

# Nucleotide Sequence Analysis of a 38.5-Kilobase-Pair Region of the Genome of Human Herpesvirus 6 Encoding Human Cytomegalovirus Immediate-Early Gene Homologs and Transactivating Functions

JOHN NICHOLAS<sup>1,2\*</sup> AND MICHELLE E. D. MARTIN<sup>2†</sup>

*Johns Hopkins Oncology Center, Baltimore, Maryland 21231,<sup>1</sup> and Division of Virology, National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom<sup>2</sup>*

Received 26 July 1993/Accepted 20 October 1993

**Human herpesvirus 6 (HHV-6) is prevalent in the human population, with primary infection occurring early in life. Its predominant CD4<sup>+</sup> T-lymphocyte tropism, its ability to activate human immunodeficiency virus type 1 (HIV-1) gene expression in vitro, and its upregulation of CD4 expression has led to speculation that HHV-6 may act as a positive cofactor in the progression of HIV infection to AIDS in individuals infected with both viruses. Previous sequencing studies of restricted regions of the 161.5-kbp genome of HHV-6 have demonstrated unequivocally that it is a member of the betaherpesvirus subgroup and have indicated that the HHV-6 genome is generally collinear with the unique long (U<sub>L</sub>) component of human cytomegalovirus (HCMV). In the work described in this report we have extended these sequencing studies by determining the primary structure of 38.5-kbp of the HHV-6 genome (genomic position 21.0 to 59.5 kbp). Within the sequenced region lie 31 open reading frames, 20 of which are homologous to positional counterparts in HCMV. Of particular significance is the identification of homologs of the HCMV UL36-38 and US22-type genes, which have been shown to encode transactivating proteins. We show that DNA sequences encoding these HHV-6 homologs were able to transactivate HIV-1 long terminal repeat-directed chloramphenicol acetyltransferase expression in cotransfection assays, thus demonstrating functional as well as structural conservation of these betaherpesvirus-specific gene products. Our data therefore confirm the close relationship between HHV-6 and HCMV and identify putative immediate-early regulatory genes of HHV-6 likely to play key roles in lytic replication and possibly also in the interactions between HHV-6 and HIV in dually infected cells.**

Human herpesvirus 6 (HHV-6) has been identified as the causative agent of exanthem subitum, a mild febrile disease of infants (40, 98). There is some evidence that HHV-6 may be associated with other diseases including lymphoproliferative disorders and hepatitis (reviewed by Thomson et al. [90]). The predominant CD4<sup>+</sup> T-cell tropism of HHV-6 both in vivo and in vitro (50, 86) and the ability of HHV-6 and human immunodeficiency virus type 1 (HIV-1) to coinfect the same cell in vitro, leading to activated HIV-1 gene expression and accelerated cell death (48), have led to speculation that HHV-6 may be involved in the progression of HIV infection to AIDS. Further supporting evidence has come from studies of effects of HHV-6 infection and cotransfection of subgenomic fragments on activation of long terminal repeat (LTR)-directed transcription (21, 25, 34, 35, 52) and from the finding that HHV-6 infection of lymphocytes upregulates the expression of CD4, the major membrane receptor for HIV-1, thereby having the potential to expand the population of cells susceptible to HIV-1 infection (47, 49). Other data fail to support a correlation between HHV-6 antibody titers and the predisposition to HIV infection or progression of HIV infection to AIDS (78), and one study has shown the suppression of HIV-1 replication by coinfection with HHV-6 (7). The bases for these apparently contradictory findings are at present unclear but are probably

due in part to differences in cell culture conditions, applied infectious doses, and the specific properties of different HHV-6 isolates.

As part of our investigations into the molecular biology of HHV-6, we have undertaken large-scale sequencing of a region of the genome of strain U1102 comprising sequences between genomic positions 21.0 and 59.5 kbp (53). The complete nucleotide sequences of several herpesviruses are known. These include the human viruses Epstein-Barr virus (EBV [4]), varicella-zoster virus (VZV [17]), herpes simplex virus type 1 (HSV-1 [55–57, 70]), and human cytomegalovirus (HCMV [11]) and the simian herpesvirus saimiri (HVS [2, 63]). These sequence analyses, together with results from selected sequencing of other herpesvirus genomes, e.g., HHV-6 (20, 27, 37, 44, 52, 61, 62, 88, 89), bovine herpesvirus 4 (6, 91, 92), and murine herpesvirus 68 (19), have provided a comprehensive and extremely informative set of data with which to study the relationships between different herpesviruses, the features common to and characteristic of viruses within a particular herpesvirus subgroup (alpha, beta, or gamma [73]), and the molecular biological determinants of their diverse in vivo biological properties. The three main subgroups of herpesviruses (alpha, beta, and gamma) were originally assigned on the basis of various in vitro and in vivo biological properties (31, 33), and it is now clear through DNA sequence analyses that these biological groupings largely reflect the phylogenetic relationships among the herpesviruses. Thus, within a subgroup, herpesvirus genetic content, gene organization, and degree of conservation of homologous encoded gene products are more similar than they are between herpesviruses of

\* Corresponding author. Mailing address: Johns Hopkins Oncology Center, 418 N. Bond St., Baltimore, MD 21231. Phone: (410) 550-6801. Fax: (410) 550-6802.

† Present address: Molecular Biology Institute, University of California Los Angeles, Los Angeles, CA 90024-1570.

different subgroups. Perhaps most notable is the discovery of subgroup-specific immediate-early (IE) regulatory genes common to all members of the same herpesvirus subgroup (see, e.g., references 55, 58, 65, 75, and 92). Certain noncoding features, e.g., lack or presence of local or global CpG dinucleotide suppression (32), are also conserved within herpesvirus subgroups. Previous sequencing studies of HHV-6 have demonstrated that the genetic content and organization of the HHV-6 genome is very similar to that found in HCMV, a member of the betaherpesvirus subgroup, and that the HHV-6 genome is essentially collinear with the unique long ( $U_L$ ) component of HCMV (20, 44, 52, 61, 89). This relationship between HHV-6 and HCMV enabled us previously to predict the location of the HHV-6 equivalent of the HCMV major IE locus, a region which in HHV-6 and all other characterized betaherpesviruses is markedly deficient in CpG dinucleotides and which encodes transactivating functions (32, 52).

Here we present the nucleotide sequence analysis of a contiguous 38.5-kbp region of the HHV-6 genome (positions 21.0 to 59.5kbp) predicted on the basis of collinearity between HHV-6 and HCMV to encode homologs of the HCMV IE UL36-38 and US22 genes (11, 14, 41, 83). The results of this work have indeed identified HHV-6 homologs of these and other positional counterparts in HCMV, thereby providing further evidence of the close genetic relationship between HHV-6 and HCMV and providing evidence of a new family of herpesvirus subgroup-specific IE regulatory genes common to the betaherpesviruses. Our data also identify HHV-6 genes with no counterparts in any other characterized herpesvirus. Furthermore, the use of effector plasmids containing the HHV-6 UL36-38 and US22 homologous coding sequences together with an HIV-1 LTR-chloramphenicol acetyltransferase (CAT) target plasmid in cotransfection assays has revealed functional as well as structural conservation of these betaherpesvirus gene products and has identified HHV-6 genes of potential importance in HHV-6-HIV-1 interactions *in vivo*.

## MATERIALS AND METHODS

**Virus and tissue culture.** Vero cell monolayers were grown at 37°C in Dulbecco modified Eagle medium containing 10% newborn calf serum and antibiotics (penicillin and streptomycin; 10 and 20 µg/ml, respectively). Peripheral blood lymphocytes were isolated from blood of healthy adults by centrifugation on Ficoll-Paque (Pharmacia) gradients and grown under 5% CO<sub>2</sub> at 37°C in RPMI 1640 medium containing 10% fetal calf serum and antibiotics. HHV-6 U1102 was propagated in phytohemagglutinin-stimulated peripheral blood lymphocytes as previously described (53).

**Virus DNA and recombinant plasmids.** All of the sequencing studies reported here were carried out with DNA derived from HHV-6 U1102. Propagation of the virus and subsequent purification of viral DNA, for use in PCR or cycle sequencing (15), were as previously described (53). Viral DNAs used for sequencing were fragments *EcoRI*-F (ED7), *EcoRI*-J, *BamHI*-Q, *EcoRI*-P (ED4), *HindIII*-H (HD9), *XhoI*-I, and two PCR-derived DNA fragments spanning the *SmaI*-G-*EcoRI*-F and *EcoRI*-P-*HindIII*-H regions (53). The isolation and characterization of the clones pED7, pED4, and pHD9 have been described previously (53). The *BamHI*-Q fragment was isolated from a λ phage clone (ZB70.6), kindly provided by I. Teo and M. Jones (88), and subcloned into the *BamHI* site of pUC19. The *EcoRI*-J fragment was cloned from HHV-6 U1102 DNA following *EcoRI* digestion and gel extraction of 5.6-kbp DNA for ligation into the *EcoRI* site of pUC19. The

identity of the cloned DNA was confirmed by Southern hybridization as described previously (53). The *SmaI*-G-*EcoRI*-F PCR fragment was cloned into the *BamHI* site of pUC19 for bacterial amplification prior to sequencing; the PCR products that made up the *EcoRI*-P-*HindIII*-H region were sequenced directly (without prior cloning) by M13 shotgun sequencing techniques (5). Plasmids pIEGP1-2 and pIEGP2-3' contain sequences encoding open reading frames (ORFs) EJLF4 to EJLF6 and EFLF1 plus EFLF2, respectively. They were constructed following PCR amplification of appropriate sequences from HHV-6 U1102 DNA by using the primer pairs IEGP1-IEGP2 and IEGP2-IEGP3 (see below). The product of the IEGP1-IEGP2-primed PCR (1.89 kbp) was digested with *HindIII* and *BamHI* (which cut within IEGP1 and IEGP2 sequences, respectively) and cloned between these sites in pUC19. PCR amplification with IEGP2 and IEGP3 produced a 5-kbp product encompassing the EFLF2 to EJLF4 coding sequences, which was digested with *BamHI* (to cleave between the EJLF6 and EPLF1 ORFs) and *HindIII* (to cleave within IEGP3 sequences). The EFLF1-plus-EFLF2-containing sequences (see Fig. 5A) were cloned between the *BamHI* and *HindIII* sites of pUC19. The expression plasmid pEPLF3 was constructed by cloning PCR-amplified EPLF3 coding sequences (with oligonucleotides LF4A and LF4B [see below]) in the appropriate orientation between the *HindIII* and *BamHI* sites of pBC12 (16). The target plasmid pLTR-CAT contains the entire LTR of HIV-1, the *cat* gene from pSV2-CAT (29), and the polyadenylation/3'-processing sequences from the HCMV major IE gene (1) cloned into a modified pUC18 vector and was a gift from A. Krigger.

**Oligonucleotides and PCR.** The following oligonucleotides were used for the isolation of specific HHV-6 sequences for cloning into bacterial vectors: IEGP1 (5'-acacaagcTTGACGCAGGTGAGTAAGTCT-3', nucleotides 6063 to 6083), IEGP2 (5'-acacggatccATAGCGTTCAAAGGTC-3', nucleotides 8955 to 8939), IEGP3 (5'-acacaagcTTCATAACA CATATACAATGGT-3', nucleotides 3909 to 3930), LF4A (5'-acacggatCCTACCCTCTTTCAATACAATT-3', nucleotides 13653 to 13674), and LF4B (5'-acacaagcctAGCCGGGCAGTTCATCGTTAT-3', nucleotides 14623 to 14603). These oligonucleotides include non-HHV-6-complementary sequences (lowercase letters) at their 5' ends, which were incorporated to generate restriction endonuclease sites at the termini of the PCR products and thus facilitate their subsequent cloning. PCR was performed on HHV-6 U1102 DNA template essentially as described previously (53), with extension times for IEGP1-IEGP2, IEGP2-IEGP3, and LF4A-LF4B primer pairs being 4, 10, and 2 min, respectively. The resulting 1.89-, 5.05-, and 0.97-kbp PCR products were cleaved with *HindIII* and *BamHI*, and appropriate fragments were isolated for cloning into pUC19 (see Fig. 5A) or pBC12.

