Binding of Cellular Proteins to the Regulatory Region of BK Virus DNA

RHEA-BETH MARKOWITZ* AND WILLIAM S. DYNAN

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

The human papovavirus BK has a noncoding regulatory region located between the divergently transcribed early and late coding regions. Many strains of BK virus (BKV) have direct DNA sequence repeats in the regulatory region, although the number and extent of these repeats varies widely between independent isolates. Until recently, little was known about the individual functional elements within the BKV regulatory region, and the biological significance of the variable repeat structure has been unclear. To characterize the interaction between sequences in the BKV regulatory region and host cell transcription factors, we have carried out DNase I footprinting and competitive binding experiments on three strains of BKV, including one strain that does not contain direct sequence repeats. We have used relatively crude fractions from HeLa cell nuclear extracts, as well as DNA affinity-purified preparations of proteins. Our results demonstrate that BK(Dunlop), BK(WW), and BK(MM) each contain multiple binding sites for a factor, NF-BK, that is a member of the nuclear factor 1 family of transcription factors. We predict the presence of three to eight binding sites for NF-BK in the other strains of BKV for which a DNA sequence is available. This suggests that the binding of this protein is likely to be required for biological activity of the virus. In addition to NF-BK sites, BK(WW) and BK(MM) each contain a single binding site for transcription factor Sp1, and BK(Dunlop) contains two binding sites for transcription factor AP-1. The AP-1 sites in BK(Dunlop) span the junction of adjacent direct repeats, suggesting that repeat formation may be an important mechanism for de novo formation of binding sites not present in a parental strain.

BK virus (BKV) is a human papovavirus first isolated in 1971 from the urine of a renal allograft patient (19). A number of different strains of BKV have subsequently been isolated, in most cases from the urine of renal allograft recipients or patients suffering from Wiskott-Aldrich syndrome, an X-linked recessive disorder characterized by defects in cellular and humoral immunity (4). Comparative DNA sequence analysis shows that the BKV isolates have extensive homologies in their protein-coding region, both to each other and to simian virus 40 (SV40) (25, 27, 56, 71). However, these studies also indicate that the noncoding regulatory regions of BKV and SV40 are quite different in sequence.

In both BKV and SV40, this noncoding regulatory region is located between the early and late transcription units, near the origin of viral DNA replication. In SV40, this region is about 350 base pairs (bp) in size and contains most or all of the *cis*-acting elements required for initiation of viral RNA synthesis. There are several repeated sequence motifs, including the 21-bp repeats and the 72-bp repeats. The latter have been shown to possess enhancer function, that is, they increase the level of transcription from the early promoter regardless of orientation and position (3, 18, 34, 44).

In BKV, the noncoding regulatory region varies in size between isolates. Repeated sequences are present in this region of most viral strains. For example, the widely used BK(Dunlop) strain contains a triplication of a 68-bp block with an 18-bp deletion within the middle repeat (56). Using the chloramphenicol acetyltransferase assay, Rosenthal et al. (54) demonstrated that these BK(Dunlop) repeats possess transcriptional enhancer function. Studies using deletion mutants constructed in vitro demonstrate that one copy of the BK(Dunlop) 68-bp block is not sufficient for early

transcription, but that activity is restored when certain sequences within the 68-bp block are repeated (64). A recent comprehensive deletion analysis of BK(prototype), which is similar to BK(Dunlop), showed a progressive stepwise decline in early gene expression, as sequences within the 68-bp blocks were deleted (11).

The repeat structure varies widely between BKV strains and in some cases is quite complex. Recently, however, a new strain, BK(WW), was obtained by isolation and molecular cloning of viral DNA directly from patient urine (55). This differs from the procedures used to isolate most other BKV strains, which generally involved passage of the virus in cell culture prior to cloning. BK(WW) is distinguished from most other BKV strains by the absence of repeated DNA sequences. The workers who isolated this strain have pointed out that the regulatory regions of many other BKV strains can be derived from BK(WW) by postulating a small number of duplication and deletion events. Such events might have occurred either in nature or during passage of the virus in culture. Specific examples of how this might have occurred with BK(prototype), BK(Dunlop), and BK(MM) will be presented in the Results section.

The transcriptional enhancer activity associated with the BKV regulatory region is almost certainly dependent on the binding of specific host cell transcription factors, as has been shown with other papovaviruses. A prior report described two recognition sites in the regulatory region of BK(Dunlop) for the TGGCA-binding protein, now thought to be functionally the same as nuclear factor 1 (NF1) (39, 46). Little is known about other proteins that may bind to this region, however, and comparative studies of different strains have not been carried out. There has been considerable speculation in the literature about the DNA sequences similar to the adenovirus E1a and SV40 enhancer core motifs which are present in BKV (23, 67), but it is not known if these actually

---

* Corresponding author.

correspond to protein-binding sites or constitute functional units of the enhancer.

We have screened protein fractions derived from HeLa cell nuclear extracts for the presence of proteins that recognize specific DNA sequences in BKV. This approach has been used in SV40 to identify many different proteins that bind the enhancer region (68). Initially, we carried out DNase I footprinting experiments with relatively crude heparin-agarose step-gradient fractions. Subsequently, the identity of specific proteins was confirmed by footprinting with DNA affinity-purified protein preparations and by competitive binding experiments.

A surprisingly simple picture emerged from these studies. The enhancer region of each of the three strains BK(Dunlop), BK(WW), and BK(MM) contained three to six binding sites for a protein we refer to as NF-BK, which is a member of the NF1 family of transcription factors. This was the only protein we detected that recognized sites in all three strains, suggesting that it is essential for biological activity of the enhancer region. An additional recognition site for the transcription factor Sp1 was detected in two strains, and two sites for the transcription factor AP-1 were detected in a third strain.

MATERIALS AND METHODS

Preparation of radiolabeled DNA probes. (i) BK(Dunlop) strain. The 216-bp *Hae*III DNA fragment containing the BK(Dunlop) regulatory region was inserted via *Bam*HI linkers into the *Bam*HI site of pUC19 to create the plasmid pDun/Bam. To prepare the probe for DNase I footprinting, the *Eco*RI-*Hin*dIII fragment of pDun/Bam containing the BK(Dunlop) sequences was excised, treated with calf intestinal phosphatase (Boehringer Mannheim Biochemicals), and purified by preparative gel electrophoresis. This fragment was then radiolabeled with T4 polynucleotide kinase and [γ-^{32}P]ATP (crude grade, 7,000 Ci/mmol; ICN Pharmaceuticals Inc.). To prepare a late-strand probe that was singly end labeled at the *Hin*dIII site, the DNA was digested with *Kpn*I to remove a small fragment of labeled DNA from the opposite end. Similarly, to prepare an early-strand probe labeled at the *Eco*RI site, the DNA was digested with *Pst*I. After the final restriction endonuclease digestion, the DNA was extracted with PCIA (phenol-chloroform-isoamyl alcohol, 25:24:1 [vol/vol]), extracted with CIA (chloroform-isoamyl alcohol, 24:1 [vol/vol]), and precipitated from ethanol.

(ii) BK(WW) strain. The plasmid pHS Bam (55) contained a 426-bp *Hha*I-*Sst*I fragment encompassing the BK(WW) regulatory region. This fragment, which as the result of the cloning procedure was flanked by *Bam*HI sites, was excised with *Bam*HI and inserted into pUC19 at the *Bam*HI site to create pWW/Bam-9 and pWW/Bam-1. These plasmids contain the WW regulatory sequences in the orientations analogous to pDun/Bam and pMM/Bam, respectively. Early-strand probes were generated from the *Eco*RI-end-labeled, *Pst*I-digested fragment from pWW/Bam-9 or the *Hin*dIII-end-labeled, *Kpn*I-digested fragment from pWW/Bam-1.

(iii) BK(MM) strain. The 291-bp *Hae*III fragment containing the BK(MM) regulatory region was inserted via *Bam*HI linkers into the *Bam*HI site of pUC19 to create the plasmid pMM/Bam. Radiolabeled DNA probes were prepared as described above. The BK(MM) fragment was inserted in an orientation opposite to that of the BK(Dunlop) fragment. Therefore, a late-strand probe that was singly end labeled at the *Eco*RI site was prepared by *Pst*I digestion, and an

early-strand probe that was singly end labeled at the *Hin*dIII site was prepared by *Kpn*I digestion.

