12-2000

Characterization of Transcripts Expressed from Human Herpesvirus 6A Strain GS Immediate-Early Region B U16-U17 Open Reading Frames

Linda Flebbe-Rehwaldt
University of Kansas Medical Center

Charles Wood
University of Nebraska-Lincoln, cwood1@unl.edu

Bala Chandran
University of Kansas Medical Center

Follow this and additional works at: http://digitalcommons.unl.edu/virologypub

Part of the Virology Commons


This Article is brought to you for free and open access by the Virology, Nebraska Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Virology Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Characterization of Transcripts Expressed from Human Herpesvirus 6A Strain GS Immediate-Early Region B U16-U17 Open Reading Frames

LINDA M. FLEBBE-REHWALDT, CHARLES WOOD, and BALA CHANDRAN

Department of Microbiology, Molecular Genetics and Immunology, The University of Kansas Medical Center, Kansas City, Kansas 66160, and School of Biological Sciences, University of Nebraska—Lincoln, Lincoln, Nebraska 68588

Received 21 April 2000/Accepted 6 September 2000

Several gene fragments of human herpesvirus 6 (HHV-6) have been shown to activate the human immunodeficiency virus (HIV) type 1 long terminal repeat (LTR). An open reading frame (ORF) designated B701 (Y. Geng, B. Chandran, S. F. Josephs, and C. Wood, J. Virol. 66:1564–1570, 1992), found within a 22-kb HHV-6A strain GS [HHV-6A(GS)] genomic fragment and a 3.8-kb Sau3AI subfragment, was shown to activate the HIV LTR. B701, also known as HHV-6 U16, is located in the immediate-early B (IE-B) region of the genome. The sequence of the 3.8-kb genomic fragment of HHV-6A(GS) is nearly identical to the published sequence of HHV-6A strain U1102, with minor differences. The HHV-6A(GS) B701 ORF (U16) was used to screen an HHV-6A(GS) cDNA library, and two different but overlapping cDNAs were identified. These cDNAs represent differently spliced transcripts ending at different polyadenylation signals. The ORFs included in the cDNAs are positionally homologous to the human cytomegalovirus (HCMV) UL36 ORF. The ORF in one cDNA was generated by splicing together in frame ORFs U17 and U16, and the second cDNA included ORFs U16 and U15. A third differentially spliced cDNA (U16+), was identified by 5′ rapid amplification of cDNA ends. The predicted protein was identical to the U16 portion of the U17/U16 spliced gene product but did not include the U17 portion. 5′ extension analyses of the mRNAs demonstrated that at least two potential transcription initiation sites were used to express the transcripts encoding U17 and U16 gene products. Single-stranded U16 and U17 gene-specific RNA probes hybridized with at least five RNA species from infected cells and demonstrated that the expression of these transcripts was differentially regulated. The U17/U16 spliced gene products were expressed at IE times after infection, but a multiply spliced gene product encoded by U16 was expressed as a late gene. The U17/U16 and the U16+ gene products transactivated the HIV LTR. Thus, while there are similarities to the HCMV UL36-UL38 gene family, some of the IE-B U17/U16 transcripts are unique to HHV-6.

Human herpesvirus 6 (HHV-6) was first isolated from patients with AIDS and lymphoproliferative disorders (28). It was subsequently shown to be the causative agent of a common childhood disease, exanthem subitum (roseola) (1, 9, 10). HHV-6 isolates segregated into two closely related subgroups, A and B, based on their antigenic properties, T-cell line infec-
tivities, restriction endonuclease cleavage, and genomic DNA sequences (1, 4, 7, 15, 31, 40). HHV-6A includes the prototype strains GS and U1102, and HHV-6B includes the prototype strain Z29 and isolates from roseola patients (1, 40). The genome of HHV-6A is approximately 160 kb long and consists of a unique long region of 140 kb flanked by direct-repeat regions of about 10 kb (15, 24, 32). Sequence analyses of the HHV-6 genome show that it is closely related to human cyto-
megalovirus (HCMV), exhibiting a colinear arrangement of genes (6, 12, 24, 25).

