JOURNAL OF VIROLOGY, Sept. 1988, p. 3250–3257
0022-538X/88/093250-08$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 62, No. 9

Conservation of Sequence and Function between the Product of the 52-Kilodalton Immediate-Early Gene of Herpesvirus Saimiri and the BMLF1-Encoded Transcriptional Effector (EB2) of Epstein-Barr Virus

JOHN NICHOLAS, URSULA A. GOMPELS, MOLLY A. CRAXTON, AND ROBERT W. HONESS*

Division of Virology, National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom

Received 19 February 1988/Accepted 26 May 1988

We present a sequence of 2,220 nucleotides from a region of the genome of herpesvirus saimiri (HVS) which includes the coding and putative regulatory sequences for the 52-kilodalton (kDa) immediate-early (IE) phosphoprotein of the virus. The amino acid sequence predicted for this protein shows it to be homologous to the EB2 transcriptional effector encoded by the BMLF1 open reading frame of Epstein-Barr virus (EBV), the IE 68-kDa protein of varicella-zoster virus, and the IE 63-kDa (α27) protein of herpes simplex virus (HSV). By measuring the function of the HVS 52-kDa-protein gene in transient expression assays, we also showed that it can substitute with comparable efficiency for the EB2 product of EBV in the EB1-dependent activation of the EBV DR promoter. The α27 gene of HSV was an inefficient *trans*-activator in similar assays. We conclude that the IE 52-kDa protein of HVS is structurally and functionally more similar to the homologous protein of the human lymphotropic virus, EBV, than to the corresponding proteins from the neurotropic viruses, varicella-zoster virus and HSV.

Herpesviruses all have large (>100 kilobase pairs) double-stranded DNA genomes which encode more than 70 proteins. However, these genomes differ markedly in their gross composition and in the nature and organization of coding and noncoding sequences. Herpesviruses are also diverse in their biological properties, and a division into three major subgroups (alpha-, beta-, and gammaherpesviruses) has been proposed on the basis of these differences (18, 36). Thus, the alphaherpesviruses are neurotropic viruses typified by herpes simplex virus (HSV) and varicella-zoster virus (VZV), the betaherpesviruses are the cytomegaloviruses or salivary gland infectious agents (e.g., human cytomegalovirus), and the gammaherpesviruses are lymphotropic viruses exemplified by the B-cell-tropic human herpesvirus, Epstein-Barr virus (EBV), and by T-cell tropic viruses of the New World monkeys (e.g., herpesvirus saimiri [HVS]). We wished to understand the molecular basis for the differences between neurotropic and lymphotropic herpesviruses and undertook an analysis of the molecular biology of HVS, a virus with a clear-cut lymphotropism in vivo which also undergoes a lytic cycle of growth in nonlymphoid cells in tissue culture (12, 31).

Studies with representatives of each of the three major herpesvirus subgroups have shown that the expression of virus genes during productive cycles of virus growth occurs in three main temporal phases (immediate-early [IE] or α; early-delayed early [DE] or β; and late or γ). The first class of genes to be expressed, the IE genes, are defined as those that can be transcribed efficiently in the absence of de novo virus protein synthesis. The IE genes must therefore include the virus-coded transcriptional effectors that are required to permit high levels of transcription of the viral DE and late genes. Despite their critical role in the regulation of virus gene expression, there is evidence that the major IE genes of herpesviruses are not products of conserved genes. Thus, detailed analyses of the regulation of transcription from the HSV genome have shown that there are five IE genes (i.e., α4 or IE 175K; α0 or IE 110K; α22 or IE 68K; α27 or IE 63K; and α47 or IE 12K), but have implicated the product of one of these genes (IE 175K) as the major indispensable *trans*-activator of DE and late transcription (28–30). However, the protein products of the major IE genes of human cytomegalovirus are not detectably related to IE 175K or to any of the other IE genes of HSV (41, 42), and there is no homolog of IE 175K encoded by the human lymphotropic herpesvirus, EBV (1, 9). At least two EBV gene products have been shown to be capable of activating lytic cycle gene expression or of activating homologous or heterologous promoters in transient expression assays; the product of a spliced transcript from the BZLF1 reading frame (EB1) mediates activation of a lytic cycle in cells latently infected with EBV and is required for the efficient *trans*-activation of an early-DE EBV promoter by the BMLF1-encoded protein (EB2) (5, 7). The product of BZLF1 has no known homolog in other herpesvirus genomes, but the product of BMLF1 is homologous to the IE 63K (α27) protein of HSV and the IE 68K (reading frame 4) protein of VZV (8, 9).

Our studies on the synthesis and properties of virus-specified proteins in cells productively infected with HVS identified a single major IE protein with an apparent molecular weight of 52,000 (52K protein) (32). Preliminary results of hybrid selection and in vitro translation experiments suggesting a location of the 52K gene in *Eco*RI-I and adjacent *Eco*RI-E (17) were confirmed by immune precipitation of the in vitro translation products with a monoclonal antibody directed against the IE 52K protein (32, 33; E. P. Smith, Ph.D. thesis, Council for National Academic Awards, London, United Kingdom, 1986; W. Hell, E. P. Smith, and R. W. Honess, unpublished data). Subsequent results from partial sequencing of *Eco*RI-I located parts of the 52K coding region and suggested that the gene was homologous to the BMLF1 (EB2) gene of EBV (14). In this

* Corresponding author.