**M13 subcloning and DNA sequencing.** Nucleotide sequences of the subgenomic fragments of HHV-6 DNA were determined by the dideoxynucleotide-chain termination method (76, 77) essentially as described by Bankier et al. (5). The region (150 bp) between *BamHI*-Q and *EcoRI*-P was sequenced directly from the HHV-6 DNA template by the method of Craxton (15). Double-stranded sequencing of specific regions of cloned HHV-6 DNA was done to confirm M13-derived sequences in "single-stranded" regions of the contiguous sequences or at fragment junctions and was carried out with the Sequenase version 2.0 kit (United States Biochemical Corp.) or by a modification of the M13 sequencing protocol (76) with Klenow DNA polymerase. The latter involved heat denaturation of double-stranded plasmid DNA

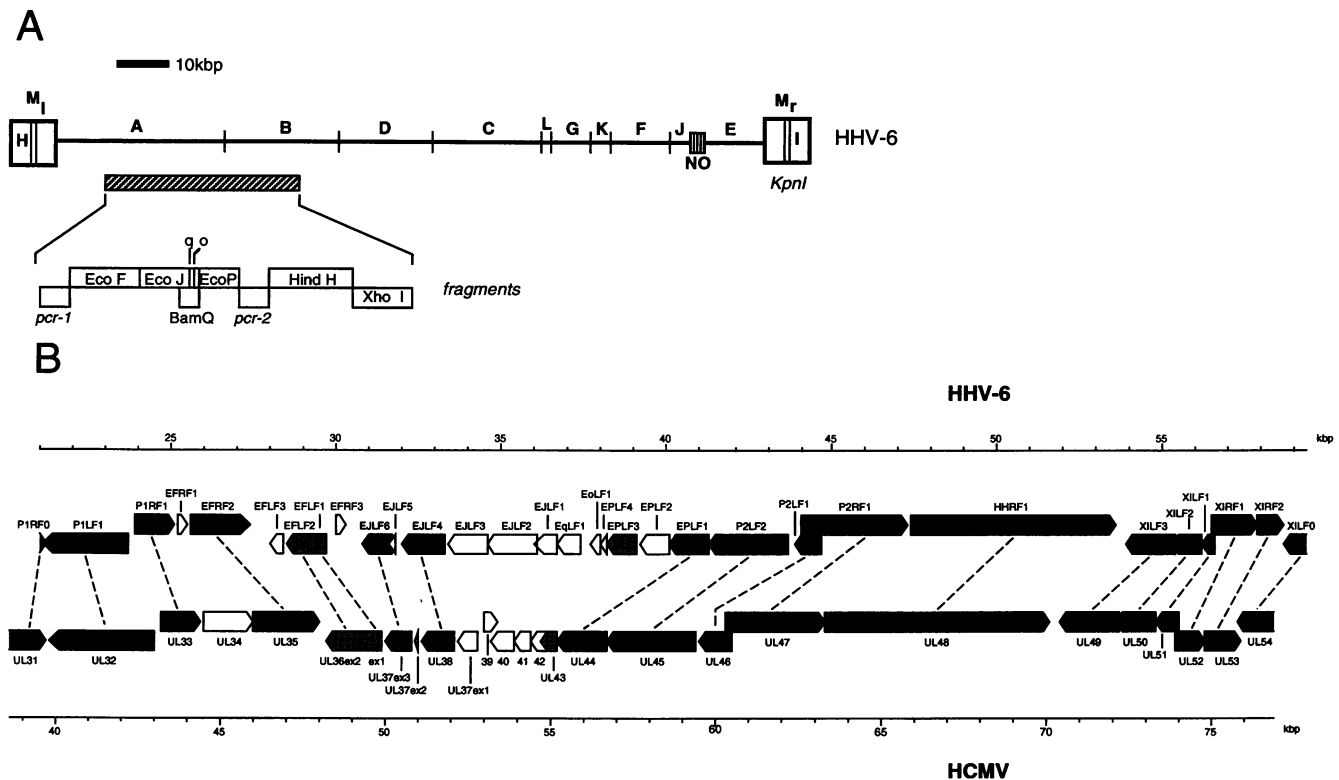


FIG. 1. Genetic organization within the sequenced region of HHV-6 U1102 and the map positions of subgenomic fragments from which the contiguous sequence was derived. (A) Representation of the HHV-6 genome as a *KpnI* restriction map (53), with the locations of the sequenced subgenomic fragments shown below (see Materials and Methods). The small gap (150 bp) between *BamHI*-Q and *EcoRI*-P was sequenced by using cycle sequencing techniques directly on HHV-6 DNA (15). (B) Diagrammatic representation of the results of analyses of the 38,491-bp sequence determined (top section), showing the detected ORFs as arrows (indicating direction of transcription), aligned with the equivalent genetic locus of HCMV (bottom section [11]). Symbols: ■, ORFs homologous between HHV-6 and HCMV; □, members of the homologous US22 family of genes (11); □, genes unique to each virus. Nomenclature of ORFs follows the convention used for EBV ORFs (4) but generally refers to the subgenomic fragments from which the sequence was derived. Thus, EFLF1 indicates the first leftward (L) reading frame (F) initiating in the *EcoRI*-F (EF) fragment of the genome, whereas EFRF1 is ascribed to the first rightward (R) reading frame in *EcoRI*-F. ORFs initiating within *pcr-1*, *pcr-2*, *EcoRI*-o, or *EcoRI*-q (panel A) are named with reference to these fragments.

(0.5 to 1.0 μg), annealing of primer (1 to 10 pmol) at 55°C for 5 min, and then sequencing as described for M13 templates.

**Assembly and analysis of the DNA sequence.** Sequences were assembled by using either the DB (DBAUTO and DBUTIL) or SAP programs of Staden (79), and the edited consensus sequence (38,491 bp from the right *SmaI* site of *SmaI*-G to the right *XhoI* site of *XhoI*-I [53]) was analyzed by using the program packages of ANALYSEQ and NIP (80, 81), Genetics Computer Group (GCG [18]), and Molecular Genetics and Sequencing (MGS; developed by B. Greer and P. Gillette, Computing Laboratory, National Institute for Medical Research, London, United Kingdom). Protein compositional analyses were undertaken by using ANALYSEP and PIP (82) or options in the MGS suite of programs. Data base searches, pairwise and multiple amino acid sequence alignments, and structural analyses of proteins were done by using FASTA (69), GAP (60), PILEUP, and PROTEINSTRUCTURE options in the GCG program package. For amino acid sequence alignments with GAP, gap and length weights were set at 3.0 and 0.1, respectively. Positional base preference and codon usage analyses were undertaken by using the ANALYSEQ (or NIP) programs of Staden (80, 81).

**DNA transfection assays.** Vero cell monolayers were transfected with plasmid DNA essentially as described previously (29). Calcium phosphate-DNA coprecipitates were applied for

10 h, the medium was then changed, and the cells were harvested 48 h later. Cell extracts were made and assayed for CAT activity as described by Gorman et al. (29), and percent acetylation was calculated by scintillation counting of appropriate regions of the chromatography plate.

**Nucleotide sequence accession number.** The 38,491 bp of HHV-6 genomic sequence reported in this article will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under accession number L25528.

## RESULTS

**DNA sequencing and identification of coding sequences.** The contiguous sequence of HHV-6 genomic DNA from positions 21.0 to 59.5kbp was derived from the subgenomic fragments *EcoRI*-F (7.0 kbp), *EcoRI*-J (5.7 kbp), *BamHI*-Q (2.9 kbp), *EcoRI*-P (3.8 kbp), *HindIII*-H (8.7 kbp), and *XhoI*-I (5.9 kbp), from PCR-derived DNA spanning the regions between *SmaI*-G and *EcoRI*-F (*pcr-1*) and between *EcoRI*-P and *XhoI*-I (*pcr-2*), and from data generated by directed sequencing using cycle-sequencing techniques (15) on HHV-6 genomic DNA to link *BamHI*-Q and *EcoRI*-P (53) (Fig. 1A). All sequence data except those for the sequence between *BamHI*-Q and *EcoRI*-P were derived by M13 shotgun cloning and sequencing techniques (5) and assembled by using the

computer programs of Staden (79) to generate contiguous sequences with an average redundancy (contig sequence thickness) of 6.27. Double-stranded sequencing of M13-derived sequence (to resolve ambiguities or check sequences) was done as described in Materials and Methods. Previously published restriction endonuclease maps (53) are in general agreement with the present sequence data, except that a 42-bp *Sall* fragment has now been identified between *Sall*-D and *Sall*-R, there is a 433-bp *Nru*I fragment between *Nru*I-H and *Nru*I-J, there are some minor changes in the *Hind*III map, and the relative positions of the *Eco*RI-o and *Eco*RI-q fragments are reversed.

ORFs were generally considered significant if they exceeded 300 nucleotides, but smaller ORFs were noted when they coincided with potentially analogous ORFs in HCMV or when they displayed features (e.g., by positional base preference analyses [81, 82]) consistent with their being coding. These include ORFs EJLF5, EoLF1, and EPLF4 (Fig. 1B). Additionally, TFASTA searches were performed to try to detect short ORFs in HHV-6 DNA specifying amino acid sequences homologous to HCMV protein sequences. The results of these searches were negative. When there were large intragenic or extensively overlapping ORFs without homologs to HCMV or other protein sequences, the larger and/or higher scoring in positional base preference analyses was identified as coding.

**Identification of ORFs homologous to HCMV genes.** The positions of the ORFs identified within the sequenced region are illustrated relative to their HCMV counterparts in Fig. 1B. Of the 31 HHV-6 ORFs identified (two partial, P1RF0 and XILF0), 20 were found to be homologous to positional counterparts in HCMV. The identities of the homologous HCMV ORFs and equivalent genes in other sequenced herpesviruses, where they occur, are indicated in Table 1 along with the positions of the HHV-6 ORFs within the sequenced region and selected structural features of the encoded proteins. ORFs P1RF0, P1LF1, P1RF1 (ORFs 1, 2/p100, and 3 [61, 62]), EFLF2 (ORF-A [25]), EPLF1 (pCD41 ORF [8, 9]), HHRF1 (37), XILF1 (ORF5 [61]), and XIRF2 and XILF0 (88) have been identified previously as complete or partial sequences in HHV-6 U1102 or GS. P1LF1 and P1RF1 are homologous to HCMV UL32 (pp150 matrix protein [36]) and UL33 (G protein-coupled receptor [12]), respectively, EFLF2 is homologous to HCMV UL36 exon 2 (11, 41), EPLF1 is homologous to HCMV UL44 (ICP36 protein family [45]), HHRF1 is homologous to HCMV UL48 (very large tegument protein [11]), and XILF0 is homologous to HCMV UL54 (DNA polymerase [43]). Other ORFs specifying homologs of characterized HCMV gene products are EFLF1 and EJLF6 (HCMV IE UL36 exon 1 and IE UL37 exon 3, respectively [14, 41]), EJLF4 (HCMV UL38 [41, 87]), and P2LF2 (HCMV UL45, ribonucleotide reductase [67]). In addition to EFLF1 and EFLF2, the HHV-6 sequence presented here encodes another homolog, EPLF3, of the HCMV US22 family of genes (11, 41, 95), distinct homologs of which have been discovered in other regions of the HHV-6 genome (20, 89). The EPLF3 coding sequences occur in a region of the HHV-6 genome positionally equivalent to the location of the UL43 US22 homolog in HCMV, although the sizes of the deduced protein products of these genes are quite different. EPLF3 is most closely related to HCMV UL24. Other HHV-6 homologs to HCMV ORFs of unknown function are P1RF0 (HCMV UL31), EFRF2 (HCMV UL35), P2RF1 (HCMV UL47), XILF3 (HCMV UL49), XILF2 (HCMV UL50), XILF1 (HCMV UL51), XIRF1 (HCMV UL52), and XIRF2 (HCMV UL53) (11).

