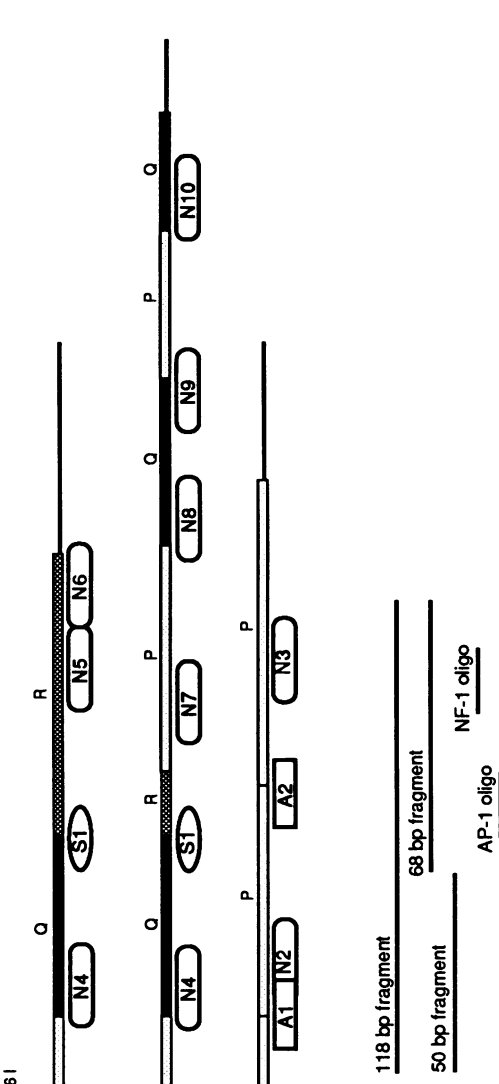

FIG. 1. Map of the regulatory regions of BKV strains and construction of recombinants. Division of sequence into P, Q, and R blocks is to aid in visualization of rearrangements (15). BK(WW) contains all three blocks, whereas BK(MM) and BK(Dun) contain partial duplications of some sequences and deletions of others. Transcription factor binding sites for previously described host cell transcription factors NF-1, AP-1, and Sp1 were determined experimentally in our previous study by using affinity-purified protein preparations (15) and are consistent with results of a competition analysis recently reported by another group (3). In our original analysis, sites N1 to N10 were protected by an NF-1-like protein, provisionally referred to as NF-BK (15). Further purification and analysis of the protein showed it to be similar, if not identical, to the previously described NF-1/CTF factor (14, 16, 22). The early mRNA start site is indicated. The BK(Dun) test fragments and synthetic AP-1 and NF-1 binding oligonucleotides were inserted into the unique *Bsu* 36I site in BK(WW), indicated by the arrow. The asymmetry of the restriction site constrains all fragments to be inserted in their natural orientations. For the AP-1 binding site, the sequence 5′-TCAGTGCATGACTCAGGGA-3′ and its complement were used; this is the natural sequence present at the A1 and A2 AP-1 binding sites (consensus recognition sequence underlined). For the NF-1 binding site, the sequence 5′-TCACATGGAATGCAGCCAAAC-3′ and its complement were used; this is the natural sequence present at the N1 and N3 binding sites (consensus recognition sequence underlined). The identities of all plasmids were confirmed by DNA sequencing.