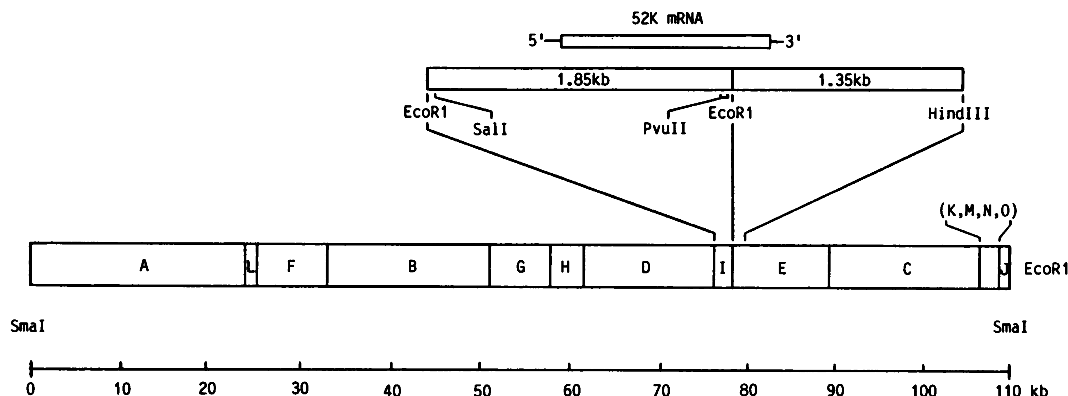


FIG. 1. Location of the major transcription unit serving the gene for the IE 52K protein of HVS relative to relevant restriction endonuclease cleavage sites within the *Eco*RI I and adjacent *Eco*RI E fragments (upper lines) from the light (L)-DNA coding sequences of the HVS genome (lower lines; the L-DNA component is conventionally defined by *Sma*I sites at 0 and 111 kilobase pairs and the conventional right end by *Eco*RI plus *Sma*I-J). kb, Kilobases.

report, we present the complete sequence of the HVS gene and show that the IE 52K protein is homologous to the EBV protein encoded by BMLF1 and that the IE 52K gene product can act as an effector of gene expression which will substitute for the EBV protein in a transient expression assay.

MATERIALS AND METHODS

Cells and transfections. Vero cells (Flow Laboratories, Inc., McLean, Va.) were grown at 37°C as monolayers on 80-cm^2 culture flasks in Dulbecco modified Eagle medium containing 10% newborn calf serum. For transfection, confluent monolayers were passaged 1:4 the previous day. Transfections were performed essentially as described by Graham and van der Eb (16), using supercoiled plasmid DNA (constructs and amounts specified in text) to transfect 10^7 cells. DNA was mixed with 0.5 ml of 250 mM $CaCl_2$, and this was added dropwise to 0.5 ml of 2× HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)-buffered saline (1% HEPES, 1.6% NaCl, 0.074% KCl, 0.25% Na_2HPO_4, 0.2% dextrose) at pH 7.05. The solution was mixed by shaking, and the DNA-calcium phosphate coprecipitate was allowed to form for 2 to 3 min before addition to the culture medium (50 ml). The cells were left at 37°C for 6 h and were then glycerol shocked by tipping off the medium, adding 1 ml of 25% glycerol in culture medium to the cells for 1 min, and rinsing twice, each time with 25 ml of fresh medium. Medium containing 10% calf serum was added to the cells, which were then incubated at 37°C until ready to harvest.

Assays of CAT activity. Cells were harvested and assayed for chloramphenicol acetyltransferase (CAT) activity essentially as described by Gorman et al. (15). At 60 to 65 h after transfection, cells were washed with phosphate-buffered saline and scraped into 1 ml of 0.05 M Tris hydrochloride (pH 7.5)–1 mM EDTA–0.15 M NaCl. The cells were pelleted by low-speed centrifugation, suspended in 100 μl of 25 mM Tris hydrochloride (pH 7.5), sonicated briefly, and centrifuged at 11,600 × *g* for 5 min. A 50-μl portion of the supernatant was added to 70 μl of 1 M Tris hydrochloride (pH 7.5)–20 μl of 4 mM acetyl coenzyme A (Sigma Chemical Co., St. Louis, Mo.)–0.2 μCi of [^{14}C]chloramphenicol (50 mCi/mmol; Amersham International) and incubated at 37°C for 90 min. Chloramphenicol and its acetylated products were extracted into 0.5 ml of ethyl acetate, which was then evaporated under a vacuum. The pellets were dissolved in 10 μl of ethyl acetate and applied to plastic-backed chromatography plates (Polygram SilG; Camlab). The different acetylated and nonacetylated forms of chloramphenicol were separated by ascending chromatography in chloroform-methanol (95:5, vol/vol) solvent, and the products were visualized by autoradiography. Percent acetylations were determined by scintillation counting, chloramphenicol-3-acetate, chloramphenicol-1-acetate, and chloramphenicol-1,3-diacetate being counted as a single acetylated product.

Plasmids. Plasmids pSVSM and pSVZ1 (encoding the EBV BMLF1-BSLF2 and BZLF1 open reading frames, respectively) and pK-CAT (containing the EBV DR promoter fused to the *cat* gene) were a gift from A. Sergeant and have been described previously (5). The plasmids pGR162 and pIGA-15 contain the HSV type 2 (HSV-2) ICP27 (27) and HSV-1 ICPO (13) genes, respectively, and pPOH3 contains the HSV-1 thymidine kinase gene promoter fused to *cat* (27). These plasmids were kindly provided by P. O'Hare. The plasmid pEcoRI-(I/E') contains the HVS IE 52K gene and was constructed by insertion of the 1.35-kilobase *Hin*dIII-*Eco*RI fragment from the left terminus of *Eco*RI-E between the *Hin*dIII and *Eco*RI sites of pUC8 (to give pEcoRI-E'), followed by ligation of *Eco*RI-I into the *Eco*RI site of pEcoRI-E'. The plasmid pEcoRI-(I*/E') contains *Eco*RI-I in the inverse of the native orientation, thus disrupting the 52K open reading frame (Fig. 1 and 2).

DNA sequencing. The *Eco*RI-I fragment of the HVS genome (Fig. 1) was obtained from a plasmid clone comprising *Eco*RI-I in pACYC184 (19). The sequence of *Eco*RI-I was determined by using the dideoxy-chain termination sequencing method of Sanger et al. (39, 40) to sequence random subfragments of *Eco*RI-I, essentially as described by Bankier and Barrell (2). Sequencing of the 3' end of the IE 52K gene, contained in *Eco*RI-E, and the *Eco*RI-I-E junction was done by using a directed approach, utilizing *Pvu*II and *Hin*dIII sites in *Eco*RI-I and *Eco*RI-E, respectively, to clone the *Eco*RI-*Hin*dIII and *Pvu*II-*Hin*dIII fragments from pEcoRI-(I/E') into M13mp19 (Fig. 2). The sequence data were assembled and analyzed as previously described (4).