DNase I footprinting. DNase I footprinting was performed essentially as described (16, 48), except that all binding reactions contained 40 μg of poly(dI-dC) alternating copolymer (Pharmacia, Inc.) per ml and 40 μg of pentadeoxynucleotide mixture [pd(N)5] oligonucleotide mixture (Pharmacia) per ml. Binding reactions were carried out on ice. Samples were then warmed to 20°C and treated with DNase I. Reactions were terminated as described, extracted with PCIA and CIA, precipitated with ethanol, and electrophoresed on 8% denaturing polyacrylamide gels (41). Sequence markers were prepared by the method of Maxam and Gilbert (41). In competitive binding experiments, the competitor DNA was mixed with the radiolabeled probe DNA, poly(dI-dC) copolymer, and pd(N)5 oligonucleotide prior to the addition of protein.

Preparation of DNA-binding proteins. Nuclei were prepared as described previously (16) using HeLa cells from 12 to 24 liters of suspension culture (approximately 5×10^9 to 1 $\times 10^{10}$ cells total). The nuclei were suspended in 2 packed-cell volumes of extraction buffer, and after centrifugation, the resulting supernatant was applied directly to a column of heparin-agarose prepared as described previously (10) (1-ml bed volume per liter of original culture). The column was eluted with a 0.15 to 0.4 M KCl step gradient, and protein-containing fractions were pooled. Protein concentrations were determined by the method of Bradford (5).

DNA affinity columns were prepared as described by Kadonaga and Tjian (33). Two complementary oligonucleotides containing the desired consensus sequences were synthesized by the β-cyanoethyl phosphoramidite method with an Applied Biosystems model 380B DNA synthesizer and were purified by preparative electrophoresis on a 20% polyacrylamide gel containing 8 M urea. The two oligonucleotides were dissolved in a solution of 50 mM Tris chloride [pH 7.6], 10 mM MgCl$_2$, 5 mM dithiothreitol, 0.1 mM spermidine, and 0.1 mM EDTA and were annealed and radiolabeled as described previously (33). The ligation reaction was carried out by incubating 440 μg of annealed oligonucleotide in a 1-ml reaction containing 66 mM Tris chloride (pH 7.5), 5 mM MgCl$_2$, 5 mM dithiothreitol, and 4 mM ATP for 24 h at 25°C. The extent of ligation was analyzed by polyacrylamide gel electrophoresis. Following purification by extraction with PCIA and CIA, and precipitation from ethanol, the oligonucleotide ligation products were dissolved in water and coupled to cyanogen bromide-activated Sepharose CL2B300 (Sigma Chemical Co.) as described. For NF-BK purification, the complementary oligonucleotides GATCTGGAATGCA GCCAA and GATCTTGGCTGCATTCCA were used. For AP-1 purification, the complementary oligonucleotides GA TCATGGTTGCTGACTAATTGAGA and GATCTCTCAA TTAGTCAGCAACCAT were used. For Sp1 purification, the complementary oligonucleotides GATCGGGGCGGG GC and GATCGCCCCGCCCC were used.

Heparin-agarose 0.15 to 0.4 M KCl-step gradient fractions were mixed with 200 μg of poly(dI-dC), diluted to a conductivity equivalent to column buffer containing 0.15 M KCl, and subjected to chromatography on the DNA affinity columns. The DNA affinity column (1-ml bed volume per 8 liters of original culture) was equilibrated with buffer Z (25 mM HEPES(K$^+$) (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid) [pH 7.8], 12.5 mM MgCl$_2$, 1 mM dithiothreitol, 20% [vol/vol] glycerol, 0.1% [vol/vol] Nonidet P-40) containing 0.15 M KCl. The extract was cycled over the column four times by gravity at a flow rate of approximately
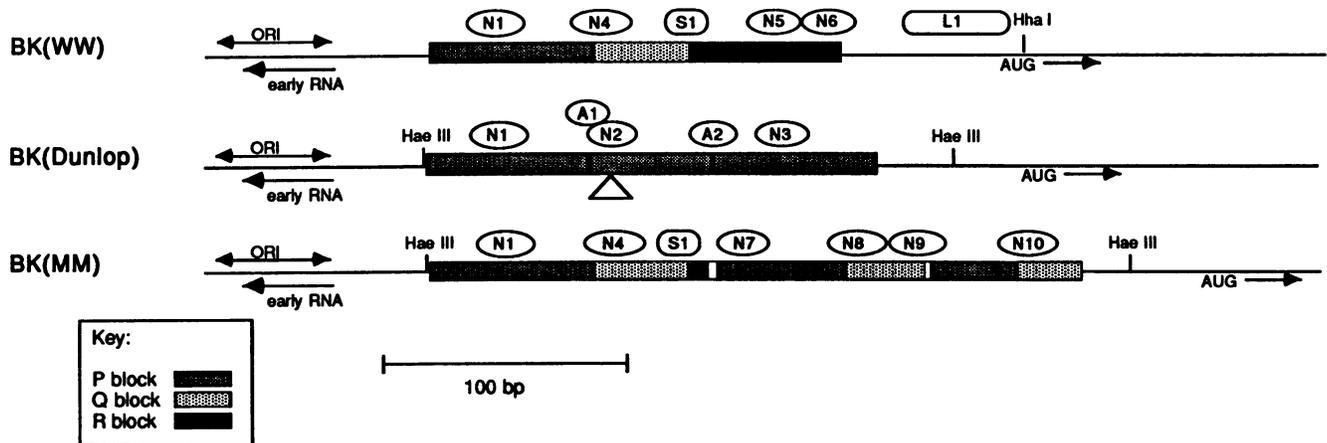
FIG. 1. Sequence relationship of BK(WW), BK(Dunlop), and BK(MM). Stippled boxes represent the regulatory region of BK(WW), which has been divided arbitrarily into three sequence blocks, P, Q, and R (see Key). Single lines represent early-region DNA to the left and late-region DNA to the right of the regulatory region. Binding sites in BK(WW) for NF-BK (N1, N4, N5, and N6), Sp1 (S1), and an unknown factor (L1) are indicated. The AUG codon for the agnoprotein, the presumptive initiation site for early RNA, and the replication origin region are marked. Restriction sites used in subcloning procedures are indicated (see Materials and Methods). BK(Dunlop) could have arisen from BK(WW) by a deletion of both the Q and R blocks and 4 bp of flanking DNA, triplication of the P block, and subsequent deletion of 18 bp within the second of the three repeated P blocks. Binding sites in BK(Dunlop) for NF-BK (N1, N2, and N3) and AP-1 (A1 and A2) and the 18-bp deletion within the middle P block in BK(Dunlop) (Δ) are indicated. BK(MM) could have arisen from BK(WW) by a series of duplication and deletion events as described in the text. Binding sites in BK(MM) for NF-BK (N1, N4, N7, N8, N9, and N10) and Sp1 (S1) are indicated. The small open boxes within the BK(MM) sequence indicate the insertions of unique DNA.

5 column volumes per h. Fractions were eluted with 5 ml of buffer Z containing 0.15 M KCl at a flow rate of 2 column volumes per h, followed by a 30-ml gradient of 0.15 to 1.75 M KCl at a flow rate of 3 column volumes per h. Fractions were collected and assayed for binding activity by DNase I footprinting. Typically, heparin-agarose fractions were first passed over the AP-1 DNA affinity column. The AP-1-depleted flowthrough from this column was passed over the Sp1 affinity column, and the AP-1- and Sp1-depleted flow-through from this column was passed over the NF-BK column. Specific DNA binding activity was detected in fractions eluting from the respective affinity columns at approximately 0.25 to 0.4 M KCl. NF-BK and AP-1 preparations were purified approximately 100-fold, and Sp1 preparations were purified approximately 30-fold.

RESULTS

Sequence relationship of BK(WW), BK(Dunlop), BK(prototype), and BK(MM). As discussed previously, a comparison of the BK(WW) regulatory region with those of BK(prototype), BK(Dunlop), BK(MM), and other BKV isolates suggests that the latter strains could have been derived from BK(WW) by a small number of sequence duplications and deletions (55, 56, 71). To illustrate how this might have occurred and to facilitate subsequent discussion of the binding sites for cellular factors, we have arbitrarily divided the WW sequence into three blocks, labeled P, Q, and R in Fig. 1. These blocks contain 68, 39, and 63 bp of DNA, respectively. BK(Dunlop) could have arisen from BK(WW) by deletion of the Q and R blocks and 4 bp of flanking DNA, triplication of the P block, and subsequent deletion of 18 bp within the second of the three repeated P blocks. BK(prototype) (not shown) could have arisen in a similar way but the deletion contains only the R block; the Q block remains. BK(WW) reportedly will not grow in cultured cells (55); BK(prototype) grows but is unstable (56); BK(Dunlop) grows relatively well. This suggests that duplication of the P

block favors viral growth in culture and that the Q and R block sequences are nonessential.