Of the 20 conserved genes between HHV-6 and HCMV within this locus, 10 are found only in these betaherpesviruses

and not in any other sequenced herpesvirus. These ORFs would thus appear to represent betaherpesvirus-specific genes, at least some of which presumably are important determinants of the specific biological characteristics of this subgroup of herpesviruses. In this respect, the HHV-6 homologs of the HCMV UL36-38 IE genes and the HHV-6 EPLF3 US22 homolog are of particular relevance as potential transcriptional regulators. Also worthy of note are the regions of divergence between these two loci in HHV-6 and HCMV and, in particular, the clustered "unique" genes in each virus comprising ORFs EJLF3 to EPLF4 in HHV-6 and UL37 exon 1 to UL42 in HCMV. Thus, there are seven ORFs in HHV-6 without detectable homologs in HCMV or any other sequenced herpesvirus and five HCMV-specific genes between the EPLF3/UL43 and EJLF4/UL38 collinear regions of the two genomes (Fig. 1B). Similar clustered unique genes in regions of divergence from collinearity have been identified in other closely related herpesviruses, for example EBV and HVS (63). In addition, HHV-6 EFRF1, EFLF3, EFRF3, and EPLF2 and HCMV UL34 represent virus-specific genes dispersed individually within these loci and partly account for the altered spatial arrangements of homologous genes in these regions (Fig. 1B).

**Proteins of predicted function encoded by the 38.5-kbp sequenced region of HHV-6.** In addition to the UL36-38 and US22 homologs in the sequenced region of HHV-6, there are 15 ORFs which are homologous to collinear genes in HCMV (Fig. 1B), 7 of these specifying previously characterized proteins or proteins of predicted function based on data from investigations of HCMV and other herpesviruses. These HHV-6 conserved genes and their homologs in HCMV, HSV-1, VZV, EBV, and HVS are listed in Tables 1 and 2 and are described briefly below. Three nonconserved HHV-6 genes specifying putative membrane-bound glycoprotein species are also described.

(i) **P1LF1 product.** The P1LF1 translation product is homologous to the HCMV UL32-encoded protein, pp150, an immunogenic, phosphorylated matrix protein (36). The corresponding protein of HHV-6, p100, is also a highly immunogenic protein and a component of the virion (97). Its gene has been mapped and sequenced previously (62). The published sequence corresponding to *Eco*RI-c and *Eco*RI-U (62) is identical to ours in all but one position (nucleotide 171). The T in this position of our sequence has been confirmed by directed sequencing of the cloned *Hind*III-S (HD2) subgenomic fragment (53). The HHV-6 p100 and HCMV pp150 proteins are 41.6% similar and 20.9% identical, as measured by GAP alignments (18, 60).

(ii) **P1RF1 product.** The P1RF1-specified protein shows significant amino acid sequence similarity to the HCMV, HVS, and cellular G protein-coupled receptor (GCR) family of proteins (12, 38, 59, 64). The HHV-6 gene was partially sequenced (110 codons) and mapped by Neipel et al. (61). The complete P1LF1 translation product is 870 amino acids long (from the first methionine), specifies a protein of 97 kDa, and displays 52.4% similarity and 25.7% identity to the HCMV GCR protein encoded by UL33, the positional counterpart of P1RF1 and one of three GCR homologs specified by the HCMV genome (12). A GCR homolog has also been found in the gammaherpesvirus HVS (64) but not in any other herpesvirus genome. Despite the relatively low degree of amino acid sequence conservation between these herpesvirus GCR proteins and between these proteins and cellular GCRs, there are several characteristic features that clearly identify this family of proteins, most notably the seven hydrophobic, putative transmembrane domains and the presence of specific conserved

TABLE 1. HHV-6 ORFs within the 38.5-kbp sequence between genomic positions 21.0 and 58.5 kbp and homologous genes in other sequenced herpesviruses

ORF <sup>a</sup>	Sense <sup>b</sup>	Position <sup>b</sup>				Length (aa) <sup>c</sup>	<i>M<sub>r</sub></i> <sup>c</sup>	Homology with HCMV (% similarity/ % identity) <sup>d</sup>	Homolog <sup>e</sup>					Comments	Reference(s)	
		Start	End	First N- terminal Met <sup>b</sup>	Poly(A) <sup>b</sup>				HCMV	HSV-1	VZV	EBV	HVS			
P1RF0	+	Start	292		833*			60.0/34.3	UL31							
P1LF1	-	2814	202	2811	163	870	97,072	41.6/20.9	UL32							
P1RF1	+	3023	4042	3089	4069	318	36,610	52.4/25.7	UL33					74	Virion protein GCR family	36, 62, 97 14, 70
EFRF1	+	4128	4448	4131	5037	106	12,003									
EFRF2	+	4507	6375	4549	6437	609	69,560	45.5/19.8	UL35						UL25 family	11
EFLF3	-	7294	6896	7225	6395	110	13,415									
EFLF2	-	8349	7495		7432	285	32,907	46.0/21.8	UL36×2						IE, US22 family	11, 20, 41, 95
EFLF1	-	8600	8184	8582	7432	133	15,537	48.5/30.3	UL36×1						IE, US22 family	11, 20, 41, 95
EFRF3	+	8958	9245	9057	9733	63	7,573									
EJLF6	-	10634	9744	10622	9648	293	33,218	43.4/18.5	UL37×3						IE, glycoprotein	45
EJLF5	-	10728	10516		9648	71	8,522									
EJLF4	-	12120	10885	12051	9648	389	43,316	44.6/24.7	UL38						IE	41, 87
EJLF3	-	13585	12305	13570	12227	422	48,688								Glycoprotein, signal/TM	
EJLF2	-	15114	13576	14874	12619	433	49,395									
EJLF1	-	15607	14975	15580	14938*	202	23,408								Glycoprotein, signal/anchor	
EqLF1	-	16387	15611	16318	15292	236	26,523								Glycoprotein, signal/TM	
EoLF1	-	16906	16628	16888	16528*	87	10,091									
EPLF4	-	17128	16910	17080	16528*	57	6,330									
EPLF3	-	18065	17100	18047	16528*	316	37,096	49.7/22.5	UL43						US22 family	11, 20, 95
EPLF2	-	19063	18158	19042	18119	295	33,334									
EPLF1	-	20241	19033	20211	18119	393	44,810	65.5/42.2	UL44	(UL42)	(16)	(BMRF1)	(59)		DNA polymerase accessory protein	22, 30, 39, 68
P2LF2	-	22682	20256	22667	20246	804	93,353	48.8/25.2	UL45	UL39	19	BORF2	61	Ribonucleotide reductase	3, 13, 24	
P2LF1	-	23731	22940	23590	22643	217	24,570	51.0/29.4	UL46	UL38	20	BORF1	62	Capsid protein	71, 72	
P2RF1	+	23106	26363	23118	29702	1,082	123,988	48.5/25.4	UL47	(UL37)	(21)	(BOLF1)	(63)			
HHRF1	+	26381	32614	26384	32672	2,077	239,945	47.5/23.9	UL48	UL36	22	BPLF1	64	Very large tegument protein	26, 28, 55	
XILF3	-	34384	32960	34369	32657	470	54,700	61.4/39.1	UL49			BFRF2	66			
XILF2	-	35159	34323	35150	34234	276	31,671	61.3/38.7	UL50	(UL34)	(24)	BFRF1	67			
XILF1	-	35556	35170	35487	34529	106	12,440	65.4/37.4	UL51							
XIRF1	+	35402	36937	35486	37520*	484	56,407	63.1/39.2	UL52	UL32	26	BFLF1	68			
XIRF2	+	36941	37735	36944	37734*	264	30,845	66.3/42.4	UL53	UL31	27	BFLF2	69			
XILF0	-	End	37787		37749			75.8/52.1	UL54	UL30	28	BALF5	9	DNA polymerase	42, 43, 88, 96	

<sup>a</sup> The nomenclature used for the HHV-6 ORFs is explained in the legend to Fig. 1.

<sup>b</sup> The positions and orientations (relative to the conventional orientation of HHV-6 [53]) of the ORFs are indicated, together with the locations of the first N-terminal ATG codons (where these occur) and first downstream polyadenylation signals (AATAAA, ATTTAA\*), within the contiguous sequence described here.

<sup>c</sup> The sizes of the ORF translation products (starting at N-terminal methionines) in amino acids (aa) and their relative molecular weights (*M<sub>r</sub>*) are shown.

<sup>d</sup> Homologous genes were identified by FASTA searches and GAP alignments (60) with the GCG program package (18). The values for percent similarity and percent identity between the HHV-6 and HCMV homologs are based on GAP alignments with gap and length weights set at 3.0 and 0.1, respectively.

<sup>e</sup> Listings of homologous genes were based on data from comparisons of other herpesvirus proteins (see, e.g., references 11, 55, and 63). Viral genes positionally (and in some cases also functionally) analogous to HHV-6 genes but with very limited sequence similarity are shown in parentheses.

amino acid residues within and around these domains (12, 38, 59, 64) (Fig. 2).

(iii) **EPLF1 product.** The HHV-6 GS equivalent of the EPLF1 gene of HHV-6 U1102 has been sequenced as a cDNA clone, pCD41 (8, 9). This cDNA clone contains sequences which in HHV-6 U1102 encode both EPLF1- (393 amino acids) and EPLF2 (295 amino acids)-specified proteins. The 3' end of the mRNA occurs at position 14655 of our sequence, 14 nucleotides downstream of the only polyadenylation signal (AATAAA) within the sequences represented in pCD41. Thus, EPLF1 and EPLF2 appear to be encoded by a bicistronic mRNA. Differences between our nucleotide sequence and that

presented by Chang and Balachandran (8, 9) result in eight amino acid differences between overlapping portions of the EPLF1 and GS p41 translation products, two amino acid differences between the EPLF2 protein and its GS equivalent, and size differences between the deduced U1102 and GS proteins.

The HHV-6 EPLF1 (p41) protein is homologous to the phosphorylated and glycosylated family of proteins, ICP36 (45), encoded by HCMV UL44 (11). Both p41 and the UL44 product are nuclear proteins, and the latter has shown to bind DNA (8, 45) and to be required for HCMV *ori*Lyt-dependent DNA replication (22, 68). The UL44 and HHV-6 EPLF1 gene

TABLE 2. Comparisons of HHV-6 ORF translation products with homologs in HCMV, HSV-1, VZV, EBV, and HVS

HHV-6 ORF	% Similarity/% identity <sup>a</sup> with:				
	HCMV	HSV-1	VZV	EBV	HVS
PIRF1	52.4/25.7				47.1/21.2
EPLF1	65.5/42.2	40.5/16.4	38.1/16.0	42.1/18.1	38.0/17.3
P2LF2	48.8/25.1	47.6/22.6	47.5/23.6	48.1/22.9	46.7/22.8
P2LF1	51.0/29.4	41.7/20.4	43.4/20.0	43.2/16.9	44.4/16.4
HHRF1	47.5/23.9	42.1/19.0	42.7/19.8	40.7/20.2	41.7/20.7
XILF3	61.4/39.1			44.9/23.7	49.1/27.0
XILF2	61.3/38.7	43.6/19.6	44.9/15.2	46.3/21.7	42.8/17.5
XIRF1	63.1/39.2	48.6/23.2	50.0/26.1	49.0/26.7	50.6/27.3
XIRF2	66.3/42.4	48.7/21.1	50.2/23.0	49.0/20.9	44.6/21.5
XILF0	75.8/52.1	55.7/32.6	58.1/32.2	56.5/34.5	54.1/36.4

<sup>a</sup> Values for percent similarity (first number) and percent identity (second number) were obtained from pairwise alignments of HHV-6 translation products with homologs in the other herpesviruses by using GAP (60) in the GCG package (18) with gap and length weight settings of 3.0 and 0.1, respectively. The values illustrate the closer relationship between HHV-6 and HCMV, a betaherpesvirus, than to the alphaherpesviruses (HSV-1 and VZV) or gammaherpesviruses (EBV and HVS).

products are likely, therefore, to be functionally analogous to their positional counterparts in HSV-1 (UL42) and EBV (BMRF1), which show DNA-binding properties, associate with viral DNA polymerase, and are essential for lytic cycle viral DNA replication (23, 30, 39, 51, 93). Regulation of HCMV UL44 gene expression is manifested in both early and late transcription initiating downstream of three different TATA boxes and in delayed translation of early mRNA (45). Since there are three TATA box motifs (TATAA) in EPLF1 upstream sequences (150 bp), it is possible that EPLF1 regulation is analogous to that of HCMV UL44.