RESULTS

Nucleotide sequence of the region of the 52K gene. Previous results from limited DNA sequencing from convenient re-

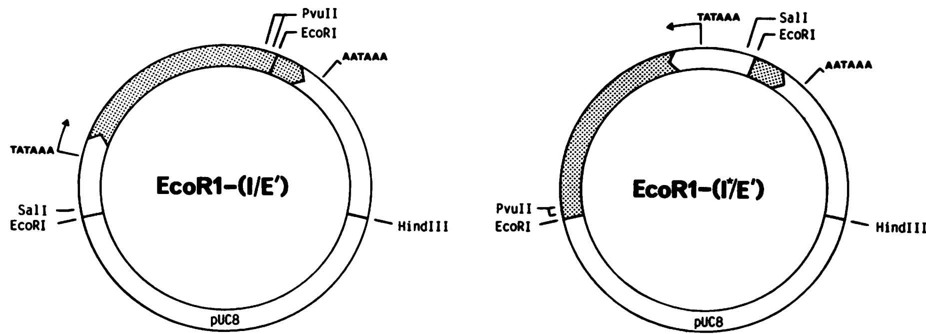


FIG. 2. Structure of a plasmid containing the intact regulatory and coding sequences of the 52K gene [pEcoRI-(I/E'); or p52K, see Fig. 5] and a plasmid in which the *Eco*RI I fragment is inverted, resulting in truncation of the reading frame for the 52K gene and removal of the appropriately orientated adenylation signal [pEcoRI-(I*/E')]. The location and orientation of TATA (TATAAA) and polyadenylation (AATAAA) consensus sequences and the open reading frame of the 52K gene (stippled areas, orientation indicated by arrowheads) are based on the DNA sequence shown in Fig. 3.

striction sites in the light (L)-DNA component of HVS DNA comfirmed the location of the 52K gene and permitted the prediction of its boundaries (14). On the basis of this prediction, we determined the complete sequence of the *Eco*RI-I fragment of HVS DNA and the adjacent region of *Eco*RI-E. This sequence and aspects of its interpretation are summarized in Fig. 3. The sequence contains two long open reading frames each transcribed from left to right on the prototype orientation of the HVS genome (Fig. 1 and 3). Both of these reading frames have a codon bias similar to that of other HVS genes (data not shown; see reference 4). The first of these reading frames (PRO1) extends through the left *Eco*RI site (position 1) to position 617, and the second, encoding the 52K gene product, starts at position 817 and ends (within *Eco*RI-E) at position 2052. Database searches of EBV- and VZV-encoded proteins (1, 8) confirmed that the product of the first HVS reading frame (PRO1) is homologous to the product of the BSLF1 reading frame of EBV and that the predicted product of the 52K reading frame is homologous to the BMLF1-encoded protein of EBV (see below and Fig. 4A).

Some interesting features are indicated on the DNA sequence. Downstream of the first open reading frame, but before the 52K coding sequences, is a poly(A) recognition sequence (position 616), followed six nucleotides 3' by a putative TATA box promoter element (TATAAA), which could serve the 52K gene. There is another poly(A) recognition site at position 2160 and a G+T-rich sequence beginning at position 2193. The likely poly(A) addition site of the 52K gene transcript would be predicted to occur between these sequence motifs (3, 26). The sequence also contains two potential binding sites for the transcription factor AP1 (22, 23) at positions 266 (CTGACTCA) and 703 (GT GACTAA), the former in inverse orientation (Fig. 3). Single AP1-binding sequences are also found in the 5'-flanking regions of the homologous genes in EBV (BMLF1 [24, 43]) and VZV (8). In the entire genome sequences of VZV (125 kilobases) and EBV (175 kilobases), there are only 5 and 13 AP1-binding sequences, respectively. Since these sites are rare within EBV and VZV and within the other regions of the HVS genome that have been sequenced (40 kilobase pairs; unpublised data), their occurrence upstream of the homologous 52K, BMLF1, and RF4 reading frames may be significant. The 5' regions of both the 52K and BMLF1 genes also contain two nonidentical palindromic sequences (a 5-mer and 11-mer in HVS, and 6-mer and 9-mer in EBV), but the significance of these motifs is unknown.

Conserved and nonconserved features of predicted sequence of 52K protein. Comparisons of the predicted amino acid sequences of the HVS reading frames with the homologous proteins of EBV and VZV by using a dot matrix program to display regions of similar sequence are illustrated in Fig. 4A. It is qualitatively apparent from these comparisons that the sequences of the HVS proteins are more similar to their EBV-encoded counterparts than they are to the homologous products of VZV. These displays also show that the most conserved regions of the proteins are in their carboxy-terminal portions and that the amino-terminal regions of the 52K protein and its homologs have some degenerate sequence elements. A more precise illustration of the similarities between the HVS 52K protein and its homologs in EBV (BMLF1 [EB2]) and VZV (RF4 [IE 68K protein]) is provided by the alignment shown in Fig. 4B and by the summaries of the main properties of these sequences given in Table 1. Thus, the N-terminal regions of these proteins contain an acidic region followed by a basic region which is rich in serine and arginine. In HVS and EBV, this region is also proline rich. The C-terminal portions of the proteins are more conserved; IE 52K and the BMLF1-encoded protein share 42% similarity of amino acid sequence, and there is 37% similarity between IE 52K and the RF4 product. Importantly, this region also contains three cysteine residues conserved in all three proteins and five conserved cysteines between IE 52K and the BMLF1 protein.