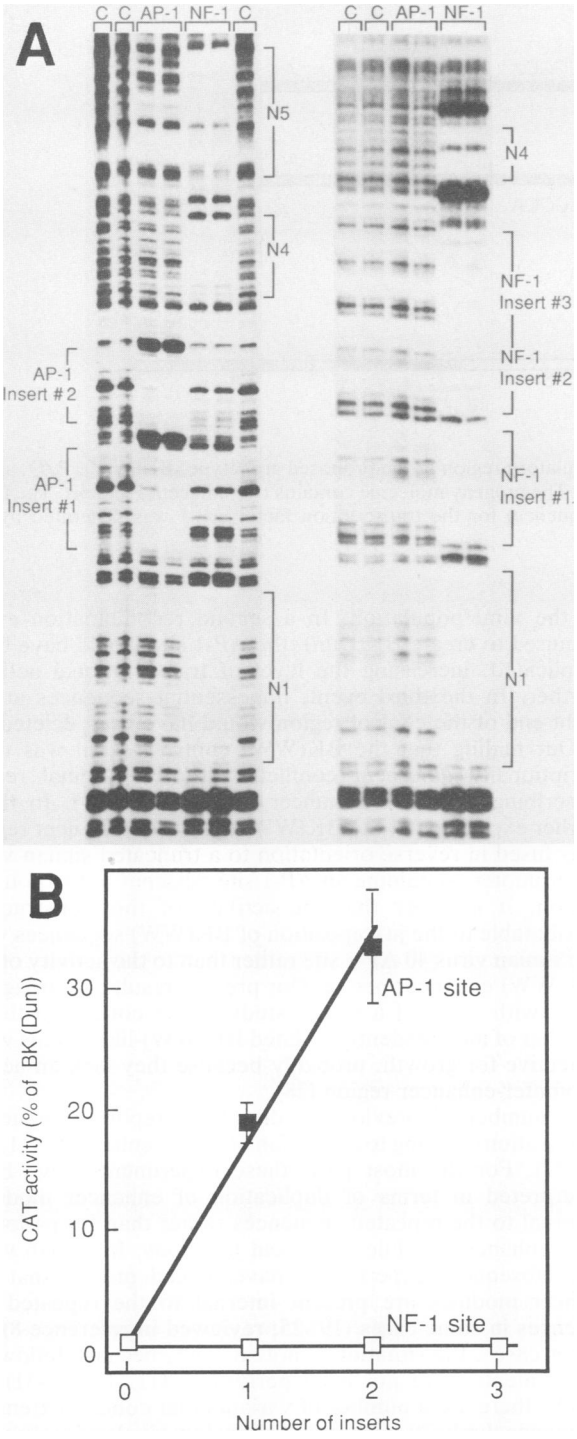

FIG. 2. Activity of pWWCAT recombinants containing synthetic AP-1 or NF-1 binding sites. Complementary oligonucleotides containing recognition sequences for AP-1 or NF-1 were inserted into the unique *Bsu* 36I site of pWWCAT as shown in Fig. 1. (A) DNase I footprinting of pWWCAT recombinants. DNase I footprinting was carried out as described elsewhere (7, 15) for the recombinants containing two copies of the AP-1 oligonucleotide (left panel) or three copies of the NF-1 oligonucleotide (right panel). A 10-μl volume of DNA affinity-purified AP-1 or NF-1 (15) was added to binding reactions where indicated. No protein extract was added for lanes marked C. Protected regions are marked by brackets. (B) CAT activity of pWWCAT recombinants. Recombinants containing one or two copies of the AP-1 oligonucleotides and one, two, or three copies of the NF-1 oligonucleotides were transfected into human WI-38 cells by using a modification of the method described by Sompayrac and Danna (27). CAT activity was assayed as described in the legend to Table 1. Error bars represent the range of observations in triplicate samples.

TABLE 1. Analysis of BKV early promoter activity in vivo[a]

BKV strain	CAT activity [% of BK(Dun)] (mean ± SD)	No. of expts
Natural		
pWWCAT	1.0 ± 1.1	4
pDunCAT	100[b]	4
pMMCAT	100 ± 13	4
Recombinants		
pW50DCAT-1	41 ± 12	3
pW68DCAT-1	23 ± 5	1
pW118DCAT-1	85 ± 25	4
pW118DCAT-2	294 ± 4	2

[a] Regulatory regions of BKV strains were subcloned in a promoterless CAT expression vector in the early orientation via a *Stu*I site conserved in the different BKV strains. The site of fusion is 53 nucleotides downstream of the BKV early RNA start site shown in Fig. 1 (4) and is proximal to the normal initiation codon for T-antigen synthesis. Recombinants were constructed as described in the text and were designated pW50DCAT-1, pW68DCAT-1, and pW118DCAT-1, where the numbers 50, 68, and 118 indicate the size of the inserted BK(Dun) fragment (Fig. 1) and the final digit indicates the number of copies of the insert. Transfection of supercoiled plasmid DNA into WI-38 human embryonic lung fibroblasts was carried out by using a modification of the method described by Sompayrac and Danna (27). Analysis of CAT enzyme activity was carried out by using the method of Nordeen et al. (17). Background was determined by assaying mock-transfected cells and was subtracted. In each experiment, duplicate or triplicate plates were assayed, averaged, and normalized to BK(Dun) activity. Results shown are normalized values from multiple independent experiments, as indicated, except for results for pW68DCAT-1, which are from duplicate plates within an experiment.