(iv) **P2LF2 product.** The P2LF2 translation product is homologous to the UL45-encoded protein of HCMV and the UL39-, gene 19-, BORF2-, and gene 61-specified products of HSV-1, VZV, EBV, and HVS, respectively (2, 4, 11, 17, 56, 63). These proteins are structurally and functionally homologous to the large subunit of ribonucleotide reductase found in other organisms. In common with HCMV, a betaherpesvirus, and in contrast to representatives of the alphaherpesviruses (HSV-1 and VZV) and gammaherpesviruses (EBV and HVS), HHV-6 appears to lack an ORF specifying the small ( $\beta$ ) subunit of the enzyme. In the alpha- and gammaherpesviruses this ORF is located immediately downstream of the  $\alpha$ -subunit gene. In HSV-1, at least, the  $\alpha$ - and  $\beta$ -polypeptides associated to form  $\alpha_2\beta_2$  heterodimers as for mammalian ribonucleotide reductases (3, 13, 24). HCMV and HHV-6 are therefore distinct from the alpha- and gammaherpesviruses in their possession of only one ribonucleotide reductase ORF, and it is notable that these betaherpesvirus proteins are more closely related to each other than to homologs in the other sequenced herpesviruses (Table 2).

(v) **P2LF1 product.** P2LF1 specifies a homolog of the UL46-, UL38-, RF20-, BORF1-, and RF62-encoded proteins of HCMV, HSV-1, VZV, EBV, and HVS, respectively. The UL38 gene product in HSV-1 has been shown to be a component of the capsid and to be important for capsid assembly (71, 72).

(vi) **HHRF1 product.** The HHRF1 translation product specifies the HHV-6 homolog of the UL48-, UL36-, RF22-, BPLF1-, and RF64-encoded proteins of HCMV, HSV, VZV, EBV, and HVS, respectively. These proteins represent the largest (>200 kDa) of the herpesvirus proteins and are structural components of the tegument (4, 11, 17, 26, 28, 55, 63).

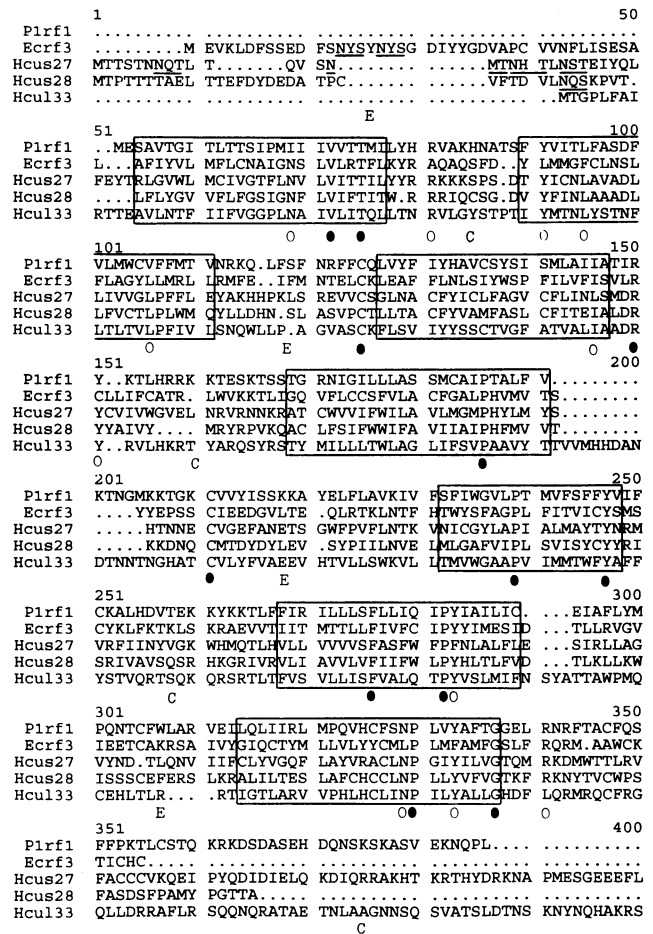


FIG. 2. Alignment of the HHV-6 PIRF1 translation product with the other detected herpesvirus GCR homologs in HVS (Ecrf3, gene 74 [2, 64]) and HCMV (US27, US28, and UL33 [12]). The alignments were based on those performed by Chee et al. (12) and on a multiple alignment output obtained by using the PILEUP option in the GCG program package (18). The sequences were aligned manually by using the LINEUP and PRETTY programs in the GCG package. The boxed regions indicate putative transmembrane domains, and the predicted cytoplasmic (C) and extracellular (E) domains are indicated below the sequence. Potential N-linked glycosylation sites occurring in the N-terminal regions of the proteins are underlined. Amino acids conserved in all five protein sequences (●) or in all but one (○) are indicated below the sequences.

The HHV-6 GS equivalent of HHRF1 has been sequenced (37).

(vii) **XILF0 product.** The partial ORF XILF0 represents part of the HHV-6 U1102 DNA polymerase gene, which has been sequenced and described previously (88).

(viii) **EJLF3, EJLF1, and EqLF1 products.** The translation products of these HHV-6 ORFs do not have detectable homologs among the HCMV-specified proteins, other herpesvirus proteins, or protein sequences in the data base. However, it seems highly probable from computer-based structural analyses (PIP [82]; GCG PEPTIDESTRUCTURE [18]) that each represents a membrane-bound glycoprotein. The N-terminal sequences of the EJLF3, EJLF1, and EqLF1 proteins are predicted to function as signal peptide sequences according to the criteria of McGeoch (54). In addition, all three proteins

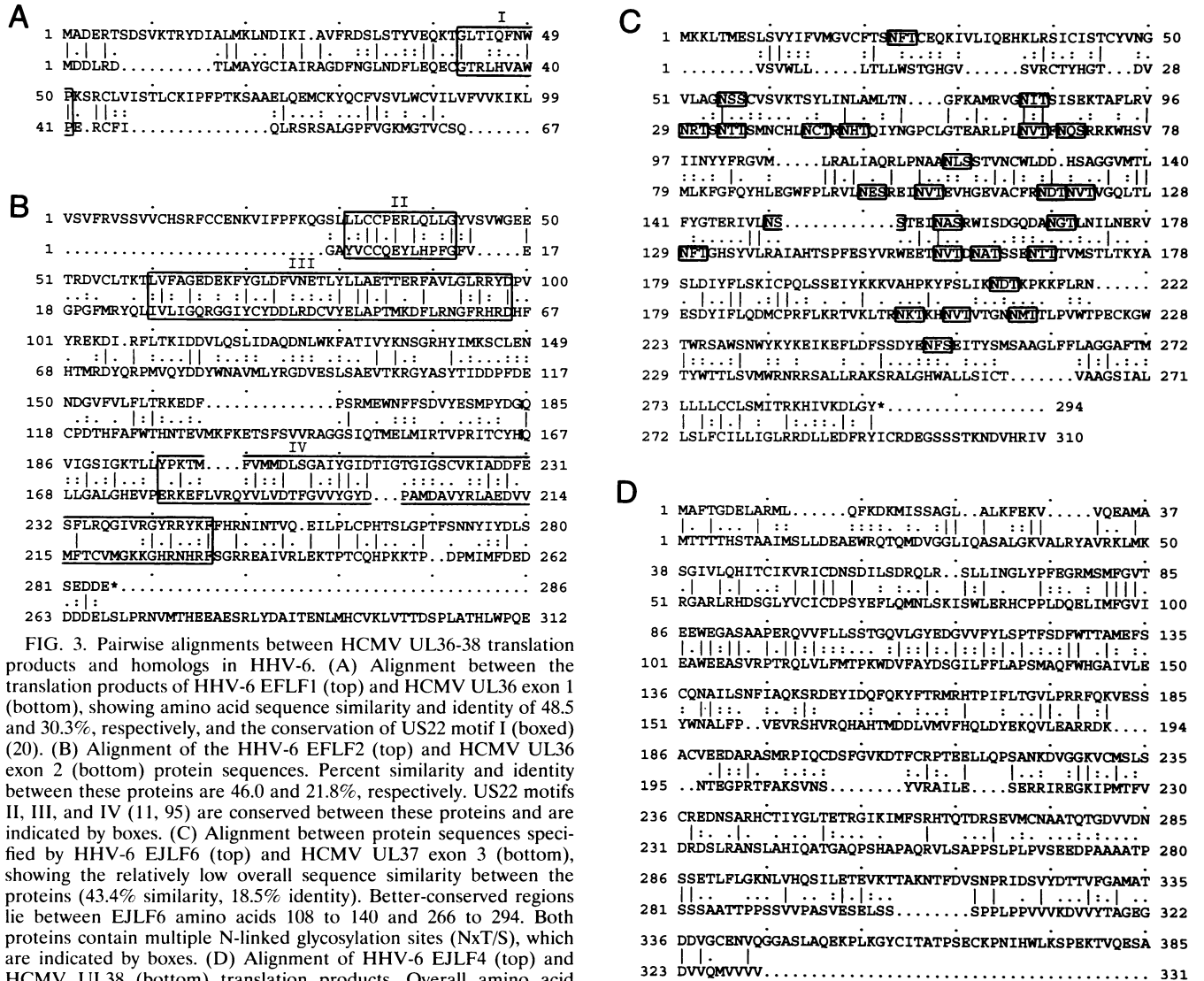


FIG. 3. Pairwise alignments between HCMV UL36-38 translation products and homologs in HHV-6. (A) Alignment between the translation products of HHV-6 EFLF1 (top) and HCMV UL36 exon 1 (bottom), showing amino acid sequence similarity and identity of 48.5 and 30.3%, respectively, and the conservation of US22 motif I (boxed) (20). (B) Alignment of the HHV-6 EFLF2 (top) and HCMV UL36 exon 2 (bottom) protein sequences. Percent similarity and identity between these proteins are 46.0 and 21.8%, respectively. US22 motifs II, III, and IV (11, 95) are conserved between these proteins and are indicated by boxes. (C) Alignment between protein sequences specified by HHV-6 EJLF6 (top) and HCMV UL37 exon 3 (bottom), showing the relatively low overall sequence similarity between the proteins (43.4% similarity, 18.5% identity). Better-conserved regions lie between EJLF6 amino acids 108 to 140 and 266 to 294. Both proteins contain multiple N-linked glycosylation sites (NxT/S), which are indicated by boxes. (D) Alignment of HHV-6 EJLF4 (top) and HCMV UL38 (bottom) translation products. Overall amino acid sequence similarity and identity between these sequences are 44.6 and 24.7%, respectively. Alignments A to D were performed by using the GAP option in the GCG package (18, 60) with gap and length weights set at 3.0 and 0.1, respectively.

possess markedly hydrophobic, putative transmembrane domains, and all contain multiple N-linked glycosylation sites (NxT/S). The EJLF3, EJLF1, and EQLF1 translation products contain 10, 4, and 3 potential glycosylation sites, respectively. In EJLF1, the transmembrane domain occurs at the C terminus of the protein and may represent an anchor sequence. Thus, all three of these HHV-6 proteins are likely to represent multiply glycosylated membrane proteins, distinct from other herpesvirus glycoprotein species.

**Homologs of the HCMV IE UL36-38 genes specified by HHV-6.** The most abundant IE transcripts of HCMV arise from the major IE locus (85), but there are additional IE transcripts that are specified by other regions of the genome. These map to positions of reading frames UL36-38, US3, and IRS1/TRS1 (11, 41, 83, 87, 94). The UL36-38 ORFs make up an IE transcription unit from which four mRNA transcripts arise from three distinct promoters (41, 87) (see Fig. 4A). The

1.7-kb (UL37 exon 1, UL38), 3.2-kb (UL37 exons 1, 2, and 3), and 1.68-kb (UL36 exons 1 and 2) mRNA species are all expressed as IE transcripts (41), whereas the 1.15-kb mRNA (UL38) is expressed with early kinetics (87). Products specified by individual UL36-38 cDNA clones or genomic sequences have recently been reported to act synergistically with other IE gene products (those of the US3 and IE2 genes) to activate heterologous or HCMV early gene expression, and the entire UL36-38 region can transactivate at least one heterologous promoter independently (14).