Nuclear localization signals have been defined for a number of nuclear proteins, including the adenovirus E1A (20) and polyomavirus large T (35) proteins. These signals are characterized by a preponderance of basic residues, lysine, and arginine, and many also contain proline residues (e.g., adenovirus type 5 E1A, KRPRP; polyomavirus large T, RKRPRP and PKKARED). Similar motifs are also found within the proline-arginine repetitive N-terminal region of IE 52K (KRERQR, KRPR) and the BMLF1-encoded protein (PRKK, KRRR) and may correspond to nuclear localization signals for these proteins. In addition, there are several potential threonyl-seryl phosphorylation acceptor sites (10), mainly within the repetitive region of the N-terminal portion of the proteins. Both of these products are known to be phosphorylated and to localize in the nuclei of infected cells (6, 34, 43).

Substitution of IE 52K protein for EB2 in activation of transcription from EBV DR promoter. Others have shown that the protein encoded by the BMLF1 open reading frame of EBV is a potent *trans*-activator of heterologous genes, its

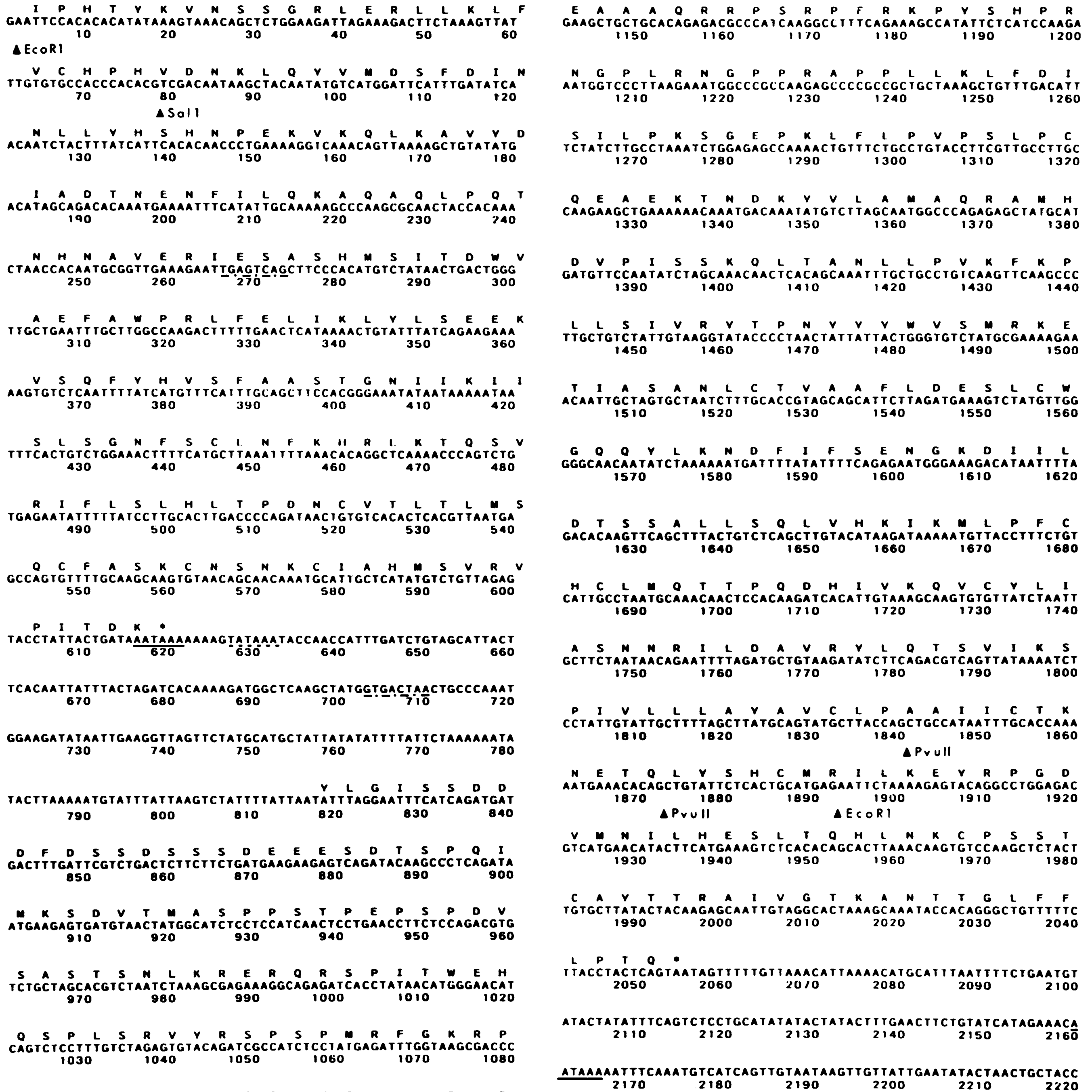
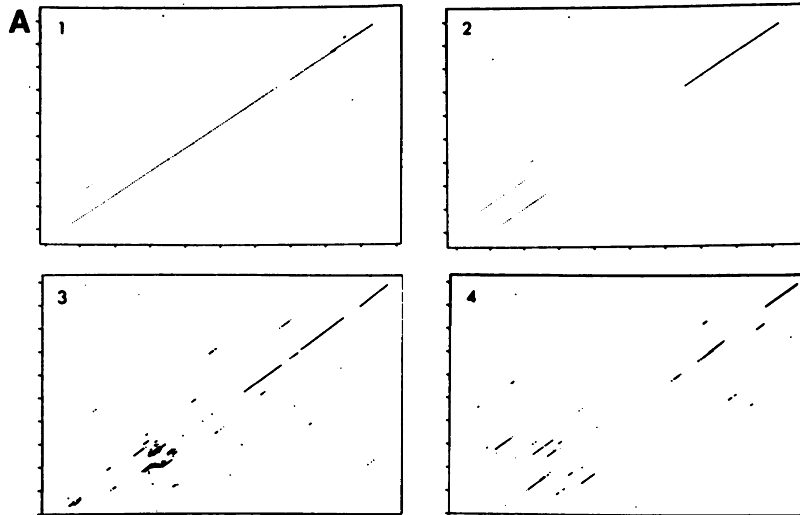