The generation of BK(MM) from BK(WW) is more complicated but can be readily visualized. The first step could have involved the deletion of the rightmost 55 bp of the R block, together with 12 bp of the adjacent unique sequence. A unit encompassing 53 bp of the P block and 34 bp of the Q block may then have been duplicated, with 4 bp of unique sequence inserted between the repeats. Another duplication encompassing 36 bp of the P block and 25 bp of the Q block may then have occurred, with 2 bp inserted between the duplicated segments, to give the final structure as shown.

Binding of cellular proteins to the regulatory region of BK(Dunlop) DNA. To develop a better understanding of the function of the different sequence blocks, we constructed a map of the binding sites for various host cell proteins that bind to this region using the technique of DNase I footprinting. Because BK(Dunlop) is the best-characterized of the BKV laboratory strains, we began our studies using this DNA. DNase I footprinting probes were prepared for both strands of DNA as described in Materials and Methods. Heparin-agarose fractions of HeLa cell nuclear extracts were used as a source of DNA-binding proteins. Previous studies have shown that fractions prepared as described (see Materials and Methods) are enriched for a broad spectrum of DNA-binding proteins and transcription factors (16, 48).

We identified five sites within the BK(Dunlop) regulatory DNA that were protected by heparin-agarose fractions of HeLa cell nuclear extract (Fig. 2A, lanes 2 through 4; Fig. 2B, lanes 3 through 5). Three of these sites, N1, N2, and N3, are located within the triplicated P blocks. N1 and N3 are identical in sequence and show complete protection. N2 differs slightly in sequence from N1 and N3 because of the previously mentioned deletion internal to the middle P block and shows only partial protection. N1 and N3 coincide with sites identified by Nowock et al. (46) as binding regions for a partially purified TGGCA-binding protein. It has been sug-
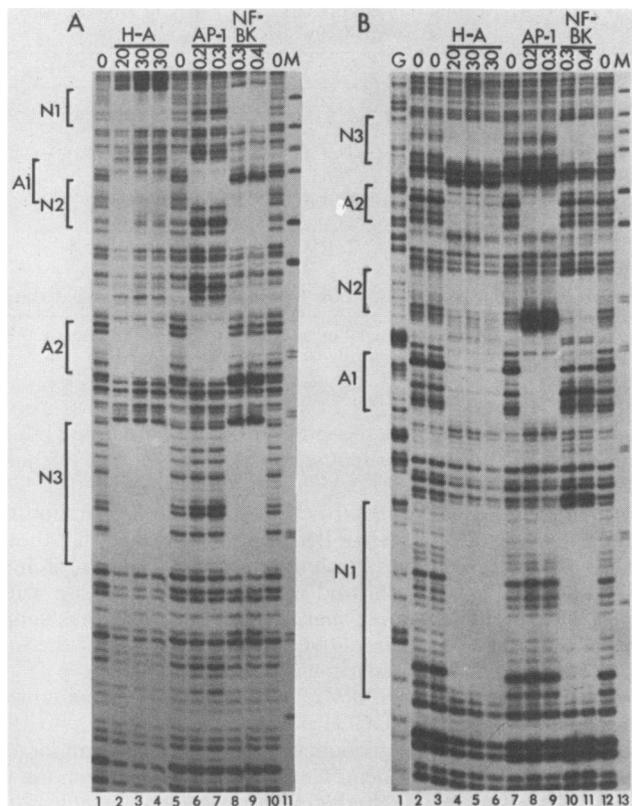
FIG. 2. DNase I footprint of the regulatory region of BK(Dunlop). DNA probes were singly end labeled and incubated with heparin-agarose fractions of HeLa cell nuclear extracts (H-A) or with DNA affinity-purified protein preparations (AP-1 and NF-BK) as indicated. DNase I cleavage was performed and products were analyzed on sequencing gels as described in Materials and Methods. Protected regions are indicated by brackets and numbered as described in the text. (A) The DNA probe was 5' end labeled on the early strand. Lanes 1, 5, and 10, No protein extract added; lanes 2, 3, and 4, heparin-agarose fractions, 20, 30, and 30 μg of protein, respectively; lanes 6 and 7, DNA affinity-purified AP-1 fractions, 0.2 and 0.3 μg of protein, respectively; lanes 8 and 9, DNA affinity-purified NF-BK fractions, 0.3 and 0.4 μg of protein, respectively; lane 11, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 217 bp). (B) The DNA probe was 5' end labeled on the late strand. Lane 1, Products of the G-specific sequencing reaction using the technique of Maxam and Gilbert (41); lanes 2, 3, 7, and 12, no protein extract added; lanes 4, 5, and 6, heparin-agarose fractions, 20, 30, and 30 μg of protein, respectively; lanes 8 and 9, DNA affinity-purified AP-1 fractions, 0.2 and 0.3 μg of protein, respectively; lanes 10 and 11, DNA affinity-purified NF-BK fractions, 0.3 and 0.4 μg of protein, respectively; lane 13, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 217 bp).

gested that this protein may be the same as NF1, a protein isolated from HeLa cell nuclei which is required for adenovirus replication (39, 45). All three sites, N1, N2, and N3, contain sequences similar or identical to the NF1 consensus, $TGG(N_{6-7})$ T/G CCAA (13), and because of this similarity, it is likely that they are protected by the same protein.

The two other protected sites, A1 and A2, span the junctions between the tandemly repeated P blocks in the BK(Dunlop) regulatory region. These junctions contain the sequence TGACTCA, corresponding to the consensus recognition sequence for AP-1, a transcription factor that binds to the control regions of SV40 DNA and the human metal-

lothionein IIA gene (37, 38). This suggests that AP-1 may be responsible for the protection seen at sites A1 and A2.

To better determine the identity of the proteins responsible for the protection of the five sites in BK(Dunlop) DNA, we isolated these proteins using sequence-specific DNA affinity column chromatography. Columns were prepared essentially as described by Kadonaga and Tjian (33). The oligonucleotide used to isolate the protein that binds sites N1, N2, and N3 contained the sequence TGGAATGCAGC CAA, which occurs in sites N1 and N3 in the BKV P block. As described above, this contains a partial match to the NF1 consensus recognition sequence (underlined). Recent reports from several laboratories suggest that NF1, TGGCA-binding protein, and other proteins are part of a family of related factors (13, 29, 39, 42, 46). In view of the present confusion of nomenclature, we will provisionally refer to the protein isolated from our column as NF-BK, until it is demonstrated whether or not it is identical to other specific members of this family. The oligonucleotide used to isolate AP-1 contained the sequence ATGGTTGCTGACTAATT GAGA from the SV40 AP-1-binding region. The AP-1 recognition sequence (underlined) in this oligonucleotide matches sequences in BK(Dunlop) at six of seven positions, but the flanking sequences are different from the flanking sequences in BK(Dunlop). Heparin-agarose fractions of a HeLa cell nuclear extract were passed over the NF-BK and AP-1 affinity columns, and purified factors were eluted with KCl gradients as described in Materials and Methods. The DNA-affinity purified preparations are purified 100-fold or more relative to the heparin-agarose fractions, as judged by the amount of protein required to give a footprint (see figure legends).

These purified protein preparations were used in DNase I footprinting experiments, using both strands of BK(Dunlop) as probes. Results are shown in Fig. 2. The affinity-purified AP-1 fractions showed complete protection of sites A1 and A2 of BK(Dunlop). In other experiments (results not shown), these same fractions gave full protection of the known AP-1 sites in SV40. These results confirm that AP-1 recognizes the A1 and A2 sites spanning the junction between adjacent P blocks of BK(Dunlop). The affinity-purified NF-BK fractions showed complete protection of sites N1, N2, and N3 of BK(Dunlop), confirming that all three sites are recognized by the same or closely related proteins. It is evident from the results with purified proteins that sites A1 and N2 overlap. Steric hindrance between proteins bound at these sites may account for the relatively weak protection at site N2 in some experiments using heparin-agarose fractions (for example, Fig. 2A, lanes 2 through 4).