In HHV-6, we have identified homologs of the UL36 exon 1, UL36 exon 2, UL37 exon 3, and UL38 ORFs of HCMV. Alignments of the HHV-6 EFLF2, EFLF1, EJLF6, and EJLF4 translation products with their homologs in HCMV are shown in Fig. 3. It is important to note that although the overall amino acid sequence similarities of these paired proteins are relatively low, there are restricted regions of the proteins that show far better conservation of amino acid sequence or that display other conserved features. For example, conservation of amino acid sequence between EJLF4 and HCMV UL38 is restricted mainly to the amino-terminal portions of the proteins (Fig.

3D). The homology between the US22 family members EFLF1 and UL36 exon 1, and EFLF2 and UL36 exon 2, is seen clearly by the possession of collinear US22 motifs (11, 20, 41). Motif I (GxxOxOxWP, where O is a hydrophobic residue and x is any residue) is present in the short EFLF1 and UL36 exon 1 translation products (Fig. 3A), whereas motifs II, III, and IV are present in the products of the downstream ORFs EFLF2 and UL36 exon 2. Motif II (OOCxLxxOG) is most clearly defined, with the larger hydrophobic domains making up motifs III and IV distal to this (Fig. 3B). In HCMV, UL36 exons 1 and 2 are spliced together in an mRNA specifying the UL36 coding sequences, thus joining all four US22 motifs in the translation product. Similar splicing events are also possible (and indeed predicted) for the HHV-6 EFLF1 and EFLF2 ORFs (see Fig. 4B). The HHV-6 EJLF6 and HCMV UL37 exon 3 ORFs show limited overall amino acid sequence similarity. However, there are restricted regions of the proteins (e.g., EJLF6 positions 108 to 140 and 266 to 294 and equivalent regions in UL37 exon 3) that are better conserved and contain conserved cysteine residues (Fig. 3C). Furthermore, in addition to their genetic collinearity, they share other notable features. First, they are each predicted to encode highly glycosylated proteins, the EJLF6 and UL37 exon 3 ORFs specifying 9 and 17 N-linked glycosylation sites, respectively. Second, each translation product specifies at its C terminus a hydrophobic domain, possibly representing an anchor sequence. Third, the EJLF6 and UL37 exon 3 translation products are of very similar size, 294 and 310 amino acids, respectively. Homologs of the first two exons of UL37 were not detected in HHV-6, despite the use of TFASTA searches (18) to try to detect the short (14-amino-acid) UL37 exon 2 equivalent in the HHV-6 sequences.

The EJLF3 ORF encodes a potentially highly glycosylated translation product (10 glycosylation sites, NxT/S, are present). It is not detectably homologous to UL37 exon 1, its positional counterpart in HCMV, but it is possible that it is functionally analogous. In common with HCMV UL37 exon 1, HHV-6 EJLF3 encodes a hydrophobic, putative signal peptide sequence at its N terminus (41, 54). Although EJLF3 may be expressed as an unspliced mRNA, splicing events analogous to those in HCMV may occur. Figure 4 indicates the positions of potential splice donor and splice acceptor sites in relation to the HHV-6 EFLF2-EJLF3 ORFs, demonstrating the possibility of analogous splicing patterns to those determined for the HCMV UL36-38 mRNAs, including the 3.2-kb spliced UL37 mRNA. The putative equivalent of UL37 exon 2 indicated in Fig. 4B represents a 13-amino-acid coding segment of the EJLF5 ORF. A notable divergence between these betaherpesvirus loci is the spacing between the UL37 exon 3 and UL36 exon 1 ORFs in HCMV and the EJLF6 and EFLF1 homologous ORFs in HHV-6 and the presence of additional potential coding sequences, EFRF3, in HHV-6.

**HHV-6 EPLF3 encodes a member of the HCMV US22 family of proteins.** The EPLF3 ORF specifies a translation product of 316 amino acids (from the first methionine) which, like the EFLF1/EFLF2-encoded protein (see above), is homologous to the 12-member US22 family of proteins in HCMV (11, 41, 95). Other US22 homologs have recently been identified in HHV-6 (20, 25, 89). US22 homologs have so far not been identified in any other herpesvirus, and it is likely that they represent betaherpesvirus-specific genes. The predicted translation products of the US22 genes vary considerably in length (143 to 788 amino acids) and have a low overall amino acid sequence similarity. The US22 proteins contain one or more of four sequence motifs (11, 20), all of which are present in the EPLF3 translation product. These are illustrated, along

with representative motifs from other US22 family members in HHV-6 and HCMV, in Fig. 5.

**Functional analyses of the EFLF2-EJLF4 and EPLF3 genes.** Products encoded by the HCMV IE UL36-38 transcription unit have recently been shown to function as transcriptional regulators of heterologous and HCMV promoters in transient cotransfection experiments (14). The duplicated US22 family members TRS1 and IRS1 of HCMV have also been shown to encode transactivating functions (83), thus demonstrating regulatory properties encoded by US22 genes other than UL36. In view of these previous findings, we wanted to test the ability of our UL36-38 and US22 homologs to regulate gene expression. In initial experiments to assess the likely role of the EPLF3 ORF, the plasmid pED4 (pEcoRI-P [53]), which contains intact ORFs EPLF1, EPLF2, and EPLF3, was cotransfected in increasing amounts (2, 5, and 10  $\mu$ g) with pLTR-CAT (1  $\mu$ g) into Vero cells. The results of these assays demonstrated that a function(s) encoded by pED4 was able to induce LTR-CAT expression up to 12.5-fold (Fig. 6B). To investigate the role of the US22 homolog EPLF3 in this induction, we also tested the effect of pEPLF3 in dose-response experiments with pLTR-CAT. The results are shown in Fig. 6C and demonstrate that these sequences are able to effect induction of LTR-CAT expression (4.4-fold at 500 ng of pEPLF3), indicating that EPLF3 does indeed encode the ED4 transactivating function. Expression constructs containing EPLF1 and EPLF2 coding sequences were nonfunctional in similar assays (data not shown).

Similar assays were performed with the effector constructions pIEGP1-2 and pIEGP2-3' (Fig. 6A) to assess the potential roles of the HHV-6 UL38/UL37 exon 3 and UL36 homologs to effect transactivation. In addition to the UL36 homologues, pIEGP2-3' contains EFRF3 coding sequences but lacks the 3' polyadenylation and processing signals necessary for expression of this ORF. Figure 6D summarizes the results of cotransfection experiments in which pLTR-CAT (1  $\mu$ g) was cotransfected with increasing doses (1 to 15  $\mu$ g) of each effector construct both individually and together. The maximum levels of induction obtained were approximately fivefold for both pIEGP1-2 and pIEGP2-3' (at 1 and 2.5  $\mu$ g, respectively) and were not increased significantly in mixed transfections, indicating that the gene products encoded by these effector constructions do not function synergistically, at least in these assays.

Our data indicate that the HHV-6 homologs of the HCMV IE UL36-38 genes are able to effect transactivation of at least one target promoter, the HIV-1 LTR, in transient-cotransfection assays, and that the EFLF1/EFLF2 (UL36 exon 1/UL36 exon 2) and EJLF6 to EJLF4 (UL37 exon 3 to UL38) regions each encode a transactivating function(s). Furthermore, a new HHV-6 US22 family member, EPLF3 (HCMV UL43 positional counterpart), also encodes a transactivating function consistent with the emerging evidence implicating the US22 gene products as IE transcriptional activators (14, 25, 83).

## DISCUSSION

The data presented in this report comprise the primary sequence of a contiguous 38.5-kbp region of the genome of HHV-6 strain U1102, its analysis, the identification of several genes conserved between HHV-6 and HCMV, and the demonstration of transactivating functions encoded by DNA sequences specifying ORFs homologous to the HCMV IE UL36-38 and UL43 genes. The sequenced region of HHV-6 is between genomic positions 21.0 and 59.5 kbp (53) and is genetically generally collinear with the positionally equivalent



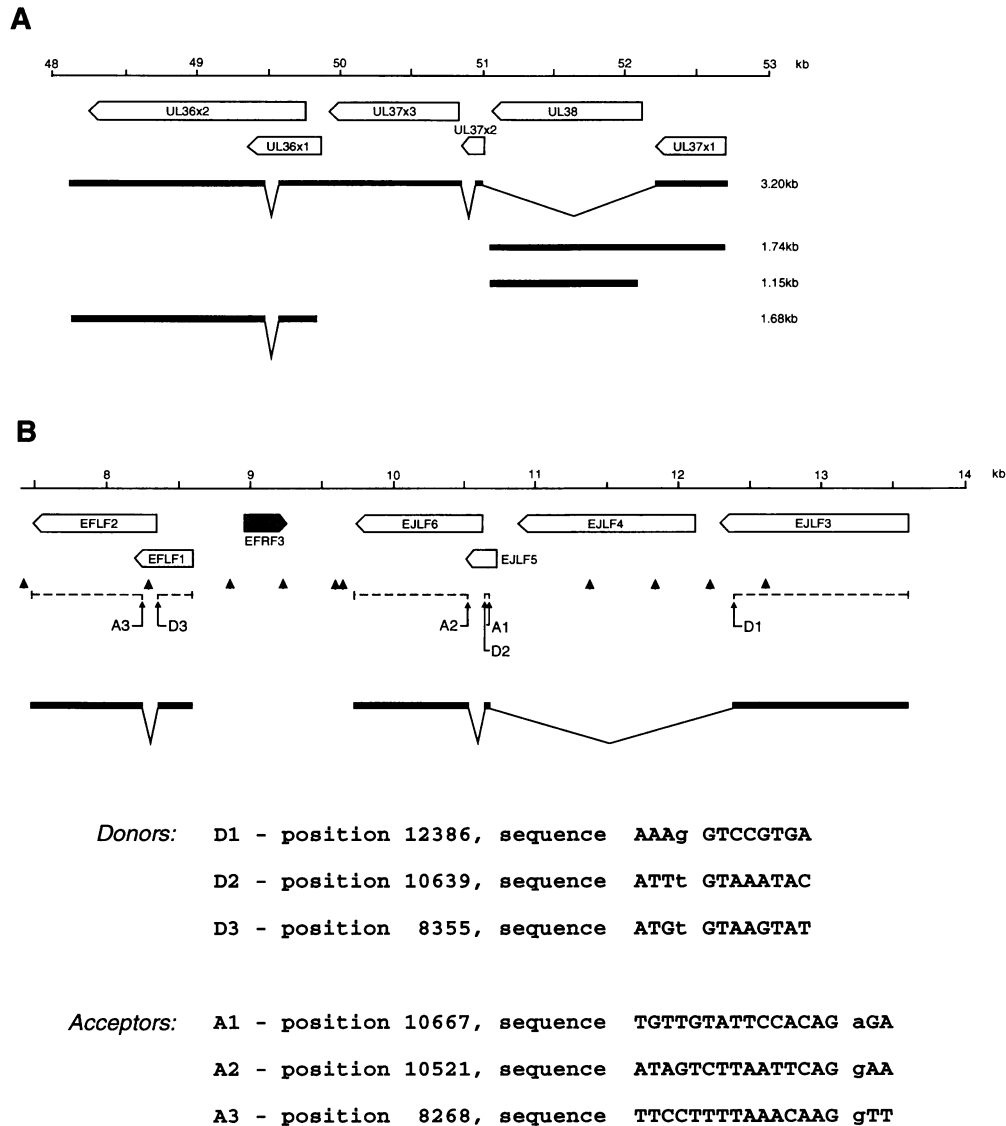


FIG. 4. Splicing patterns within the HCMV UL36-38 locus and the possibility of analogous splicing events in the homologous region of HHV-6. (A) HCMV ORFs UL36-38 are indicated together with the mRNA species encoding the corresponding gene products. Thick lines represent exons, and their splicing is depicted by the adjoining lines. These data were derived from Kouzarides et al. (41) and Tenney and Colberg-Poley (87). The positions of the ORFs within the HCMV genomic sequence are indicated by the scale (top). (B) HHV-6 homologous genetic locus, showing the conserved ORFs EFLF2 (UL36 exon 2), EFLF1 (UL36 exon 1), EJLF6 (UL37 exon 3), and EJLF4 (UL38), along with ORFs potentially analogous to HCMV UL37 exon 2 (EJLF5) and UL37 exon 1 (EJLF3), and the nonconserved EFRF3 ORF (shaded). Numbering on the scale (top) refers to the sequence described in this paper. The locations of potential splice donor (D) and splice acceptor (A) sites are indicated by arrows, and details of their sequence and coordinates are listed below. All polyadenylation signals (AATAAA) in the negative strand are indicated by arrowheads. Putative mRNA species derived from splicing between the indicated splice donor-acceptor sites are indicated by thick lines (exons) and adjoining lines representing splicing. Such mRNA structures for HHV-6 are purely speculative.

region within the  $U_L$  component of the genome of HCMV (11). Within the sequenced region of HHV-6, we have detected homologs of the HCMV ORFs UL31, UL32 (phosphorylated matrix protein), UL33 (GCR homolog), UL35, UL36 exons 1 and 2 (IE, US22 family), UL37 exon 3 (IE glycoprotein), UL38 (IE), UL43 (US22 family), UL44 (DNA polymerase accessory protein), UL45 (ribonucleotide reductase), UL46 (capsid protein), UL47, UL48 (very large tegument protein), UL49, UL50, UL51, UL53, and UL54 (DNA polymerase). These HHV-6 and HCMV conserved genes lie in equivalent relative positions and orientations within the ge-

nomes of these betaherpesviruses. Our data therefore confirm and extend previous reports indicating general collinearity between HHV-6 and HCMV  $U_L$  (20, 44, 52, 61).