FIG. 3. Sequence of 2,220 nucleotides from the *Eco*RI I and the adjacent *Eco*RI E fragments of HVS DNA (Fig. 1) with the predicted protein sequences of the major open reading frames (single-letter code, upper case) and the location of some noncoding features annotated on the sequence. The sequence contains two major open reading frames; the first (PRO1) crosses the left-hand *Eco*RI site of *Eco*RI-I and is homologous to 250 residues from the carboxy terminus of the BSLF1 reading frame of EBV. The second open reading frame encodes the 52K protein (IE 52K) and is homologous to the BMLF1 reading frame of EBV (see Fig. 4A and B). Adenylation signals (AATAAA), the TATA box (TATAAA), and AP1-binding consensus sequences (TGAGTCAG and GTGACTAA) are underlined as indicated.

own promoter, and at least one EBV early gene (5, 24, 43). Specifically, the BMLF1-BSLF2-encoded protein (EB2) is able to induce transcription from the DR promoter (21) contained in the plasmid pK-CAT when cotransfected together with a plasmid containing the EBV BZLF1 open reading frame into EBV-negative cells (5). This *trans*-activation is seen only in the presence of BZLF1. The results shown in Fig. 5A demonstrate that the 52K protein can substitute for EB2 in the activation of pK-CAT in transient expression assays. Samples (1, 5, and 10 μg) of pEcoRI-(I/E') (p52K) were cotransfected into Vero cells with 1 μg of pSVZ1 (containing the BZLF1 open reading frame under the



B

```
                                                      40
HVS  ---------------------YLGISSDDDFDSSDSSSDE
EBV  -ESHILELEAVSDTNTDCDLDPMEGSEEHSTDGEISSSEE
VZV  MASASIPTDPDVSTICEDFMNLLPDEPSDDFALEVTDWAN

                                                      80
     EESDTSPQ--IMKSDVTMASPPSTP---------------
     EDEDPTPAHAIPARPSSVVITPTSASFVIPRKKWDLQDKT
     DEAIGSTPGEDSTTSRTVYVERTADTAYNPRYSKRRHGRR

                                                     120
     ----------------------EPSPDVSASTSNLKRERQ
     VTLHRS-------PLCRDEDEKEETGNSSYTRGHKRRRGE
     ESYHHNRPKTLVVVLPDSNHHGGRDVETGYARIERGHRRS

                                                     160
     RSPITWE------HQSPLSRVYRSPSPMRFGKRPRISSNS
     VHGCTDESYGKRRHLPPGARAPRAPRAPRVPRAPR--SPR
     SRSYNTQSSRKHRDRSLSNRRRRPTT------------PP

                                                     200
     TSRSCKTSWADRVREAAAQRRPSRPFRKPYSHPRNGPLRN
     APRSNRATRGPRSESRGAGRSTRKQARQERSQR---PLPN
     AMTTGERNDQTHDESYRL-RFSKRDARRERI---------

                                                     240
     GPPRAPPLLKLFDISILPKSGEPKLFLPVPSLPCQEAEKT
     KPWFDMSLVKP-----VSKITFVTLPSPLASLTLEPIQDP
     RKEYDIPVDRITGRAIEV------VSTAGASVTIDSVRHL

                                                     280
     NDKYVLAMAQRAMHDVPIS-----------------SKQL
     FLQSMLAVA--AHPEIGAW-----------------QKVQ
     DETIEKLVVRYATIQEGDSWASGGCFPGIKQNTSWPELML

                                                     320
     TANLLPVKFKPLLSIVRYTPNYYYWVSMRKETIASANLCT
     PRHELRRSYKTLREFFTKSTNKDTWLDARMQAIQNAGLCT
     YGHELYRTFESYKM------DSRIARALRERVIRGESLIE

                                                     360
     VAAFLDESLCWGQQYL-KNDFIFSENGK-DIILDTSSALL
     LVAMLEETIFWLQEITYHGDLPLAP--AEDILLACAMSLS
     ALESADELLTWIKMLAAKN-LPIYTNNPIV-----ATSKS

                                                     400
     SQLVHKIKMLPFCHCLMQTTPQDHIVKQVCYLIASNNRIL
     KVILTKLKELAP--CFLPNTRDYNFVKQLFYITCATARQN
     LLENLKLKLGPFVRCLLLNRDNDLGSRTLPELLRQ---QR

                                                     440
     DAVRYLQTSVIKSPIVLLLAYAVCLPAAII--CTKNETQL
     KVVETLSSSYVKQPLCLLAAYAAVAPAYINANCRRRHDEV
     FSDITCITTYMFVMIARIANIVVRGSKFVEYDDISCNVQV

                                                     480
     YSHCMRILKEYRPGDVMNILHESLTQHLNKCPSSTCAYTT
     EFLGH-YIKNYNPGTLSSLLTEAVETHTRDCRSASCSRLV
     -------LQEYTPGSCLAGVLEALITHQRECGRVECTLST

     RAIVGTKANTTGLFFLPTQ---  HVS
     RAILSPGTGSLGLFFVPGLNQ-  EBV
     WAGHLSDARPYGKYFKCSTFNC  VZV
```

TABLE 1. Summary of the general properties of the nucleotide and the predicted amino acid sequences of the HVS 52K gene and of the homologous genes of EBV and VZV

Properties	HVS	EBV	VZV
Gene product	IE 52K	BMLF1 (EB2)	RF4 (IE 68K)
Size of open reading frame (base pairs)	1,236	1,377	1,356
Composition (%G+C)	41	57	46
Observed %CpG/expected %CpG	1.38/4.11	4.47/8.21	5.90/5.35
Protein molecular wt	46,305	51,347	51,542
Hydrophobicity			
Entire sequence	−22.8	−68.8	−54.1
N-terminal half	−45.9	−72.5	−57.7
C-terminal half	+23.8	+3.8	+3.6
% Contribution of selected amino acids to N-terminal/C-terminal halves			
P (5.1)[a]	13.1/4.4	11.4/4.8	5.3/3.5
S (7.6)	16.5/7.3	11.4/6.1	9.7/7.5
R (5.0)	9.2/3.4	11.8/6.1	14.6/7.1
C (2.5)	1.0/4.9	1.3/3.5	0.4/4.0
% Similarity to HVS sequence			
N-terminal half		26	27
C-terminal half		42	37
C-terminal 50 amino acids		52	43