Binding of cellular proteins to the regulatory region of BK(WW) DNA. We next examined the binding of cellular protein fractions to BK(WW) DNA. DNase I footprinting probes were prepared for the early strand of cloned viral DNA as described in Materials and Methods. Heparin-agarose fractions of HeLa cell nuclear extracts and DNA affinity column preparations of AP-1 and NF-BK were prepared as previously described. In addition, fractions containing transcription factor Sp1 were prepared by DNA affinity chromatography, by using an oligonucleotide containing the Sp1 consensus recognition sequence GGGG CGGGGC.

The results of these experiments are shown in Fig. 3. A number of protected regions were seen using the heparin-agarose fractions (lanes 5 and 6). At the top of the gel, site N1 corresponds to the same P block footprint seen with BK(Dunlop). Sites A1, N2, A2, and N3 are not present in
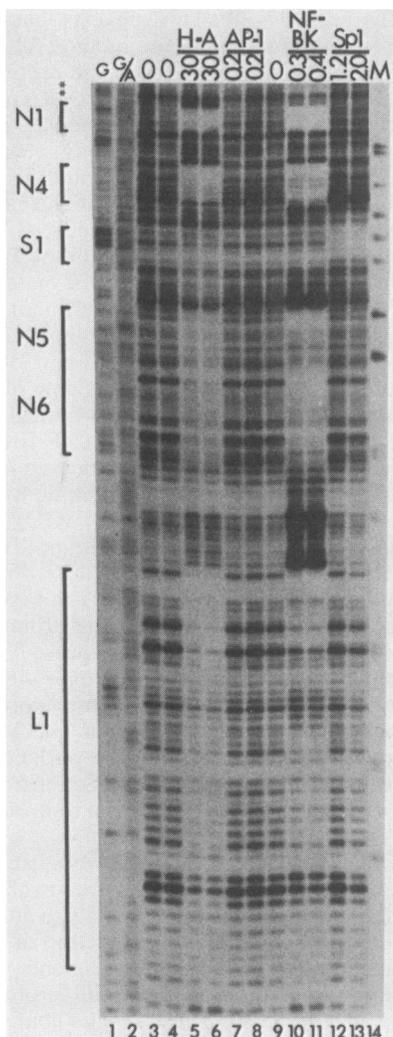
FIG. 3. DNase I footprint of the regulatory region of BK(WW). The DNA probe was 5' end labeled on the early strand. For experimental details, see the legend to Fig. 2 and Materials and Methods. Protected regions are indicated by brackets and numbered as described in the text. Asterisks indicate a possible Sp1 recognition sequence where no binding was detected. Lane 1, Products of the Maxam-Gilbert G-specific sequencing reaction; lane 2, products of the Maxam-Gilbert G- and A-specific sequencing reaction; lanes 3, 4, and 9, no protein extract added; lanes 5 and 6, heparin-agarose fractions, 30 μg of protein; lanes 7 and 8, DNA affinity-purified AP-1 fractions, 0.2 μg of protein; lanes 10 and 11, DNA affinity-purified NF-BK fractions, 0.3 and 0.4 μg of protein, respectively; lanes 12 and 13, DNA affinity-purified Sp1 fractions, 1.2 and 2.0 μg of protein, respectively; lane 14, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 242 bp).

BK(WW); these sites arose in BK(Dunlop) as the result of the P block triplication. The next three protected sites that are seen in BK(WW) have therefore been labeled N4, N5, and N6. N4 spans the junction between the P and Q blocks, N5 is within the R block, and N6 spans the righthand junction of the R block. All three sites contain significant matches to the P block NF-BK recognition site (Table 1), suggesting that all are recognized by the same protein. The final protected site in BK(WW) DNA, labeled L1, is located to the right of the R block, immediately upstream from the start codon for the BKV agnoprotein. This site has some

TABLE 1. Recognition sequences for NF-BK in BK(Dunlop), BK(WW), and BK(MM)[a]

Site	Sequence
N1	TGGAATGCAGCCAAA
N2	GGGAATGCAGCCAAA
N4	TGGGCAGC C/A GCCAGT
N5	TGGAAACTGGCCAAA
N6	TGGCTGCTTTCCACT
N9	TGAAACCATGCCAAA
L1	TGGCCTTGTCCCCAG
Consensus	TGGAA T/A G/C C/T A/T GCCAAA

[a] The following NF-BK recognition sites have been omitted because their sequence is identical to sites already listed: N3, N7, N8 and N10.

similarity to the NF-BK recognition site in the P block but is more divergent than the sequences in sites N2, N4, N5, and N6 (Table 1).

Affinity-purified NF-BK fractions gave full protection of sites N1, N4, N5, and N6 of BK(WW), confirming that these sites are recognized by the same or closely related proteins. By contrast, site L1 showed only limited protection with purified NF-BK fractions, suggesting that NF-BK has some affinity for these sequences but may not account for the full protection seen with heparin-agarose fractions.

We noticed that the BK(WW) regulatory region contained a sequence GGGGCGGGGT that is a good match to the consensus recognition sequence for the transcription factor Sp1. This sequence is located at the junction between the Q and R blocks. Although this region was not significantly protected by heparin-agarose fractions, we did observe complete protection with affinity-purified Sp1, and have labeled the site S1 (Fig. 3, lanes 12 and 13). Because there is both a match to the consensus and observed protection with purified protein, we believe this is a bona fide Sp1 recognition site. The relative lack of protection with heparin-agarose fractions may reflect a low concentration of Sp1 in this particular preparation. This was the only site observed for Sp1; we did not detect protection of a putative Sp1 site (AGGGAGGGAGC) in the P block (Fig. 3, asterisks) using either heparin-agarose or DNA affinity-purified fractions.

We also tested BK(WW) DNA with affinity-purified AP-1. As expected, no protection of BK(WW) DNA was observed with AP-1 fractions (Fig. 3, lanes 7 and 8). The AP-1 sites that were seen with BK(Dunlop) DNA spanned the junction between the triplicated P blocks, and this triplication is not present in BK(WW).

Binding of cellular proteins to the regulatory region of BK(MM) DNA. DNase I footprinting experiments with BK(MM) DNA are shown in Fig. 4. As in the previous experiments, we used heparin-agarose fractions of HeLa cell nuclear extracts, as well as NF-BK, AP-1, and Sp1 fractions purified by DNA affinity column chromatography.

Six regions were protected by heparin-agarose fractions (Fig. 4A, lane 5; Fig. 4B, lanes 5 and 6). Site N1 (Fig. 4B only; not shown in Fig. 4A) is the same P block footprint seen in the other strains. Site N4 is at the junction of the P and Q blocks and corresponds to the N4 site in BK(WW) DNA. Sites N7, N8, and N10 arise as the results of duplications of portions of the P and Q blocks. N9, however, is unique to BK(MM) and arises as the result of a novel Q-P junction containing a 2-bp insert. This unusual creation of a protein-binding site across a repeat junction is analogous to the situation in BK(Dunlop), where novel AP-1 sites were created at the P-P junctions.
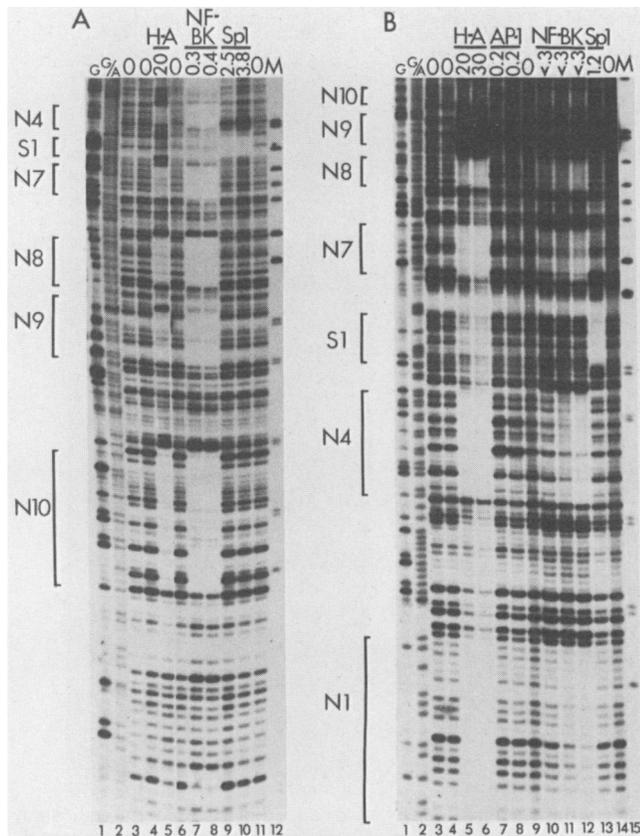
FIG. 4. DNase 1 footprint of the regulatory region of BK(MM). For experimental details, see the legend to Fig. 2 and Materials and Methods. (A) The DNA probe was 5' end labeled on the early strand. Lane 1, Products of the Maxam-Gilbert G-specific sequencing reaction; lane 2, products of the Maxam-Gilbert G- and A-specific sequencing reaction; lanes 3, 4, 6, and 11, no protein extract added; lane 5, heparin-agarose fractions, 20 μg of protein; lanes 7 and 8, DNA affinity-purified NF-BK fractions, 0.3 and 0.4 μg of protein, respectively; lanes 9 and 10, DNA affinity-purified Spl fractions, 2.5 and 3.8 μg of protein, respectively; lane 12, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 242 bp). (B) The DNA probe was 5' end labeled on the late strand. Lane 1, Products of the Maxam-Gilbert G-specific sequencing reaction; lane 2, products of the Maxam-Gilbert G- and A-specific sequencing reaction; lanes 3, 4, 9, and 14, no protein extract added; lanes 5 and 6, heparin-agarose fractions, 20 and 30 μg of protein, respectively; lanes 7 and 8, DNA affinity-purified AP-1 fractions, 0.2 μg of protein; lanes 10, 11, and 12, DNA affinity-purified NF-BK fractions, 3, 5, and 10 μl, respectively; the highest amount tested (10 μl) contained less than 0.3 μg of protein; lane 13, DNA affinity-purified Spl fractions, 1.2 μg of protein; lane 15, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 309 bp).