The results of sequence analyses also identify 11 genes of HHV-6 with no detectable homology to HCMV or other herpesvirus genes or to any other sequences in the data base. These HHV-6-specific genes occur mainly as a cluster of ORFs between the UL36-38 and UL43 homologs (HHV-6 EJLF4-EPLF3, positions 33 to 38 kbp), with others singly dispersed among HHV-6/HCMV conserved genes (Fig. 1B). Thus, the regions within the genomes of HHV-6 and HCMV encoding

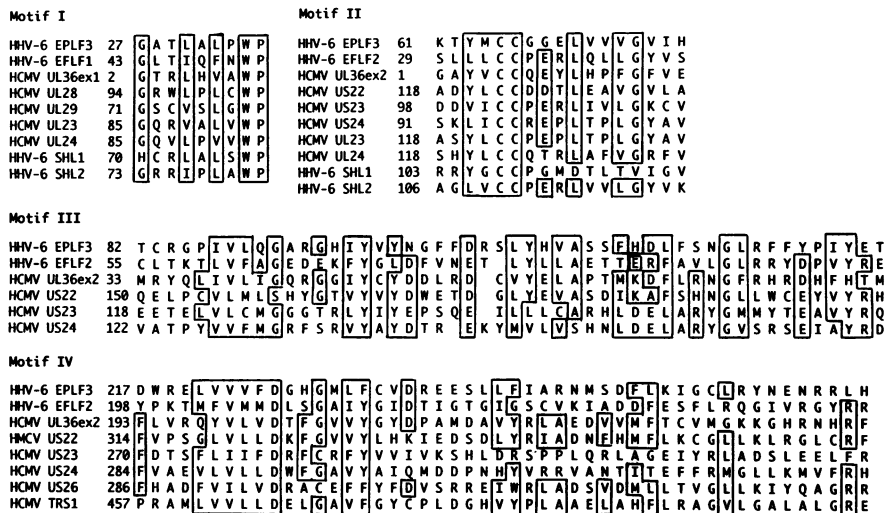


FIG. 5. The HHV-6 EPLF3 translation product contains all four motifs characteristic of the US22 protein family. The sequences of the EPLF3-encoded US22 motifs (20, 41, 95) are shown aligned with corresponding motifs in other US22 family members in HHV-6 (20) and HCMV (11). Numbers indicate amino acid positions within the proteins (initiating at the first methionine) of the first amino acid of each sequence presented. US22 motifs I, II, III, and IV are shown separately, and their characteristics are described in the text.

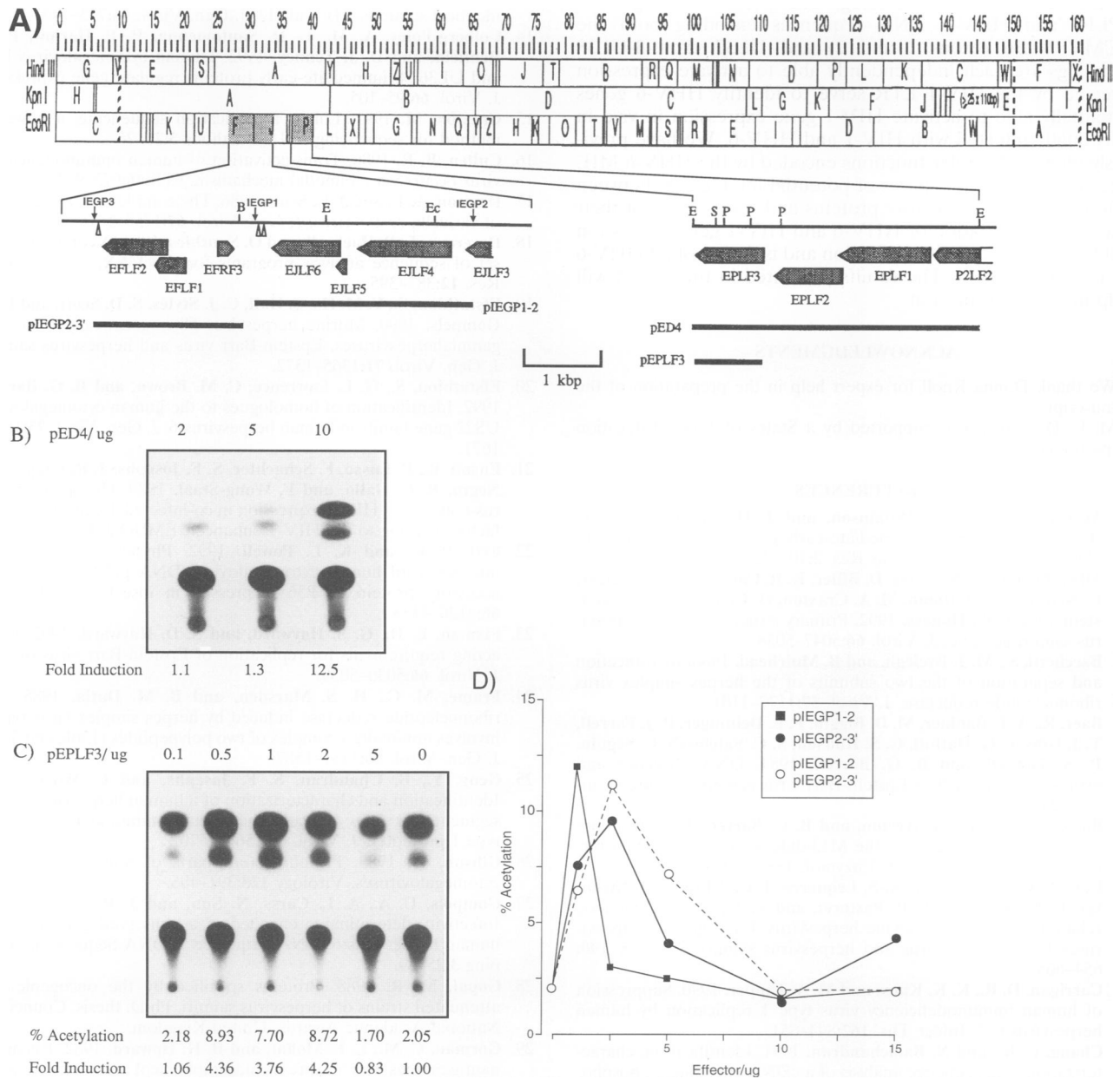
the betaherpesvirus-specific IE UL36-38-type and UL43 (US22)-type genes are loci in which HHV-6 and HCMV “unique” genes are also located.

The HHV-6 and HCMV unique genes within the genetic loci presented here may therefore represent genes whose products are important contributors to the specific biological characteristics of each virus. In contrast, the conserved betaherpesvirus-specific genes represented by HCMV UL36-38 and UL43 and HHV-6 EPLF2-EJLF4 and EPLF3 (Fig. 1B) within these loci are likely to specify functions relevant to biological properties and requirements of all members of the subgroup. The clustering of herpesvirus unique genes within otherwise collinear regions between herpesviruses of the same subgroup has been noted previously for HVS and EBV (2, 63). Of relevance to this discussion is the finding that the two major regions of divergence between these gammaherpesviruses occur at positions which in EBV encode genes and *cis* elements important for episomal viral replication in persistently infected cells. HVS lacks equivalents of all of the EBV latency genes, *oriP* (EBV latent replication origin), and the BZLF1 gene (encoding Zta) that regulates latent-lytic switching. This divergence between HVS and EBV indicates fundamental differences in the mechanisms of maintenance of and reactivation from latency in these two gammaherpesviruses, possibly reflecting the distinct latent tropisms of HVS (T cell) and EBV (B cell). Within these regions of divergence in HVS are found HVS-specific genes, including cyclin and GCR homologs, which could conceivably play important roles either during latency or during lytic cycle replication (63, 64). Thus, there is a precedent for the presence of clustered herpesvirus-specific genes at regions of divergence from collinearity between viral genomes of the same herpesvirus subgroup and evidence that such unique genes can be major determinants of differences in biological properties between these viruses. The HHV-6 EJLF3-EPLF4 and HCMV UL39-UL42 genes may represent such biologically important genes.

Of particular significance is the finding that HHV-6 encodes homologs of the HCMV UL36-38 genes and also contains a positional and structural homolog of the HCMV UL43 gene, a member of the US22 gene family. In HHV-6 and HCMV, the

US22 genes occur in multiple copies (11, 20, 89). An HHV-6 US22 homolog positionally equivalent to HCMV UL43 has not previously been reported, and the EFLF2-EJLF4 genes of HHV-6 described here represent the first homologs of the HCMV UL36-38 IE genes to be detected. Therefore, in addition to the betaherpesvirus major IE gene loci, which contain conserved coding and noncoding features (11, 32, 52, 58, 75, 84), it is likely that the HCMV UL36-38 genes represent betaherpesvirus-conserved IE genes specific to this subgroup of herpesviruses. Analogous subgroup-specific IE genes in the alphaherpesviruses are equivalents of HSV-1 IE68, IE110, and IE175 (55) and in the gammaherpesviruses are equivalents of EBV BRLF1 (4, 65, 92). These subgroup-specific IE genes are likely to specify important functions with respect to biological characteristics and requirements common to all members within each herpesvirus subgroup. The virus-specific IE genes, on the other hand, probably encode transregulatory proteins specifically suited to function within specific host cells and/or within the context of the particular viral genome. These virus-specific IE genes include EBV BZLF1 (Zta [4, 10, 46, 74]), HVS IE-G (66), and bovine herpesvirus 4 IE1 (91). The finding of HCMV IE UL36-38 homologs in HHV-6 suggests that these genes correspond to betaherpesvirus-conserved IE regulatory genes. However, conclusive evidence of this will have to await appropriate sequence, transcriptional, and functional data from other betaherpesviruses.