[a] Mean percent contribution of indicated amino acids in a large sample of eucaryotic proteins; from reference 25.

control of the simian virus 40 early promoter-enhancer) and 1 μg of pK-CAT as the target. The pK-CAT plasmid was also cotransfected with pSVSM (containing the EB2-coding sequences under the control of the simian virus 40 early promoter-enhancer) and with either pSVZ1 or pUC8 DNA alone. In all transfections, the amount of DNA was kept constant by the addition of pUC8 DNA.

The level of uninduced expression from pK-CAT seen in the absence of pSVSM or p52K was very low (<1% acetylation of chloramphenicol). However, when either pSVSM or p52K was cotransfected with pK-CAT and pSVZ1, there was a very substantial induction of CAT activity (approximately 20-fold and 70-fold, respectively, at 1 μg of *trans*-activator). That this activation by p52K is not due merely to a nonspecific effect of transfection of these DNA sequences was demonstrated by replacing p52K by pEcoRI-(I*/E') in these transfections. This plasmid contains *Eco*RI-I inverted relative to *Eco*RI-E' sequences, thus disrupting the most conserved terminal 50 amino acids of IE 52K (Fig. 2; Table 1), and was unable to induce expression from pK-CAT in this assay (data not shown). Also, the activation of pK-CAT by IE 52K is dependent on the presence of pSVZ1, comparable to the results of Chevallier-Greco et al. (5) using EB2 as a *trans*-activator.

Despite the dependence on the BZLF1-encoded protein (EB1) for EB2- and IE 52K-mediated *trans*-activation of expression from pK-CAT, the BMLF1-encoded product by itself is able to *trans*-activate heterologous promoters in transient expression assays (24, 43), indicating that the dependence on EB1 for induction of the DR promoter reflects the properties of this promoter, rather than those of the IE 52K and EB2 *trans*-activators. The results shown in Fig. 5B show that IE 52K is able to *trans*-activate the HSV thymidine kinase promoter in the absence of EB1. We also demonstrated that HSV ICP0, which is capable of efficient *trans*-activation of many heterologous as well as homologous promoters, is unable to induce high levels of CAT expression from pK-CAT.

HSV homolog of IE 52K, ICP27, does not efficiently *trans*-activate the DR promoter. The HSV homolog of the HVS IE 52K protein is the 63 kilodalton (kDa) IE protein ICP27 (8). We therefore tested the ability of ICP27 (contained in the plasmid pGR162) to activate the DR promoter of pK-CAT in cotransfections with pSVZ1. The results of these experiments (Fig. 5C) demonstrated that ICP27 is able to activate the EBV DR promoter, and that this activation depends on the presence of pSVZ1, as is the case for IE 52K- and EB2-induction of pK-CAT. Increasing the dose of pGR162 from 5 to 15 μg did not increase the level of induction seen, suggesting that maximum levels of induction were reached at a dose of 5 μg of pGR162 and that ICP27 is a relatively poor inducer of DR transcription.

FIG. 4. (A) Dot-matrix representations of comparisons between amino acid sequences of the first (PRO1; encoded by nucleotides 1 to 617 of Fig. 3) and second (reading frame for the 52K gene) open reading frames from the HVS DNA sequence with sequences of homologous proteins encoded by EBV and VZV. In each panel, the HVS amino acid sequence is represented on the vertical axis. Panel 1, Comparison between residues 8 to 212 of PRO1 and residues 638 to 842 of the EBV reading frame BSLF1. Panel 2, Comparison between PRO1 and residues 876 to 1079 of VZV reading frame 6. Panel 3, Comparison between residues 1 to 412 of the open reading frame for the 52K protein of HVS and residues 1 to 459 of the EBV reading frame BMLF1. Panel 4, Comparison of this 52K protein sequence with that of residues 1 to 452 from the IE 68-kDa product of reading frame 4 of VZV. The dot-matrix program used is similar in principle to the DIAGON program of Staden (40a) and has been described previously (4, 14). The comparisons illustrated were obtained with a sliding window of 31 residues and with all scores in excess of 10.8 per residue being recorded by a dot. (B) Alignments of the protein sequence from the major open reading frame for the 52K protein (HVS) with homologous proteins of EBV and VZV. Gaps introduced to give these alignments are shown by dashes, identical residues in pairwise comparisons are boxed, and residues which are the same in all three sequences are indicated by filled circles.