[b] Standard deviation is not applicable.

inserted fragments were tested for their ability to direct CAT activity in transfection experiments (Table 1; other data not shown). All recombinants tested showed substantially increased levels of activity. The insertion of single copies of the 50- and 68-bp fragments increased gene expression of 41 and 23% of the BK(Dun) level, respectively. Insertion of the 118-bp fragment increased BK(WW) activity to essentially the same level as BK(Dun) activity. These results strongly suggest that the high level of transcriptional activity observed for BK(Dun) is attributable to positively acting sequence elements present in both the 50- and 68-bp fragments. Our experiments show no evidence that negatively acting elements are present in the R block of BK(WW) since the recombinant containing the 118-bp insertion was rescued to nearly the BK(Dun) level of activity. However, these experiments do not formally rule out negatively acting elements that might become apparent if sequences were joined in other arrangements. In this respect, we note that other investigators have reported evidence for cell-type-specific negative elements in a different strain of BKV that was not tested in our experiments (10).

Each of the test fragments used in this experiment contained more than one binding site for known transcription factors and, in addition, could have contained regulatory elements not detected in the binding studies. To identify the precise sequences responsible for transcriptional activation, synthetic 21-bp oligonucleotides representing AP-1 or NF-1

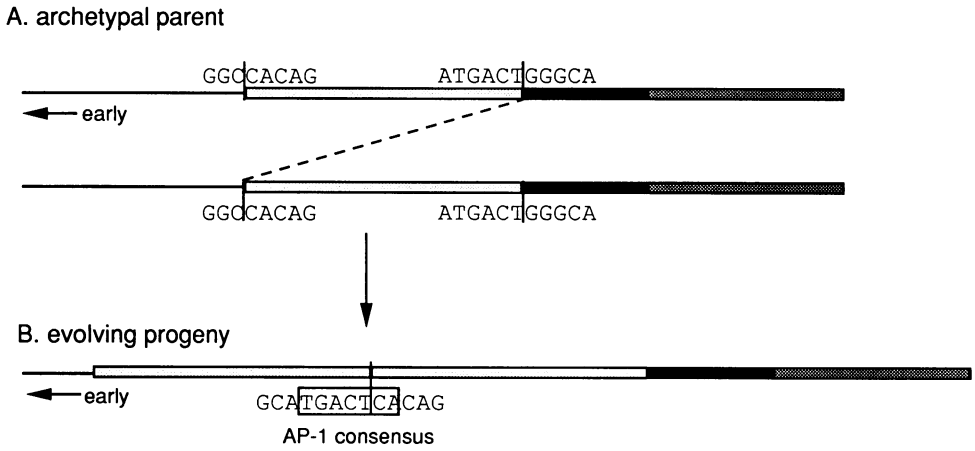

FIG. 3. Evolution of AP-1 binding sites across repeat junctions. (A) Regulatory region of the proposed archetype, BK(WW). P, Q, and R blocks are drawn as in Fig. 1. DNA sequences at junctions are indicated. (B) The progeny molecule contains two adjacent P blocks. The DNA sequence spanning the new junction is shown. A binding site (boxed sequence) for the transcription factor AP-1 was generated by the rearrangement.

protein-binding sequences from BK(Dun) were inserted into BK(WW). The method of construction and, in particular, the *Bsu* 36I site used for insertion were the same as in the previous set of experiments. DNase 1 footprint analysis was carried out for recombinants containing two synthetic AP-1 sites or three synthetic NF-1 sites (Fig. 2A). Both the naturally occurring and the synthetic protein binding sites were bound by the appropriate purified protein.

The addition of a single AP-1 binding site to BK(WW) resulted in a large increase in early gene activity (Fig. 2B). The addition of two AP-1 binding sites increased gene expression further, to 36% of the level seen with BK(Dun). When one, two, or three additional NF-1 protein binding sites were inserted, no increase in activity was seen. These experiments demonstrate that the inserted AP-1 sites can complement the sequences already present in BK(WW) to form the functional unit required for high levels of BKV early gene expression. The increase in activity cannot be due to a spacing effect alone, since insertion of NF-1 oligonucleotides of the same size showed no increase in activity. It should be noted that, although the increase in activity was clear, the absolute level of expression in templates reconstructed with synthetic oligonucleotides did not reach the same level as with natural restriction fragments. This possibly reflects the effects of transcription factor binding site spacing and context, both of which differed significantly in the oligonucleotide reconstructed strains from the pattern seen in BK(Dun).