Finally, in this report we have presented data from cotransfection assays that provide evidence of HHV-6 EPLF3- (US22 homolog), EFLF1/EFLF2- (UL36 homologs), and EJLF6/EJLF4 (UL37 exon-3/UL38 homologs)-specified transactivating functions (Fig. 6). Thus, the effector plasmids pEPLF3, pIEGP1-2 and pIEGP2-3' were each able to transactivate HIV-1 LTR-directed CAT expression during cotransfections into Vero cells. These results are consistent with the notion that the HHV-6 homologs of the HCMV IE UL36-38 genes are themselves expressed with IE kinetics and function as transactivators of HHV-6 lytic cycle gene expression. The same may also be true of EPLF3. The data obtained with pIEGP2-3' are also in agreement with previous results of Geng et al. (25), who identified a transactivating function encoded by the GS



**FIG. 6.** Transactivating functions specified by HHV-6 DNA sequences encoding homologs of HCMV UL36-38 and US22 genes. (A) Effector gene constructs containing HHV-6 homologs of HCMV UL36-38 and UL43 (US22) genes. The HHV-6 ORFs within the relevant region of the genome are shown as large arrows, together with the positions of restriction endonuclease cleavage sites (B, *Bam*HI; E, *Eco*RI; Ec, *Eco*RV; P, *Pst*I; S, *Sal*I), PCR primers (IEGP1, IEGP2, IEGP3), and polyadenylation signals (AATAAA [arrowheads]). Sequences amplified by PCR and cloned into appropriate sites within the polylinker of pUC19 or into pBC12 (see Materials and Methods) are indicated below the ORFs (solid bars). Plasmids pIEGP1-2 and pIEGP2-3' contain EJLF6, EJLF5, and EJLF4 ORFs and EFLF1, EFLF2, and EFRF3 ORFs, respectively. 3' polyadenylation and processing signals for EFRF3 are not present in pIEGP2-3'. (B) Induction of CAT activity by pED4 in cotransfections with 1 µg of pLTR-CAT and 2, 5, or 10 µg of pED4. Fold inductions were calculated from the results of cotransfecting pUC18 (10 µg) with LTR-CAT (not shown). The total amount of DNA in each transfection was made up to 11 µg with pUC18. (C) Transactivation of pLTR-CAT by pEPLF3. CAT activities induced in cotransfections of pLTR-CAT (1 µg) with increasing amounts (0, 0.1, 0.5, 1.0, 2.0, and 5.0 µg) of pEPLF3 are indicated above each lane. Percent acetylations and calculated fold inductions (relative to pUC18 control [last lane]) are indicated below each lane. Total amounts of DNA used in each transfection were made up to 16 µg with pUC18, and Vero cells were transfected as described in Materials and Methods. (D) Activation of pLTR-CAT expression by pIEGP1-2 and pIEGP2-3'. CAT activity (percent acetylation of chloramphenicol) induced in Vero cells cotransfected with pLTR-CAT (1 µg) and different amounts (0, 1, 2.5, 5, 10, and 15 µg) of pIEGP1-2, pIEGP2-3', or pIEGP1-2 and pIEGP2-3'. Total amounts of DNA were made up to 31 µg with pUC18.

equivalent of EFLF1/2, although in the absence of extensive flanking sequence data, the regulatory gene was not recognized as the HCMV UL36 homolog. The present findings that EPLF3 and HHV-6 DNA sequences encoding both the HCMV UL36 (exons 1 and 2) and UL37 exon 3-UL38 homologs are each independently able to activate expression directed by the HIV-1 LTR serve to identify HHV-6 genes potentially able to activate HIV-1 gene expression *in vivo* in cells dually infected with HIV-1 and HHV-6. We have previously identified similar functions encoded by the HHV-6 MIE locus (52). Detailed analysis of potential interactions between different HHV-6 regulatory proteins and assessments of their roles in the regulation of HHV-6 and HIV-1 gene expression will have to await the identification and isolation of all HHV-6 IE regulatory genes. The results presented in this report will help in realizing this aim.

#### ACKNOWLEDGMENTS

We thank Donna Knoll for expert help in the preparation of the manuscript.

M. E. D. Martin was supported by a States of Jersey Education Department grant.

#### REFERENCES

- Akrigg, A., G. W. G. Wilkinson, and J. D. Oram. 1985. The structure of the major immediate-early gene of human cytomegalovirus strain AD169. *Virus Res.* **2**:107-121.
- Albrecht, J. C., J. Nicholas, D. Biller, K. R. Cameron, B. Biesinger, C. Newman, S. Wittman, M. A. Craxton, H. Coleman, B. Fleckenstein, and R. W. Honess. 1992. Primary structure of the herpesvirus saimiri genome. *J. Virol.* **66**:5047-5058.
- Bacchetti, S., M. J. Eveleigh, and B. Muirhead. 1986. Identification and separation of the two subunits of the herpes simplex virus ribonucleotide reductase. *J. Virol.* **57**:1177-1181.
- 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.
- Bankier, A. T., K. M. Weston, and B. G. Barrell. 1987. Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymol.* **155**:51-93.
- Bublot, M., P. Lamonte, A.-S. Lequarre, J.-C. Albrecht, J. Nicholas, B. Fleckenstein, P.-P. Pastoret, and E. Thiry. 1992. Genetic relationships between bovine herpesvirus 4 and gammaherpesviruses Epstein-Barr virus and herpesvirus saimiri. *Virology* **190**:654-665.
- Carrigan, D. R., K. K. Knox, and M. A. Tapper. 1990. Suppression of human immunodeficiency virus type 1 replication by human herpesvirus-6. *J. Infect. Dis.* **162**:844-851.
- Chang, C. K., and N. Balachandran. 1991. Identification, characterization, and sequence analysis of a cDNA encoding a phosphoprotein of human herpesvirus 6. *J. Virol.* **65**:2884-2894.
- Chang, C. K., and N. Balachandran. 1991. Identification, characterization, and sequence analysis of a cDNA encoding a phosphoprotein of human herpesvirus 6. *J. Virol.* **65**:7085. (Erratum.)
- Chang, Y.-N., D. L.-Y. Dong, G. S. Hayward, and S. D. Hayward. 1990. The Epstein-Barr virus Zta transactivator: a member of the bZIP family with unique DNA-binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif. *J. Virol.* **64**:3358-3369.
- Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchinson III, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. M. Weston, and B. G. Barrell. 1990. Analysis of the protein coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.* **154**:125-169.
- Chee, M. S., S. C. Satchwell, E. Preddie, K. M. Weston, and B. G. Barrell. 1990b. Human cytomegalovirus encodes three G protein-coupled receptor homologues. *Nature (London)* **344**:774-777.
- Cohen, E. A., J. Charron, J. Perret, and Y. Langelier. 1985. Herpes simplex virus ribonucleotide reductase induced in infected BHK-21/C13 cells: biochemical evidence for the existence of two non-identical subunits, H1 and H2. *J. Gen. Virol.* **66**:733-745.
- Colberg-Poley, A. M., L. D. Santomenna, P. P. Harlow, P. A. Benfield, and D. J. Tenney. 1992. Human cytomegalovirus US3 and UL36-38 immediate-early proteins regulate gene expression. *J. Virol.* **66**:95-105.
- Craxton, M. 1991. Linear amplification sequencing, a powerful method for sequencing DNA. *Methods* **3**:20-26.
- Cullen, B. R. 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. *Cell* **46**:973-982.
- Davison, A. J., and J. E. Scott. 1986. The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* **67**:1759-1816.
- Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387-395.
- Efstathiou, S., Y. M. Ho, S. Hall, C. J. Styles, S. D. Scott, and U. A. Gompels. 1990. Murine herpesvirus 68 is genetically related to gammaherpesviruses, Epstein-Barr virus and herpesvirus saimiri. *J. Gen. Virol.* **71**:1365-1372.
- Efstathiou, S., G. L. Lawrence, C. M. Brown, and B. G. Barrell. 1992. Identification of homologues to the human cytomegalovirus US22 gene family in human herpesvirus-6. *J. Gen. Virol.* **73**:1661-1671.
- Ensoli, B., P. Lusso, F. Schachter, S. F. Josephs, J. Rappaport, F. Negro, R. C. Gallo, and F. Wong-Staal. 1989. Human herpesvirus-6 increases HIV-1 expression in co-infected T cells via nuclear factors binding to the HIV-1 enhancer. *EMBO J.* **8**:3019-3027.
- Ertl, P. F., and K. L. Powell. 1992. Physical and functional interaction of human cytomegalovirus DNA polymerase and its accessory protein (ICP36) expressed in insect cells. *J. Virol.* **66**:4126-4133.
- Fixman, E. D., G. S. Hayward, and S. D. Hayward. 1992. *trans*-acting requirements for replication of Epstein-Barr virus ori-Lyt. *J. Virol.* **66**:5030-5039.
- Frame, M. C., H. S. Marsden, and B. M. Dutia. 1985. The ribonucleotide reductase induced by herpes simplex virus type 1 involves minimally a complex of two polypeptides (136K and 38K). *J. Gen. Virol.* **66**:1581-1587.
- Geng, Y., B. Chandran, S. F. Josephs, and C. Wood. 1992. Identification and characterization of a human herpesvirus 6 gene segment that *trans* activates the human immunodeficiency virus type 1 promoter. *J. Virol.* **66**:1564-1570.
- Gibson, W. 1983. Protein counterparts of human and simian cytomegaloviruses. *Virology* **128**:391-406.
- Gompels, U. A., A. L. Carss, N. Sun, and J. R. Arrand. 1992. Infectivity determinants encoded in a conserved gene block of human herpesvirus-6. DNA Sequence J. DNA Sequencing Mapping **3**:25-39.
- Gopal, M. R. 1988. Proteins specified by the oncogenic and attenuated strains of herpesvirus saimiri. Ph.D. thesis. Council for National Academic Awards, United Kingdom.
- Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
- Gottlieb, J., A. I. Marcy, D. M. Coen, and M. D. Challberg. 1990. The herpes simplex virus type 1 UL42 gene product: a subunit of DNA polymerase that functions to increase processivity. *J. Virol.* **64**:5976-5987.
- Honess, R. W. 1984. Herpes simplex and "the herpes complex": diverse observations and a unifying hypothesis. *J. Gen. Virol.* **65**:2077-2107.
- Honess, R. W., U. A. Gompels, B. G. Barrell, M. Craxton, K. R. Cameron, R. Staden, Y.-N. Chang, and G. S. Hayward. 1989. Deviations from expected frequencies of CpG dinucleotides in herpesvirus DNAs may be diagnostic of differences in the states of their latent genomes. *J. Gen. Virol.* **70**:837-855.
- Honess, R. W., and D. H. Watson. 1977. Unity and diversity in the herpesviruses. *J. Gen. Virol.* **37**:15-37.
- Horvat, R. T., C. Wood, and N. Balachandran. 1989. Transactivation of human immunodeficiency virus promoter by human herpesvirus 6. *J. Virol.* **63**:970-973.