All six of the sites protected by heparin-agarose fractions in BK(MM) DNA contain close matches to the NF-BK consensus recognition sequence (Table 1). We tested the BK(MM) probes in DNase I footprinting experiments using affinity-purified NF-BK and found, as expected, full protection of sites N1, N4, N7, N8, N9 and N10, with essentially the same boundaries as seen with the heparin-agarose fractions.

BK(MM) contains an Spl consensus recognition sequence coinciding with the single Q-R junction present in this strain. This sequence is protected by purified Sp-1 (Fig. 4A, lanes 9

and 10; Fig. 4B, lane 13). This site is analogous to the single Spl site in BK(WW) and therefore has been labeled S1.

Competitive binding experiments. Experiments with fractions purified by DNA sequence affinity chromatography, together with sequence comparisons, strongly suggests that the observed footprints at sites N1 through N10 are attributable to a single factor, NF-BK. To confirm this, we carried out competitive binding experiments using double-stranded oligonucleotides containing either the NF-BK recognition sequence from site N1 or a heterologous sequence from the promoter region of an unrelated virus. In these experiments, competitor oligonucleotide was mixed with probe DNA, heparin-agarose fractions were added, complexes were allowed to form, and DNase I footprinting was carried out as before.

Figure 5A shows a competitive binding experiment using BK(Dunlop) probe. Significant inhibition of NF-BK binding at sites N1 and N3 is evident with 20 nM specific competitor, whereas no inhibition is seen with the heterologous oligonucleotide at concentrations as high as 310 nM. Inhibition of binding to site N2 is more difficult to see, because protection is only partial even in the absence of competitor. However, it appears that inhibition occurs with 6 to 20 nM of the specific competitor, and that there is no inhibition with the heterologous competitor. These results suggest that, as expected, sites N1, N2, and N3 of BK(Dunlop) are protected by the same or closely related protein species.

In contrast, the binding of AP-1 at sites A1 and A2 of BK(Dunlop) DNA is not inhibited by any of the concentrations of the NF-BK oligonucleotide that we tested. Some inhibition of AP-1 binding was seen at the highest concentration tested of the heterologous oligonucleotide, probably because this oligonucleotide contains a sequence (TGACGTG) that partially matches the AP-1 consensus recognition sequence. This result confirms that sites A1 and A2 are protected by a protein that is different from the one that binds to sites N1, N2, and N3.

Competitive binding studies were extended to BK(WW) DNA (Fig. 5B). In this case, binding to four protected regions, N1, N4, N5, and N6, was inhibited by the NF-BK oligonucleotide. There was no inhibition with the heterologous oligonucleotide. As with BK(Dunlop), these results suggest that the same protein is binding to all sites. There was some inhibition of binding to site L1 with the NF-BK oligonucleotide, but the inhibition was not complete, even at 120 nM. This is consistent with the results of the binding experiments using purified factors and suggests that protection in the L1 region arises predominantly from the binding of a different factor.

Competitive binding studies with BK(MM) DNA are shown in Fig. 5C. Here, there are six protected regions, N1, N4, N7, N8, N9, and N10. Binding to sites N1, N4, N7, and N8 is inhibited by 20 nM NF-BK oligonucleotide, whereas there is no inhibition by 310 nM heterologous oligonucleotide. Again, this suggests that these sites are protected by a single factor, NF-BK. Some inhibition of binding is evident at sites N9 and N10. These are more difficult to see because protection of these sites with heparin-agarose fractions in the absence of competitor is not complete (Fig. 5, lanes 3 and 4; Fig. 4A, lane 5). However, some inhibition occurs at 20 nM specific competitor at site N9 (note the disappearance of the hypersensitive band just above the protected region) and site N10. No inhibition of protection is seen with the heterologous oligonucleotide at 310 nM. Similar results were obtained in several independent experiments, suggesting that
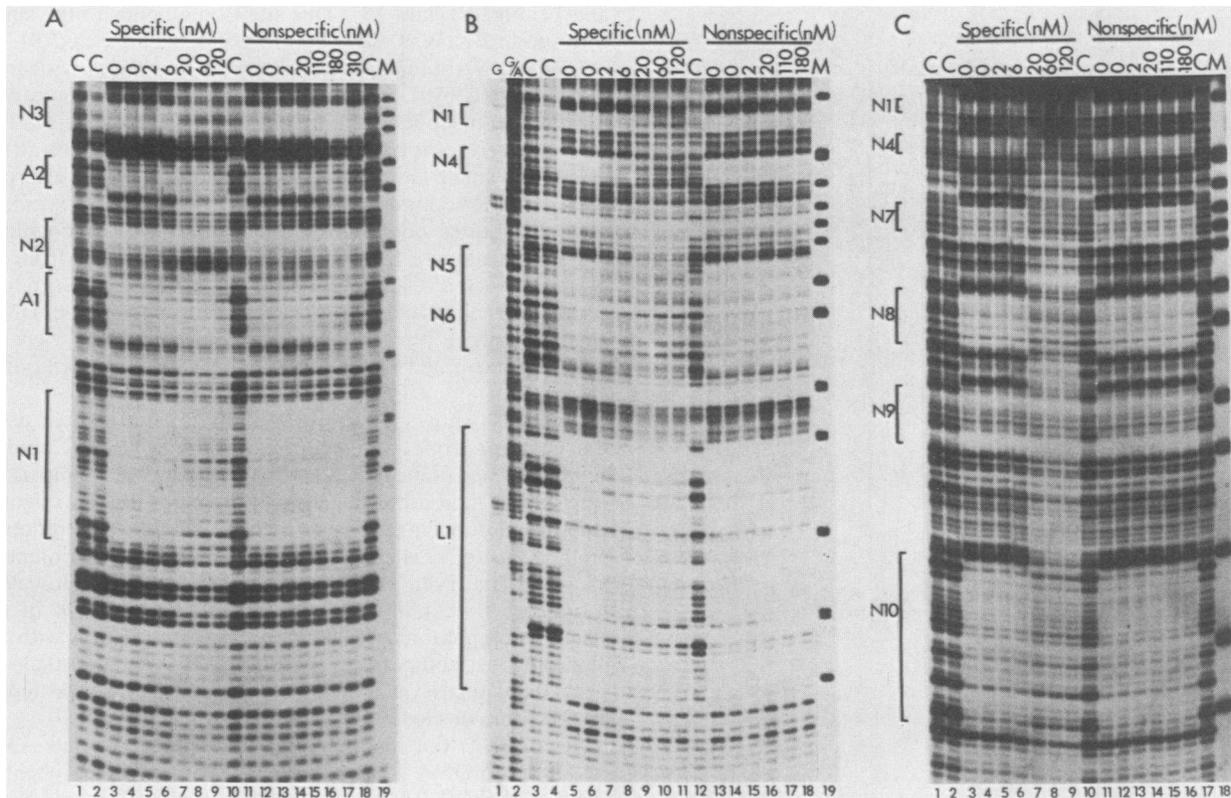
FIG. 5. Competitive binding to NF-BK sites in the regulatory region of BK virus. Competitor DNA was added in the amounts indicated to a mixture containing singly end-labeled DNA probes, poly(dI-dC) copolymer, and pd(N)5 oligonucleotides as described in the text and Materials and Methods. The double-stranded oligonucleotide containing the NF-BK recognition sequence from site N1 (5' TGGAAT GCAGCCAA 3') was used as the specific competitor. A double-stranded oligonucleotide containing a sequence present in the human T-cell lymphotropic virus type 1 promoter (5' TGGGCTAGGCCCTGACGTGTCCCCCTGAAGACAAA 3') was used as the nonspecific competitor. Heparin-agarose fractions (30 µg) of HeLa cell nuclear extracts was added to the DNA mixture, DNase I cleavage was performed, and products were analyzed on sequencing gels as described in Materials and Methods. Protected regions are indicated by brackets and enumerated as described in the text. (A) The DNA probe was 5' end labeled on the late strand of BK(Dunlop). Lanes 1, 2, 10, and 18, No protein extract added (control [C]); lanes 3 through 9, NF-BK-specific competitor DNA was added in the following concentrations: 0, 0, 2, 6, 20, 60, and 120 nM, respectively; lanes 11 through 17, nonspecific competitor DNA was added in the following concentrations: 0, 0, 2, 20, 110, 180, and 310 nM, respectively; lane 19, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 201 bp). (B) The DNA probe was 5' end labeled on the early strand of BK(WW). Lane 1, Products of the Maxam-Gilbert G-specific sequencing reaction; lane 2, products of the Maxam-Gilbert G- and A-specific sequencing reaction; lanes 3, 4, and 12, no protein extract added (control [C]); lanes 5 through 11, NF-BK-specific competitor DNA was added in the following concentrations: 0, 0, 2, 6, 20, 60, and 120 nM, respectively; lanes 13 through 18, nonspecific competitor DNA was added in the following concentrations: 0, 0, 2, 20, 110, and 180 nM, respectively; lane 19, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 309 bp). (C) The DNA probe was 5' end labeled on the early strand of BK(MM). Lanes 1, 2, 10, and 17, No protein extract added (control [C]); lanes 3 through 9, NF-BK-specific competitor DNA was added in the following concentrations: 0, 0, 2, 6, 20, 60, and 120 nM, respectively; lanes 11 through 16, nonspecific competitor DNA was added in the following concentrations: 0, 0, 2, 20, 110, and 180 nM, respectively; lane 18, marker (M), HpaII digestion products of pBR322 DNA (largest fragment shown is 309 bp).