It has been proposed that laboratory strains of BKV are derived from a common BK(WW)-like archetype (24). We suggest that the first step in the evolution of the BK(Dun) enhancer was the illegitimate recombination event diagrammed in Fig. 3, which juxtaposed two distant sequences from the regulatory region of a BK(WW)-like ancestor. Five base pairs of a consensus AP-1 site are contributed by sequences to the left of the junction, and two base pairs are contributed by sequences to the right of the junction. Affinity-purified AP-1 binds to this sequence junction (Fig. 2A) (15) but does not bind under the same conditions to any sequences present in the BK(WW) regulatory region (15). We expect that this intermediate containing a single AP-1 site would have substantially greater transcriptional activity than its parent and potentially could have been selected for in the viral population. In a second recombination event required to create BK(Dun), the AP-1 site would have been duplicated, increasing the level of transcriptional activity further. In the third event, nonessential sequences at the right end of the control region would have been deleted.

Our finding that the BK(WW) control region was transcriptionally quiescent conflicts with the original report describing an active enhancer in BK(WW) (24). In these earlier experiments, the BK(WW) upstream enhancer region was fused in reverse orientation to a truncated simian virus 40 promoter containing an AP-1 site adjacent to the point of fusion. It is likely that the activity of this construct is attributable to the juxtaposition of BK(WW) sequences with the simian virus 40 AP-1 site rather than to the activity of the BK(WW) enhancer per se. Our present results are in agreement with those of a recent study which concluded that a number of independently isolated BK(WW)-like strains were defective for growth, probably because they lack an active promoter-enhancer region (28).

A number of previous studies have reported sequence duplications leading to formation of active enhancers (11, 12, 29–31). For the most part, these experiments have been interpreted in terms of duplication of enhancer modules internal to the repeated sequences rather than formation of new enhancer modules at repeat junctions. In simian virus 40, subsequent experiments have indeed proven that enhancer modules are present internal to the repeated sequences in some cases (19, 25; reviewed in reference 8). In other cases, the situation is not so clear, because follow-up experiments have not been performed (11, 12, 29–31). In BKV, there are a number of variants that contain extensive sequence duplications but that show low levels of early gene expression or low growth rates. These include BK(RS) (R. Steinman, R.-B. Markowitz, A. Del Vecchio, U. Shabon, W. S. Dynan, and R. Ricciardi, submitted for publication), BK(pm526) (30), and multiple variants isolated from a systemic lupus erythematosus patient (28). These observations suggest that duplications alone are not sufficient to activate transcription in BKV. Our studies argue that, at least in some cases, specific junctions formed between duplicated sequences are critical for increased transcriptional activity.

We do not yet know to what extent the results with BK(Dun) can be generalized to other strains of BKV or to

other viruses. Preliminary experiments show that restriction fragments spanning repeat junctions in BK(MM) contain positively acting enhancer elements, although the precise sequences responsible have not yet been identified. The formation of a new promoter element spanning a repeat junction was also noted some years ago in a bacterial virus. The *c*17 mutant of bacteriophage λ contains a 9-bp duplication which allows constitutive expression of DNA replication functions (21).

We thank J. Schneringer for technical assistance and R. Sclafani for bringing to our attention the λc17 mutant.

This work was supported by Public Health Service grant CA 44958 from the National Cancer Institute to W.S.D.