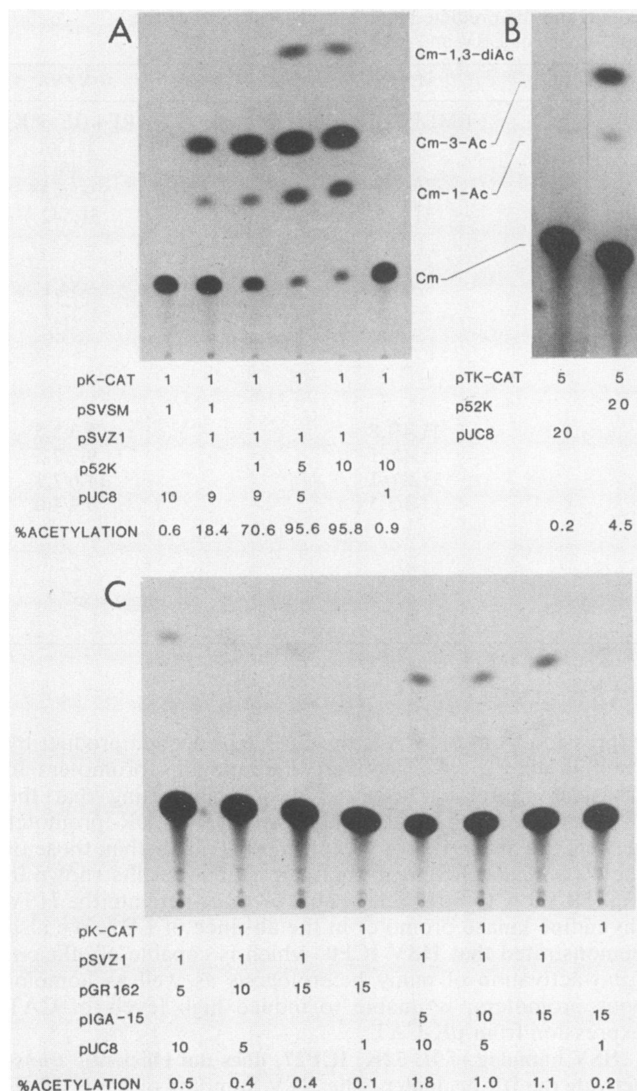


FIG. 5. Comparisons of *trans*-activation of EBV and HSV DE promoters by the IE 52K protein of HVS and homologous and nonhomologous transcriptional effectors of EBV and HSV. The figure shows measures of the CAT activity (Materials and Methods) accumulated in Vero cells transfected with the EBV DR promoter linked to the *cat* gene (pK-CAT) (A and C) or the HSV thymidine kinase promoter linked to the *cat* gene (pTK-CAT) (B) together with the indicated amounts (micrograms of each plasmid per 10^7 cells) of effector and control (pUC8) plasmids. In this figure, p52K refers to the construct containing the intact 52K gene [pEcoRI-(I/E') plasmid, Fig. 2]. The pEcoRI-(I*/E') plasmid had no effect in these assays (data not shown). Cm, Chloramphenicol; Cm-1-Ac, chloramphenicol-1-acetate; Cm-3-Ac, chloramphenicol-3-acetate; Cm-1,3-diAc, chloramphenicol-1,3-diacetate.

DISCUSSION

The complete sequence of the coding and upstream sequences of the 52K IE gene of HVS (Fig. 3) confirms and extends previous conclusions (14) that the product of this gene is homologous to the BMLF1-encoded protein (EB2) of EBV, the IE 63-kDa protein (ICP27 [27]) of HSV, and the IE 68-kDa protein of VZV (RF4). Moreover, we also showed that the HVS gene product, like its counterpart in EBV (5), is an efficient activator of expression from the EBV DR promoter in short-term expression assays. Although evidence from studies in vivo (37) and in vitro (11) suggests that ICP27 of HSV has effects on the regulation of late gene expression, plasmids bearing this gene and those expressing the 110-kDa IE gene of HSV (ICPO; a nonhomologous protein with the ability to increase expression from a range of promoters) were very inefficient activators of expression from the EBV DR promoter. Although these results are consistent with the possibility that the greater degree of protein sequence similarity between the HVS and EBV proteins is reflected in aspects of their function as effector molecules, there is no evidence that these effector functions involve a common mechanism. The observation that the actions of both these HVS and EBV genes on the DR promoter require the presence of the BZLF1 region (EB1) of EBV appears striking. However, this requirement was also a feature of the low-level activities observed in response to ICP27- and ICPO-expressing plasmids and may therefore be a precondition for activation of expression from this promoter by nonidentical pathways. Previous studies have also shown that the BMLF1 product can activate a range of other promoters without any requirement for BZLF1 (e.g., see references 24, 38, and 43), and we have found that the 52K gene can also increase expression from a number of heterologous promoters in the absence of EB2 (e.g., E3-, E4-, and major late promoters of adenovirus type 2; J. Nicholas, unpublished data). A detailed comparison of the essential sequences in these effectors and in responsive promoters will be necessary to establish the extent to which the conserved elements in the sequences of these proteins determine similarities in their effector functions.

ACKNOWLEDGMENTS

We thank P. O'Hare and A. Sergeant for generous gifts of plasmids and for their help and advice in the course of this project. We also thank D. McGeoch and A. J. Davison for computer tapes of sequences from HSV and VZV in advance of the appearance of these sequences in the public databases.

LITERATURE CITED

1. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310**:207–211.
2. **Bankier, A. T., and B. G. Barrell.** 1983. Shotgun sequencing, p. 1–33. *In* R. A. Flavell (ed.), Techniques in the life sciences, vol. B5. Elsevier/North-Holland Scientific Publishers, Ltd., Limerick, Republic of Ireland.
3. **Birnstiel, M. L., M. Busslinger, and K. Straub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41**: 349–359.
4. **Cameron, K. R., T. Stamminger, M. Craxton, W. Bodemer, R. W. Honess, and B. Fleckenstein.** 1987. The 160,000-M_r virion protein encoded at the right end of the herpesvirus saimiri genome is homologous to the 140,000-M_r membrane antigen encoded at the left end of the Epstein-Barr virus genome. J. Virol. **61**:2063–2070.
5. **Chevallier-Greco, A., E. Manet, P. Chavrier, C. Mosnier, J. Dallie, and A. Sergeant.** 1986. Both Epstein-Barr virus (EBV)-encoded trans-acting factors, EB1 and EB2, are required to activate transcription from an EBV early promoter. EMBO J. **5**: 3243–3249.
6. **Cho, M.-S., K. T. Jeang, and G. S. Hayward.** 1985. Localization of the coding region of an Epstein-Barr virus early antigen and inducible expression of the 60-kilodalton nuclear protein in transfected fibroblast cell lines. J. Virol. **56**:852–859.
7. **Countryman, J., and G. Miller.** 1985. Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a