protection at N9 and N10 is due, in part, to the binding of NF-BK, but that other proteins may also contribute.

DISCUSSION

One of the ongoing questions with BKV is to what degree the enhancer region is similar to that of the related and better-characterized virus SV40. The present results show the differences between the two systems. The distinguishing feature of the enhancer region in BKV is the presence of multiple binding sites for NF-BK; these were the only protein recognition sites that we detected in all three BKV strains tested. The ubiquity of NF-BK binding sites in BKV stands in contrast to SV40, where NF1-related protein-binding sites have not been reported in the enhancer region.

Although different strains of BK virus have undergone complex rearrangements, multiple binding sites for NF-BK

are always present. Where sites have been lost through deletions, new sites have been formed by duplications and in one case, by a duplication coupled with a small insertion. BK(WW), which does not contain repeated sequences in the regulatory region, has four NF-BK sites: N1 within the P block, N4 at the P-Q junction, N5 within the R block, and N6 at the R block-late region junction. BK(Dunlop) has a deletion that results in the loss of sites N4, N5, and N6. However, the P block triplication has resulted in the generation of two additional NF-BK binding sites, N2 and N3. BK(MM) has a deletion, relative to BK(WW), that results in the loss of sites N5 and N6, but duplications have resulted in the formation of four additional sites, N7, N8, N9, and N10.

We can predict the location of NF-BK sites in a number of other BKV strains on the basis of our results with BK(WW), BK(Dunlop), and BK(MM). Thus, we expect that BK(pro-

totype) contains the same NF-BK sites as BK(Dunlop) plus an N4-like site at the P-Q junction. BK(IR), isolated from a pancreatic adenoma of beta islet cells (7, 49), contains a total of eight NF-BK binding sites: three N1-like binding sites within the P block, three N4-like sites at P-Q junctions, one N5-like site within the R block, and one N6-like site at the R block-late region junction. BK(pm526), a small-plaque variant isolated from a BK(prototype)-induced hamster pineocytoma, contains three potential NF-BK binding sites, one N1-like site within the P block and two N4-like sites at P-Q junctions. BK(pm527), a different small-plaque variant, contains five potential NF-BK sites, including three N4-like sites and two N1-like sites (61–65).

There are a number of lines of evidence that suggest that the NF-BK sites are important for growth in culture. All strains of BKV apparently contain multiple sites. Often, there are many sites within relatively short stretches of DNA, for example, eight sites within 320 bp in BK(IR) and six sites within 370 bp in BK(MM). In enhancer-defective mutants containing only a single P block, duplication of a region encompassing the N1 site was observed (64). A role of NF-BK binding in early transcription of BKV is also consistent with recent findings showing that the binding of the same or a related protein is essential for mouse mammary tumor virus promoter function (6, 35, 42).

A systematic linker scan mutational analysis of the BKV early enhancer is found in the companion article (12). To eliminate the effect of sequence redundancy, these workers introduced their mutations into a viral background that has only a single P block. The background is identical to BK(WW), except that the R block is missing. This analysis shows a correlation between early transcriptional activity and the presence of sequences in the N1 and N4 sites (12 [note that the N4 site overlaps the genetically defined c element]). This is further evidence that the NF-BK sites defined by footprinting are functional elements of the BKV enhancer.

The recognition sequence for NF-BK (Table 1) is a subset of the recognition sequence for NF1, TGG(N$_{6-7}$) T/G CCAA (10, 13, 40). NF1 was originally described as a host protein required for adenovirus replication (45). NF1 appears to be functionally equivalent to the TGGCA-binding protein of the laboratory of Nowock and Sippel (39, 46, 47). Other work has suggested that NF1 is identical to the CCAAT transcription factor (29, 31) that binds in the upstream region of several promoters, although this remains controversial, and a recent study showed that there are a "multiplicity of CCAAT box binding proteins" that bind to these sequences in different promoters (14, 42). In view of this complexity, we have chosen to refer to the BKV-binding protein as NF-BK, until its identity is better established.

The presence of two binding sites for the transcription factor AP-1 that cross the junction between adjacent P blocks in BK(Dunlop) was surprising. Apparently, these binding sites were created at the time the P block triplication was formed. The possibility that functionally important elements span repeat junctions has been largely overlooked in prior studies of the BKV regulatory region, perhaps because this phenomenon is not known to occur in the SV40 enhancer.

The three AP-1 recognition sites in SV40, none of which are located at repeat junctions, are believed to contribute to induction of SV40 RNA synthesis by tetradecanoyl phorbol acetate (1, 38). The AP-1-binding sites in BKV may have similar functions. Studies with deletion mutants of BK(prototype) indicate that there is some loss of early gene expression when an AP-1 recognition sequence (site A2 in our nomenclature) is removed (11). AP-1 recognition sites are clearly not essential for viral viability, however, since they are not present in all strains.

BK(WW) and BK(MM) have a single binding site for the transcription factor Sp1 (17) at the junction of the Q and R blocks. This site is completely protected by affinity-purified Sp1, and has a 10 of 10 bp match to a known Sp1 site in the IE3 promoter of herpes simplex virus (30). This site is not present in BK(Dunlop) and is partially deleted in BK(prototype). In addition to the Sp1 site at the Q-R junction, there is a sequence AGGGAGGAGC at the early-proximal end of the P block that has an 8 of 10 match (underlined) to the Sp1 recognition consensus (32). Although there is genetic evidence that the region containing this potential Sp1 site may be important for early promoter function (11, 12), we have never detected binding at this site by affinity-purified Sp1 fractions in any of the strains tested. This suggests either that it is a very weak Sp1 site or that it is a recognition site for a different protein.

We have also observed protection of a late-proximal region in BK(WW) that overlaps the start codon for agnoprotein. This site, which we call L1, is protected by heparin-agarose fractions. Although there is some protection at L1 with purified NF-BK, competitive binding experiments suggest that protection at L1 arises primarily from a different factor. The L1 sequence is also present in the viral DNA of BK(Dunlop) and BK(MM), although it was not present in the subcloned probes used in our binding experiments. In a study using deletion mutants, deletion of L1 sequences had no effect on early gene expression (11), suggesting that the function of this site, if any, might be related to late gene expression. In SV40, there is evidence that sequences near and within the late coding region affect late gene expression (2, 20, 52, 53, 57, 60).