LITERATURE CITED

1. **Cassill, J. A., K. L. Deyerle, and S. Subramani.** 1989. Unidirectional deletion and linker scan analysis of the late promoter of the human papovavirus BK. Virology **169**:172–181.
2. **Cassill, J. A., and S. Subramani.** 1988. The late promoter of the human papovavirus BK is contained within the early promoter enhancer region. Virology **166**:175–185.
3. **Chakraborty, T., and G. C. Das.** 1989. Identification of HeLa cell nuclear factors that bind to and activate the early promoter of human polyomavirus BK in vitro. Mol. Cell. Biol. **9**:3821–3828.
4. **Deyerle, K. L., J. A. Casill, and S. Subramani.** 1987. Analysis of the early regulatory region of the human papovavirus BK. Virology **158**:181–193.
5. **Deyerle, K. L., and S. Subramani.** 1988. Linker scan analysis of the early regulatory region of human papovavirus BK. J. Virol. **62**:3378–3387.
6. **Dhar, R., C. Lai, and G. Khoury.** 1978. Nucleotide sequence of the DNA replication origin for human papovavirus BKV: sequence and structural homology with SV40. Cell **13**:345–358.
7. **Dynan, W. S.** 1987. DNase I footprinting as an assay for mammalian gene regulatory proteins, p. 75–87. *In* J. K. Setlow (ed.), Genetic engineering. Plenum Publishing Corp., New York.
8. **Dynan, W. S.** 1989. Modularity in promoters and enhancers. Cell **58**:1–4.
9. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
10. **Grinnell, B. W., D. T. Berg, and J. D. Walls.** 1988. Negative regulation of the human polyomavirus BK enhancer involves cell-specific interaction with a nuclear repressor. Mol. Cell. Biol. **8**:3448–3457.
11. **Hara, K., Y. Oya, H. Kinoshita, F. Taguchi, and Y. Yogo.** 1986. Sequence reiteration required for the efficient growth of BK virus. J. Gen. Virol. **67**:2555–2559.
12. **Herr, W., and J. Clarke.** 1986. The SV40 enhancer is composed of multiple functional elements that can compensate for one another. Cell **45**:461–470.
13. **Howley, P. M.** 1980. Molecular biology of SV40 and the human polyomaviruses BK and JC, p. 489–550. *In* G. Klein (ed.), Viral oncology. Raven Press, New York.
14. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48**:79–89.
15. **Markowitz, R.-B., and W. S. Dynan.** 1988. Binding of cellular proteins to the regulatory region of BK virus DNA. J. Virol. **62**:3388–3398.
16. **Markowitz, R.-B., J. Schneringer, S. Tolbert, A. Küppers, and W. S. Dynan.** 1989. Functional organization of the BK virus transcriptional control region, p. 191–206. *In* L. P. Villarreal (ed.), Common mechanisms of transformation by small DNA tumor viruses. American Society for Microbiology, Washington, D.C.
17. **Nordeen, S. K., P. P. Green III, and D. M. Fowlkes.** 1987. A rapid, sensitive, and inexpensive assay for chloramphenicol acetyltransferase. DNA **6**:173–178.
18. **Nowock, J., U. Borgmeyer, A. W. Puschel, R. A. W. Rupp, and A. E. Sippel.** 1985. The TGGCA protein binds to the MMTV-LTR, the adenovirus origin of replication, and the BK virus enhancer. Nucleic Acids Res. **13**:2045–2061.
19. **Ondek, B., A. Shepart, and W. Herr.** 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. **6**:1017–1025.
20. **Padgett, B.** 1981. Human papovaviruses, p. 339–370. *In* J. Tooze (ed.), DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Rosenberg, M., D. Court, H. Shimatake, C. Brady, and D. L. Wulff.** 1978. The relationship between function and DNA sequence in an intercistronic regulatory region in phage λ. Nature (London) **272**:414–423.
22. **Rosenfeld, P. J., and T. J. Kelly.** 1986. Purification of nuclear factor 1 by DNA recognition site affinity chromatography. J. Biol. Chem. **261**:1398–1408.
23. **Rosenthal, N., M. Dress, P. Gruss, and G. Khoury.** 1983. BK viral enhancer element and a human cellular homolog. Science **222**:749–755.
24. **Rubinstein, R., N. Pare, and E. H. Harley.** 1987. Structure and function of the transcriptional control region of nonpassaged BK virus. J. Virol. **61**:1747–1750.
25. **Schirm, S., J. Jiricny, and W. Schaffner.** 1987. The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. Genes Dev. **1**:65–74.
26. **Seif, I., G. Khoury, and R. Dhar.** 1979. The genome of human papovavirus BKV. Cell **18**:963–977.
27. **Sompayrac, L. M., and K. J. Danna.** 1981. Efficient infection of monkey cells with DNA of simian virus 40. Proc. Natl. Acad. Sci. USA **78**:7575–7578.
28. **Sugimoto, C., K. Hara, F. Taguchi, and Y. Yogo.** 1989. Growth efficiency of naturally occurring BK virus variants in vivo and in vitro. J. Virol. **63**:3195–3199.
29. **Swimmer, C., and T. Shenk.** 1984. A viable simian virus 40 variant that carries a newly generated sequence reiteration in place of the normal duplicated enhancer element. Proc. Natl. Acad. Sci. USA **81**:6652–6656.
30. **Watanabe, S., and K. Yoshiike.** 1986. Evolutionary changes of transcriptional control region in a minute-plaque viable deletion mutant of BK virus. J. Virol. **59**:260–266.
31. **Weber, F., J. de Villiers, and W. Schaffner.** 1984. An SV40 "enhancer trap" incorporates exogenous enhancers or generates enhancers from its own sequences. Cell **36**:938–992.
32. **Yang, R. C. A., and R. Wu.** 1979. BK virus DNA: complete nucleotide sequence of a human tumor virus. Science **206**: 456–462.
33. **Yoshiike, K., and K. Takemoto.** 1986. Studies with BK virus and monkey lymphotrophic papovavirus, p. 295–326. *In* N. P. Salzman (ed.), The papovaviridae. Plenum Publishing Corp., New York.