35. **Horvat, R. T., C. Wood, S. F. Josephs, and N. Balachandran.** 1991. Transactivation of the human immunodeficiency virus promoter by human herpesvirus 6 (HHV-6) strains GS and Z-29 in primary human T lymphocytes and identification of transactivating HHV-6 (GS) gene fragments. *J. Virol.* **65**:2895-2902.
36. **Jahn, G., T. Kouzarides, M. Mach, B.-C. Scholl, B. Plachter, B. Traupe, E. Preddie, S. C. Satchwell, B. Fleckenstein, and B. G. Barrell.** 1987. Map position and nucleotide sequence of the gene for the large structural phosphoprotein of human cytomegalovirus. *J. Virol.* **61**:1358-1367.
37. **Josephs, S. F., D. V. Ablashi, S. Z. Salahuddin, L. L. Jagodzinski, F. Wong-Staal, and R. C. Gallo.** 1991. Identification of the human herpesvirus 6 glycoprotein H and putative large tegument protein genes. *J. Virol.* **65**:5597-5604.
38. **Kerlavage, A. R., C. M. Fraser, F.-Z. Chung, and J. C. Venter.** 1986. Molecular structure and evolution of adrenergic and cholinergic receptors. *Proteins* **1**:287-301.
39. **Kiehl, A., and D. I. Dorsky.** 1991. Cooperation of EBV DNA polymerase and EA-D (BMRF1) in vitro and colocalization in nuclei of infected cells. *Virology* **184**:330-340.
40. **Knowles, W. A., and S. D. Gardener.** 1988. High prevalence of antibody to human herpesvirus-6 and seroconversion associated with rash in two infants. *Lancet* **ii**:912-913.
41. **Kouzarides, T., A. T. Bankier, S. C. Satchwell, E. Preddy, and B. G. Barrell.** 1988. An immediate-early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology* **165**:151-164.
42. **Kouzarides, T., A. T. Bankier, S. C. Satchwell, K. Weston, P. Tomlinson, and B. G. Barrell.** 1987. Sequence and transcription analysis of the human cytomegalovirus DNA polymerase gene. *J. Virol.* **61**:125-133.
43. **Larder, B. J., S. D. Kemp, and G. Darby.** 1987. Related functional domains in virus DNA polymerases. *EMBO J.* **6**:169-175.
44. **Lawrence, G. L., M. Chee, M. A. Craxton, U. A. Gompels, R. W. Honess, and B. G. Barrell.** 1990. Human herpesvirus 6 is closely related to human cytomegalovirus. *J. Virol.* **64**:287-299.
45. **Leach, F. S., and E. S. Mocarski.** 1989. Regulation of cytomegalovirus late-gene expression: differential use of three start sites in the transcriptional activation of ICP36 gene expression. *J. Virol.* **63**:1783-1791.
46. **Lieberman, P. M., J. M. Hardwick, J. Sample, G. S. Hayward, and S. D. Hayward.** 1990. The Zta transactivator involved in induction of lytic cycle gene expression in Epstein-Barr virus-infected lymphocytes binds both AP-1 and ZRE sites in target promoter and enhancer regions. *J. Virol.* **64**:1143-1155.
47. **Lusso, P., A. De Maria, M. Malnati, F. Lori, S. E. De Rocco, M. Baseler, and R. C. Gallo.** 1991. Induction of CD4 and susceptibility to HIV-1 infection in human CD8<sup>+</sup> T lymphocytes by human herpesvirus 6. *Nature (London)* **349**:533-535.
48. **Lusso, P., B. Ensoli, P. D. Markham, D. V. Ablashi, S. Z. Salahuddin, E. Tschachler, F. Wong-Staal, and R. C. Gallo.** 1989. Productive dual infection of human CD4<sup>+</sup> T lymphocytes by HIV-1 and HHV-6. *Nature (London)* **337**:370-373.
49. **Lusso, P., M. S. Malnati, A. Garzino-Demo, R. W. Crowley, E. O. Long, and R. C. Gallo.** 1993. Infection of natural killer cells by human herpesvirus-6. *Nature (London)* **362**:458-462.
50. **Lusso, P., P. D. Markham, E. Tschachler, F. Di Marzo Veronese, S. Z. Salahuddin, D. V. Ablashi, S. Pahwa, K. Krohn, and R. C. Gallo.** 1988. In vitro cellular tropism of human B-lymphotropic virus (human herpesvirus-6). *J. Exp. Med.* **167**:1659-1670.
51. **Marsden, H. S., M. E. M. Campbell, L. Haarr, M. C. Frame, D. S. Parris, M. Murphy, R. G. Hope, M. T. Muller, and C. M. Preston.** 1987. The 65,000-M, DNA-binding and virion *trans*-inducing proteins of herpes simplex virus type 1. *J. Virol.* **61**:2428-2437.
52. **Martin, M. E. D., J. Nicholas, B. J. Thomson, C. Newman, and R. W. Honess.** 1991. Identification of a transactivating function mapping to the putative immediate-early locus of human herpesvirus 6. *J. Virol.* **65**:5381-5390.
53. **Martin, M. E. D., B. J. Thomson, R. W. Honess, M. A. Craxton, U. A. Gompels, M.-Y. Liu, E. Littler, J. R. Arrand, I. Teo, and M. D. Jones.** 1991. The genome of human herpesvirus-6: maps of unit-length and concatemeric genomes for nine restriction endonucleases. *J. Gen. Virol.* **72**:157-168.
54. **McGeoch, D. J.** 1985. On the predictive recognition of signal peptide sequences. *Virus Res.* **3**:271-286.
55. **McGeoch, D. J.** 1989. The genomes of the human herpesviruses: contents, relationships, and evolution. *Annu. Rev. Microbiol.* **43**:235-265.
56. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* **69**:1531-1574.
57. **McGeoch, D. J., A. Dolan, S. Donald, and F. J. Rixon.** 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. *J. Mol. Biol.* **181**:1-13.
58. **Messerle, M., B. Buhler, G. M. Keil, and U. H. Koszinowski.** 1992. Structural organization, expression, and functional characterization of the murine cytomegalovirus immediate-early gene 3. *J. Virol.* **66**:27-36.
59. **Nathans, J.** 1987. Molecular biology of visual pigments. *Annu. Rev. Neurosci.* **10**:163-194.
60. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J. Mol. Biol.* **48**:443-453.
61. **Neipel, F., K. Ellinger, and B. Fleckenstein.** 1991. The unique region of the human herpesvirus 6 genome is essentially collinear with the U<sub>L</sub> segment of human cytomegalovirus. *J. Gen. Virol.* **72**:2293-2297.
62. **Neipel, F., K. Ellinger, and B. Fleckenstein.** 1992. Gene for the major antigenic structural protein (p100) of human herpesvirus 6. *J. Virol.* **66**:3918-3924.
63. **Nicholas, J., K. R. Cameron, H. Coleman, C. Newman, and R. W. Honess.** 1992. Analysis of nucleotide sequence of the rightmost 43kbp of herpesvirus saimiri (HVS) L-DNA: general conservation of genetic organization between HVS and Epstein-Barr virus. *Virology* **188**:296-310.
64. **Nicholas, J., K. R. Cameron, and R. W. Honess.** 1992b. Herpesvirus saimiri encodes homologues of G protein-coupled receptors and cyclins. *Nature (London)* **355**:362-365.
65. **Nicholas, J., L. S. Coles, C. Newman, and R. W. Honess.** 1991. Regulation of the herpesvirus saimiri (HVS) delayed-early 110-kilodalton promoter by HVS immediate-early gene products and a homologue of the Epstein-Barr virus R *trans*-activator. *J. Virol.* **65**:2457-2466.
66. **Nicholas, J., E. P. Smith, L. Coles, and R. Honess.** 1990. Gene expression in cells infected with gammaherpesvirus saimiri: properties of transcripts from two immediate-early genes. *Virology* **179**:189-200.
67. **Nikas, I., J. McLaughlan, A. J. Davison, W. R. Taylor, and J. B. Clements.** 1986. Structural features of ribonucleotide reductase. *Proteins* **1**:376-384.
68. **Pari, G. S., M. A. Kacica, and D. G. Anders.** 1993. Open reading frames UL44, IRS1/TRS1, and UL36-38 are required for transient complementation of human cytomegalovirus *ori*<sub>Lyt</sub>-dependent DNA synthesis. *J. Virol.* **67**:2575-2582.
69. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444-2448.
70. **Perry, L. J., and D. J. McGeoch.** 1988. The DNA sequences of the long repeat region and adjoining parts of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* **69**:2831-2846.
71. **Pertuiset, B., M. Boccara, J. Cebrian, N. Berthelot, S. Chousterman, F. Puvion-Dutilleul, J. Sisman, and P. Sheldrick.** 1989. Physical mapping and nucleotide sequence of a herpes simplex virus type 1 gene required for capsid assembly. *J. Virol.* **63**:2169-2179.
72. **Rixon, F. J., M. D. Davison, and A. J. Davison.** 1990. Identification of the genes encoding two capsid proteins of herpes simplex virus type 1 by direct amino acid sequencing. *J. Gen. Virol.* **71**:1211-1214.
73. **Roizman, B., R. C. Desrosiers, B. Fleckenstein, C. Lopez, A. C. Minson, and M. J. Studert.** 1992. The family herpesviridae: an update. *Arch. Virol.* **123**:425-449.
74. **Rooney, C. M., D. T. Rowe, T. Ragot, and P. J. Farrell.** 1989. The

- spliced BZLF1 gene of Epstein-Barr virus (EBV) transactivates an early EBV promoter and induces the virus productive cycle. *J. Virol.* **63**:3109–3116.
75. **Sandford, G. R., K. Ho, and W. H. Burns.** 1993. Characterization of the major locus of immediate-early genes of rat cytomegalovirus. *J. Virol.* **67**:4093–4103.
76. **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.
77. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
78. **Spira, T. H., L. J. Bozeman, K. C. Sanderlin, D. T. Warfield, P. M. Feorino, R. C. Holman, J. E. Kaplan, D. B. Fishbein, and C. Lopez.** 1990. Lack of correlation between human herpesvirus-6 infection and the course of human immunodeficiency virus infection. *J. Infect. Dis.* **161**:567–570.
79. **Staden, R.** 1982. Automation of the computer handling of the gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res.* **10**:4731–4751.
80. **Staden, R.** 1984. Graphic methods to determine the function of nucleic acid sequences. *Nucleic Acids Res.* **12**:521–538.
81. **Staden, R.** 1984. Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes. *Nucleic Acids Res.* **12**:551–567.
82. **Staden, R.** 1986. The current status and portability of our sequence handling software. *Nucleic Acids Res.* **14**:217–231.
83. **Stasiak, P. C. and E. S. Mocarski.** 1992. Transactivation of the cytomegalovirus ICP36 gene promoter requires the  $\alpha$  gene product TRS1 in addition to IE1 and IE2. *J. Virol.* **66**:1050–1058.
84. **Stenberg, R. M., P. R. Witte, and M. F. Stinski.** 1985. Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate-early region 1. *J. Virol.* **56**:665–675.
85. **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.
86. **Takahashi, K., S. Sonoda, K. Higashi, T. Kondo, H. Takahashi, M. Takahashi, and K. Yamanishi.** 1989. Predominant CD4 T-lymphocyte tropism of human herpesvirus 6-related virus. *J. Virol.* **63**:3161–3163.
87. **Tenney, D. J., and A. M. Colberg-Poley.** 1991. Expression of the human cytomegalovirus UL36-38 immediate-early region during permissive infection. *Virology* **182**:199–210.
88. **Teo, I. A., B. E. Griffin, and M. D. Jones.** 1991. Characterization of the DNA polymerase gene of human herpesvirus 6. *J. Virol.* **65**:4670–4680.
89. **Thomson, B. J., and R. W. Honess.** 1992. The right end of the unique region of the genome of human herpesvirus-6 (HHV-6) U1102 contains a candidate immediate-early gene enhancer and a homologue of the human cytomegalovirus US22 gene family. *J. Gen. Virol.* **73**:1649–1660.
90. **Thomson, B. J., M. E. D. Martin, and J. Nicholas.** 1991. The molecular and cellular biology of human herpesvirus-6. *Rev. Med. Virol.* **1**:89–99.
91. **van Santen, V. L.** 1991. Characterization of the bovine herpesvirus 4 major immediate-early transcript. *J. Virol.* **65**:5211–5224.
92. **van Santen, V. L.** 1993. Characterization of a bovine herpesvirus 4 immediate-early RNA encoding a homolog of the Epstein-Barr virus R transactivator. *J. Virol.* **67**:773–784.
93. **Vaughan, P. J., D. J. M. Purifoy, and K. L. Powell.** 1985. DNA-binding protein associated with herpes simplex virus DNA polymerase. *J. Virol.* **53**:501–508.
94. **Weston, K.** 1988. An enhancer element in the short unique region of human cytomegalovirus regulates the production of a group of abundant immediate early transcripts. *Virology* **162**:406–416.
95. **Weston, K., and B. G. Barrell.** 1986. Sequence of the short unique region, short repeats, and part of the long repeats of human cytomegalovirus. *J. Mol. Biol.* **192**:177–208.
96. **Wong, S. W., A. F. Wahl, P.-M. Yuan, N. Arai, B. E. Pearson, K. Arai, D. Korn, M. W. Hunkapiller, and T. S.-F. Wang.** 1988. Human DNA polymerase  $\alpha$  gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. *EMBO J.* **7**:37–47.
97. **Yamamoto, M., J. B. Black, J. A. Stewart, C. Lopez, and P. E. Pellet.** 1990. Identification of a nucleocapsid protein as a specific serological marker of human herpesvirus 6 infection. *J. Clin. Microbiol.* **28**:1957–1962.
98. **Yamanishi, K., T. Okuno, K. Shiraki, M. Takahashi, T. Kondo, Y. Asano, and T. Kurata.** 1988. Identification of human herpesvirus-6 as a causal agent for exanthem subitum. *Lancet* **i**:1065–1067.