There has been considerable speculation about the role of motifs homologous to the SV40 enhancer core (67) and the E1a enhancer core (23) that are present within the regulatory region of BKV. The SV40 enhancer core sequence, TG GAAAGT, was originally defined by correlating loss of enhancer function with specific point mutations located within the 72-bp repeats and is now known to be the binding site for one or more cellular proteins (28, 43, 70). On the basis of the presence of similar sequences in Moloney murine sarcoma virus, BK(Dunlop), and polyomavirus, the SV40 enhancer core consensus (G)TGG A/T A/T A/T (G) was proposed (67). In the light of current knowledge, however, this fairly degenerate consensus may not be meaningful. It has been pointed out by Weber et al. (66) that the TGG A/T A/T A/T G motif occurs in at least 38 places within the SV40 genome. Moreover, the sequence in BK(Dunlop) matches (underlined) SV40 in only four of the eight positions (ATGGTTTG).

Our experiments show no evidence that the SV40 enhancer core homology in BKV is recognized by cellular proteins. Although the homology overlaps the late end of the NF-BK binding sites within the P block (sites N1, N2, N3, N7, and N9), only part of the sequence is protected. In addition, identical patterns of protection are seen with crude heparin-agarose fractions and DNA affinity-purified NF-BK, suggesting that NF-BK is the only factor binding in this region.

A second enhancer motif thought to be present in BKV, the E1a enhancer core, was originally defined by deletion mutants affecting expression of the adenovirus type 5 early region 1A (E1a) transcription unit (23). A consensus sequence

(A/C GGAAGTGA A/C) was proposed on the basis of sequence homologies with 13 other transcription units, BK (Dunlop) among them. This sequence (AGGAAAGTGCA) (matches are underlined) occurs in BK(Dunlop) at the late end of the P blocks and overlaps binding sites A1 and A2. However, not all the sequence is protected at these sites, and identical footprints are seen with both heparin-agarose fractions and DNA affinity-purified AP-1. The E1a enhancer core sequence is also present in BK(MM) at the early side of sites N4, N8, and N10 and in BK(WW) at the early side of sites N4, N5, and N6. Although the complete sequence is protected at sites N5 and N6, protection at these other sites does not cover the entire sequence, and identical footprints are seen with both heparin-agarose fractions and DNA affinity-purified NF-BK. These results suggest that the presence of E1a enhancer core homologies in BKV is actually due to the similarity between this consensus and the AP-1 and NF-BK recognition sequences and does not reflect the presence of an additional cellular E1a enhancer core-binding protein.

The companion article (12) reaches similar conclusions about the SV40 and adenovirus E1a homologies on the basis of independent evidence. For the most part, mutations that specifically affect these sequences have a less than twofold effect on early-direction transcription.

One question that may be raised about the studies presented here is whether the map of factor-binding sites is complete or whether some sites may have been overlooked. It is possible that some factors are present in the HeLa cell nucleus in very limiting amounts. In addition, some factors might not be recovered efficiently in our extraction or preliminary fractionation steps or might not bind to DNA under the conditions used for footprinting. Genetic evidence suggests that a GC-rich sequence of the early proximal end of the P block is required for early promoter function, but we have not detected protein binding in this region (11, 12). With this exception, however, the results of our binding studies are generally consistent with prior mutational analyses of the BKV regulatory region (11, 64) and with the linker scan analysis in the companion article (12). More than half of the DNA sequence in the regulatory region is protected by the factors defined here.

As noted previously, a hypothesis has been put forward by Rubinstein et al. (55) that BK(WW), which does not contain sequence repeats and does not grow in culture, represents the authentic virus as it exists in the human population. If this is correct, the occurrence of sequence repeats and deletions in different strains of BKV may be attributable to selective pressure in culture. The spontaneous generation of sequence repeats in response to selective pressure is consistent with the known tendency of enhancer mutants of BKV and SV40 to revert by sequence duplication (24, 64). It is also consistent with the observation that primary isolates of BKV usually require an extended period of incubation in culture before the first cytopathic effects are observed (7, 8, 15, 19, 21, 22, 26, 36, 50, 51, 58, 59, 69). This would allow time for repeats to be generated and fixed in the population.

It is unlikely that rearrangements during passage in culture are the only mechanism for repeat formation in BKV. BK(prototype) and BK(Dunlop), which were isolated independently, have identical repeat structures and differ only in the endpoint of a deletion on the late side of the regulatory region. Moreover, a third strain of BKV, BK-17, that was isolated by molecular cloning directly from the kidney tissue of an accident victim is identical to BK(Dunlop) in the regulatory region (R. Frisque, personal communication).

These findings demonstrate that sequence repeats are present in BKV in vivo in at least some individuals.

LITERATURE CITED

1. Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlick, and M. Karin. 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated transacting factor. Cell 49:729–739.
2. Ayer, D., and W. S. Dynan. 1988. Simian virus 40 major late promoter: a novel tripartite structure that includes intragenic sequences. Mol. Cell. Biol. 8:2021–2033.
3. Banerji, J., S. Rusconi, and W. Schaffner. 1981. Expression of a β-globin gene is enhanced by remote SV40 sequences. Cell 27:299–308.
4. Blaese, R. M., W. Strober, R. S. Brown, and T. A. Waldman. 1968. The Wiskott-Aldrich syndrome. A disorder with a possible defect in antigen processing or recognition. Lancet i:1056–1061.
5. Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.
6. Buetti, E., and B. Kuhnel. 1986. Distinct sequence elements involved in the glucocorticoid regulation of the mouse mammary tumor virus promoter identified by linker scanning mutagenesis. J. Mol. Biol. 190:379–389.
7. Caputo, A., A. Corallini, M. P. Grossi, L. Carra, P. G. Balboni, M. Negrini, G. Milanesi, G. Federspil, and G. Barbanti-Brodano. 1983. Episomal DNA of a BK virus variant in a human insulinoma. J. Med. Virol. 12:37–49.
8. Coleman, D. V., M. R. Wolfendale, R. A. Daniel, N. K. Dhanjal, S. D. Gardner, P. E. Gibson, and A. M. Field. 1980. A prospective study of human polyomavirus infection in pregnancy. J. Infect. Dis. 142:1–8.
9. Davison, B. L., T. Leighton, and J. C. Rabinowitz. 1979. Purification of Bacillus subtilis RNA polymerase with heparin-agarose. J. Biol. Chem. 254:9220–9226.
10. de Vries, E., W. van Driel, S. J. L. van den Heuvel, and P. C. van der Vliet. 1987. Contactpoint analysis of the HeLa nuclear factor I recognition site reveals symmetrical binding at one side of the DNA helix. EMBO J. 6:161–168.
11. Deyerle, K. L., J. A. Cassill, and S. Subramani. 1987. Analysis of the early regulatory region of the human papovavirus BK. Virology 158:181–193.
12. Deyerle, K. L., and S. Subramani. 1988. Linker scan analysis of the early regulatory region of human papovavirus BK. J. Virol. 62:3378–3387.
13. Diffley, J. F. X., and B. Stillman. 1986. Purification of a cellular, double-stranded DNA-binding protein required for initiation of adenovirus DNA replication by using a rapid filter-binding assay. Mol. Cell. Biol. 6:1363–1373.
14. Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis. 1987. A multiplicity of CCAAT box-binding proteins. Cell 50:863–872.
15. Dougherty, R. M., and H. S. DiStefano. 1974. Isolation and characterization of a papovavirus from human urine (38131). Proc. Soc. Exp. Biol. Med. 146:481–487.
16. 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.
17. Dynan, W., and R. Tjian. 1983. The promoter-specific transcrip-

tion factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell 35:79–87.

18. Fromm, M., and P. Berg. 1983. Simian virus 40 early- and late-region promoter functions are enhanced by the 72-bp repeat inserted at distant locations and inverted orientations. Mol. Cell Biol. 3:991–999.

19. Gardner, S. D., A. M. Field, D. V. Coleman, and B. Hulme. 1971. New human papovavirus (BK) isolated from urine after renal transplantation. Lancet i:1253–1257.

20. Ghosh, P. K., M. Piatak, J. E. Mertz, S. M. Weisman, and P. Lebowitz. 1982. Altered utilization of splice sites and 5' termini in late RNAs produced by leader region mutants of simian virus 40. J. Virol. 44:610–624.