small cloned subfragment of heterogeneous viral DNA. Proc. Natl. Acad. Sci. USA **82**:4085–4089.
8. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella-zoster virus. J. Gen. Virol. **67**:1759–1816.
9. **Davison, A. J., and P. Taylor.** 1987. Genetic relations between varicella-zoster virus and Epstein-Barr virus. J. Gen. Virol. **68**: 1067–1079.
10. **Edelman, A. M., D. K. Blumenthal, and E. G. Krebs.** 1987. Protein serine/threonine kinases. Annu. Rev. Biochem. **56**:567–613.
11. **Everett, R. D.** 1986. The products of herpes simplex type 1 (HSV-1) immediate-early genes 1, 2, and 3 can activate gene expression in trans. J. Gen. Virol. **67**:2507–2513.
12. **Fleckenstein, B., and R. C. Desrosiers.** 1982. Herpesvirus saimiri and herpesvirus ateles, p. 253–352. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
13. **Gelman, I. H., and S. Silverstein.** 1985. Identification of immediate-early genes from herpes simplex virus that trans-activate the virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **82**: 5265–5269.
14. **Gompels, U. A., M. A. Craxton, and R. W. Honess.** 1988. Conservation of gene organization in the lymphotropic herpesviruses herpesvirus saimiri and Epstein-Barr virus. J. Virol. **62**: 757–767.
15. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyl transferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
16. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
17. **Hell, W., S. Modrow, and H. Wolf.** 1985. Mapping of herpesvirus saimiri proteins on the viral genome: proteins dependent and not dependent on viral DNA synthesis. J. Virol. **56**:414–418.
18. **Honess, R. W.** 1984. Herpes simplex and "The Herpes Complex": diverse observations and a unifying hypothesis. J. Gen. Virol. **65**:2077–2107.
19. **Knust, E., S. Schirm, W. Dietrich, W. Bodemer, E. Kolb, and B. Fleckenstein.** 1984. Cloning of herpesvirus saimiri DNA fragments representing the entire L-region of the genome. Gene **25**: 281–289.
20. **Krippl, B., B. Furguson, N. Jones, M. Rosenberg, and H. Westphal.** 1985. Mapping the functional domains in adenovirus E1A proteins. Proc. Natl. Acad. Sci. USA **82**:7480–7484.
21. **Laux, G., K. Freese, and G. W. Bornkamm.** 1985. Structure and evoluation of two related transcription units of Epstein-Barr virus carrying small tandem repeats. J. Virol. **56**:987–995.
22. **Lee, W., A. Hasslinger, M. Karin, and R. Tjian.** 1987. Activation of transcription by two factors that bind promoter and enhancer sequences of human metallothionein gene and SV40. Nature (London) **325**:368–372.
23. **Lee, W., P. Mitchell, and R. Tjian.** 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell **41**:349–359.
24. **Leiberman, P. M., P. O'Hare, and G. S. Hayward.** 1986. Promiscuous *trans*-activation of gene expression by an Epstein-Barr virus-encoded early nuclear protein. J. Virol. **60**:140–148.
25. **Maruyama, T., T. Gojobori, S. Aota, and T. Ikemura.** 1986. Codon usage tabulated from the GenBank genetic sequence data. Nucleic Acids Res. **45**:151–197.
26. **McLaughlan, J., G. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. **13**:1347–1368.
27. **O'Hare, P., and G. S. Hayward.** 1985. Evidence for a direct role for both the 175,000- and 110,000-molecular weight immediate-early proteins of herpes simplex virus in the *trans*-activation of delayed-early promoters. J. Virol. **53**:751–760.
28. **Pearson, R. H., S. Bacchetti, and J. S. Smiley.** 1985. Cells that constitutively express the herpes simplex virus immediate-early protein ICP4 allow efficient activation of virus delayed-early genes in trans. J. Virol. **54**:414–421.
29. **Preston, C. M.** 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant *tsK*. J. Virol. **29**:275–284.
30. **Preston, V. G.** 1981. Fine-structure mapping of herpes simplex virus type 1 temperature-sensitive mutations within the short repeat region of the genome. J. Virol. **39**:150–151.
31. **Randall, R. E., R. W. Honess, and P. O'Hare.** 1983. Proteins specified by herpesvirus saimiri: identification and properties of virus-specified polypeptides in productively-infected cells. J. Gen. Virol. **64**:19–35.
32. **Randall, R. E., C. Newman, and R. W. Honess.** 1984. A single major immediate-early virus gene product is synthesized in cells productively infected with herpesvirus saimiri. J. Gen. Virol. **65**:1215–1219.
33. **Randall, R. E., C. Newman, and R. W. Honess.** 1984. Isolation and characterization of monoclonal antibodies to structural and nonstructural herpesvirus saimiri proteins. J. Virol. **52**:872–883.
34. **Randall, R. E., C. Newman, and R. W. Honess.** 1985. Asynchronous expression of the immediate-early protein of herpesvirus saimiri in populations of productively infected cells. J. Gen. Virol. **66**:2199–2213.
35. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44**:77–85.
36. **Roizman, B.** 1982. The family Herpesviridae: general description, taxonomy, and classification, p. 1–23. In B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
37. **Sacks, W. R., C. Greene, D. P. Aschman, and P. A. Schaffer.** 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. **55**:796–805.
38. **Sample, J., G. Lancz, and M. Nonoyama.** 1986. Mapping of genes in *Bam*HI fragment M of Epstein-Barr virus DNA that may determine the fate of viral infection. J. Virol. **57**:145–154.
39. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

40a. **Staden, R.** 1982. An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Res. **10**:2951–2961.
41. **Stenberg, R. M., D. R. Thomsen, and M. F. Stinski.** 1984. Structural analysis of the major immediate-early gene of human cytomegalovirus. J. Virol. **49**:190–199.
42. **Stinski, M. F., D. R. Thomsen, R. M. Stenberg, and L. C. Goldstein.** 1983. Organization and expression of the immediate-early genes of human cytomegalovirus. J. Virol. **46**:1–14.
43. **Wong, K.-M., and A. J. Levine.** 1986. Identification and mapping of Epstein-Barr virus early antigens and demonstration of a viral gene activator that functions in *trans*. J. Virol. **60**:149–156.