21. Gibson, P. E., and S. D. Gardner. 1983. Strain differences and some serological observations on several isolates of human polyomaviruses. Prog. Clin. Biol. Res. 105:119–132.

22. Goudsmit, J., M. L. Baak, K. W. Slaterus, and J. van der Noordaa. 1981. Human papovavirus isolated from urine of a child with acute tonsillitis. Br. Med. J. 293:1363–1364.

23. Hearing, P., and T. Shenk. 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. Cell 33:695–703.

24. 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.

25. 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.

26. Howley, P. M., G. Khoury, J. C. Byrne, K. K. Takemoto, and M. M. Martin. 1975. Physical map of the BK virus genome. J. Virol. 16:959–973.

27. Howley, P. M., M. F. Mullarkey, K. K. Takemoto, and M. A. Martin. 1975. Characterization of human papovavirus BK DNA. J. Virol. 15:173–181.

28. Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight. 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. 1:133–146.

29. 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.

30. Jones, K. A., and R. Tjian. 1985. Sp1 binds to promoter sequences and activated HSV "immediate-early" gene transcription in vitro. Nature (London) 317:179–182.

31. Jones, K. A., K. R. Yamamoto, and R. Tjian. 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell 42:559–572.

32. Kadonaga, J. T., K. A. Jones, and R. Tjian. 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. 11:20–23.

33. Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA 83:5889–5893.

34. Khoury, G., and P. Gruss. 1983. Enhancer elements. Cell 33:313–314.

35. Kuhnel, B., E. Buetti, and H. Diggelman. 1986. Functional analysis of the glucocorticoid regulatory elements present in the mouse mammary tumor virus long terminal repeat. J. Mol. Biol. 190:367–378.

36. Lecatsas, G., and J. A. L. Van Wyk. 1978. DNA viruses in urine after renal transplantation. S. Afr. Med. J. 53:787–788.

37. Lee, W., A. Haslinger, M. Karin, and R. Tjian. 1987. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature (London) 325:368–372.

38. Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor AP-1 binds to TPA-inducible enhancer elements. Cell 49:741–752.

39. Leegwater, P. A. J., P. C. van der Vliet, R. A. W. Rupp, J. Nowock, and A. E. Sippel. 1986. Functional homology between the sequence-specific DNA-binding proteins nuclear factor I from HeLa cells and the TGGCA protein from chicken liver. EMBO J. 5:381–386.

40. Leegwater, P. A. J., W. van Driel, and P. C. van der Vliet. 1985. Recognition site of nuclear factor I, a sequence-specific DNA-binding protein from HeLa cells that stimulates adenovirus DNA replication. EMBO J. 4:1515–1521.

41. Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.

42. Miksicek, R., U. Borgmeyer, and J. Nowock. 1987. Interaction of the TGGCA-binding protein with upstream sequences is required for efficient transcription of mouse mammary tumor virus. EMBO J. 6:1355–1360.

43. Mitchell, P. J., C. Wang, and R. Tjian. 1987. Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell 50:847–861.

44. Moreau, P., R. Hen, B. Wasylyk, R. Everett, M. P. Gaub, and P. Chambon. 1981. The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res. 9:6047–6068.

45. Nagata, K., R. A. Guggenheimer, T. Enomoto, J. H. Lichy, and J. Hurwitz. 1982. Adenovirus DNA replication in vitro: identification of a host factor that stimulates synthesis of the preterminal protein-dCMP complex. Proc. Natl. Acad. Sci. USA 79:6438–6442.

46. 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.

47. Nowock, J., and A. E. Sippel. 1982. Specific protein-DNA interaction at four sites flanking the chicken lysozyme gene. Cell 30:607–615.

48. Nyborg, J., W. S. Dynan, I. S. Y. Chen, and W. Wachsman. 1988. Binding of host cell factors to DNA sequences in the HTLV-I LTR: implications for viral gene expression. Proc. Natl. Acad. Sci. USA 85:1457–1461.

49. Pagnani, M., M. Negrini, P. Reschiglian, A. Corallini, P. G. Balboni, S. Scherneck, G. Macino, G. Milanesi, and G. Barbanti-Brodano. 1986. Molecular and biological properties of BK virus-IR, a BK virus variant isolated from a human tumor. J. Virol. 59:500–505.

50. Pater, M. M., A. Pater, and G. Di Mayorca. 1981. Genome analysis of MG virus, a human papovavirus. J. Virol. 39:968–972.

51. Pauw, W., and J. Choufoer. 1978. Isolation of a variant of BK virus with altered restriction endonuclease pattern. Arch. Virol. 57:35–42.

52. Piatak, M., P. K. Ghosh, L. C. Norkin, and S. M. Weissman. 1983. Sequences locating the 5' ends of the major simian virus 40 mRNA late forms. J. Virol. 48:503–520.

53. Piatak, M., K. N. Subramanian, P. Roy, and S. M. Weissman. 1981. Late messenger RNA production by viable simian virus 40 mutants with deletions in the leader region. J. Mol. Biol. 153:589–618.

54. Rosenthal, N., M. Dress, P. Gruss, and G. Khoury. 1983. BK viral enhancer element and a human cellular homolog. Science 222:749–755.

55. 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.

56. Seif, I., G. Khoury, and R. Dhar. 1979. The genome of human papovavirus BKV. Cell 18:963–977.

57. Somasekhar, M. B., and J. E. Mertz. 1985. Sequences involved in determining the locations of the 5' ends of the late RNAs of simian virus 40. J. Virol. 56:1002–1013.

58. Takemoto, K. K., A. S. Rabson, M. F. Mularkey, R. M. Blaese, C. F. Garon, and D. Nelson. 1974. Isolation of papovavirus from brain tumor and urine of a patient with Wiskott-Aldrich syndrome. J. Natl. Cancer Inst. 53:1205–1207.

59. Ter Schegget, J., C. J. A. Sol, E. Wouters Baan, J. Van Der Noordaa, and H. Van Ormondt. 1985. Naturally occurring BK virus variants (JL and Dik) with deletions in the putative early enhancer-promoter sequences. J. Virol. 53:302–305.

60. Villarreal, L. P., R. T. White, and P. Berg. 1979. Mutational alterations within the simian virus 40 leader segment generate

altered 16S and 19S mRNAs. J. Virol. 29:209–219.

61. **Watanabe, S., S. Kotake, A. Nozawa, T. Muto, and S. Uchida.** 1982. Tumorigenicity of human BK papovavirus plaque isolates, wild-type and plaque morphology mutant, in hamsters. Int. J. Cancer 29:583–586.

62. **Watanabe, S., and K. Yoshiike.** 1982. Change of DNA near the origin of replication enhances the transforming capacity of human papovirus BK. J. Virol. 42:973–985.

63. **Watanabe, S., and K. Yoshiike.** 1985. Decreasing the number of 68-base-pair tandem repeats in the BK virus transcriptional control region reduces plaque size and enhances transforming capacity. J. Virol. 55:823–825.

64. **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.

65. **Watanabe, S., K. Yoshiike, A. Nozawa, Y. Yuasa, and S. Uchida.** 1979. Viable deletion mutant of human papovavirus BK that induces insulinomas in hamsters. J. Virol. 32:934–942.

66. **Weber, F., J. de Villiers, and W. Schaffner.** 1984. An SV40 "enhancer trap" incorporates exogenous enhancers or gener-

ates enhancers from its own sequences. Cell 36:938–992.

67. **Weiher, H., M. Konig, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science 219:626–631.

68. **Wildeman, A. G., M. Zenke, C. Schatz, M. Wintzerith, T. Gundstrom, H. Matthes, K. Takahashi, and P. Chambon.** 1986. Specific protein binding to the simian virus 40 enhancer in vitro. Mol. Cell. Biol. 6:2098–2105.

69. **Wright, P. J., G. Bernhardt, E. O. Major, and G. di Mayorca.** 1976. Comparison of the serology, transforming ability, and polypeptide composition of human papovaviruses isolated from urine. J. Virol. 17:762–775.

70. **Xiao, J.-H., I. Davidson, M. Macchi, R. Rosales, M. Vigneron, A. Staub, and P. Chambon.** 1987. In vitro binding of several cell-specific and ubiquitous nuclear proteins to the GT-I motif of the SV40 enhancer. Genes Dev. 1:794–807.

71. **Yang, R. C. A., and R. Wu.** 1979. BK virus DNA: complete nucleotide sequence of a human tumor virus. Science 206:456–462.