Several studies have described potential interactions between HHV-6 and human immunodeficiency virus type 1 (HIV-1) (3, 11, 13, 14, 16, 17, 18, 21, 42). HHV-6A(GS) and HHV-6B(Z29) transactivate the expression of the chla-
romphilic acetyltransferase (CAT) reporter gene under the reg-
ulatory control of the HIV long terminal repeat (HIV LTR) in human peripheral blood mononuclear cells and in T-cell lines (14, 16, 17, 42). Our studies and others have identified at least six different cloned HHV-6 genomic fragments that can trans-
activate the HIV LTR (39). In our studies, the highest level of transactivation was seen with a 22-kb BamHI genomic fragment (pZVB70) of HHV-6A(GS) (14, 16, 17) and with 3.8- and 1.8-kb subfragments of pZVB70 (14). Within the 1.8-kb fragment, an open reading frame (ORF) encoding a 258-ami-
no-acid (aa) protein mediated the transactivation (14). The first methionine encoded by this ORF was in the position correspon-
ding to 115 aa downstream from the start of the ORF, and the region of the ORF corresponding to 143-aa carboxyl terminus of the product encoded a predicted protein of about 19-kDa. This ORF was designated B701, and its product was shown to activate CAT expression from an HIV-
CAT plasmid in a cotransfection assay (14). The product of ORF B701 did not show any significant homology to the other viral proteins with transactivating functions, such as ICP0 and ICP4 of herpes simplex virus type 1, but did share some weak homology with the US22 family of immediate-early (IE) and early genes from HCMV. ORF B701 is positionally homolo-
gous to HCMV UL36 ORF exon 2 (15, 25). Comparison with the HHV-6A(U1102) published sequence shows that the HHV-6A(GS) ORF B701 corresponds to ORF U16. There are two potential TATA boxes upstream of the 143-aa coding region of ORF B701 (Fig. 1). Whether the mRNA for the putative B701 protein is transcribed from these promoter re-
regions or is part of a larger protein generated by splicing is not
known. In HCMV, the UL36 coding sequence has been shown to be generated by an in-frame splicing of the UL36 exons 1 and 2 (6, 12, 36, 37), and sequence analysis of the HHV-6 genome predicts a similar splicing site between the HHV-6A ORFs U17 and U16 (25). Using reverse transcriptase RT PCR, we show evidence of splicing in the transcripts encoded by the HHV-6A(GS) ORFs U16 and U17. Our data demonstrate that a family of transcripts is generated from the IE-B region of HHV-6, similar to the HCMV UL36-UL38 family, the U17/U16 family of transcripts has members that are unique to HHV-6. Others are late gene products. Thus, while there are some similarities to the HCMV UL36-UL38 family of genes, two cDNAs of 1.9 and 1.8 kb including ORF B701 were identified from an HHV-6A(GS) cDNA library. The 5' rapid amplification of cDNA ends (5'-RACE) technique was used to identify the 5' ends of transcripts generated from U17 and U16. Two different TATA boxes located upstream of ORF U17 were identified as potential transcriptional start sites. An additional CDNA generated with a unique splicing event was also identified by 5'-RACE. RT-PCR results suggest that some of the transcripts from this region are IE gene products and others are late gene products. Thus, while there are some similarities to the HCMV UL36-UL38 family, the U17/U16 family of transcripts has members that are unique to HHV-6.

MATERIALS AND METHODS

Cells, virus, and infection procedures. The human T-cell line J-Jhan was maintained as a suspension culture in RPMI 1640 (JRH Biosciences, Lenexa, Kan.) containing 10% heat-inactivated Fetal Clone (JRH Biosciences) and antibiotics. HHV-6A(GS) was propagated in J-Jhan cells by standard procedures described previously (2, 4, 5). The percentage of infected cells was determined by immunofluorescence staining using HHV-6 monoclonal antibodies specific for early and late proteins (2, 4, 5). Cell-free virus was prepared by concentrating the culture supernatant collected from HHV-6A(GS)-infected J-Jhan cells (2, 4, 5). For infection experiments, 5 x 10^6 J-Jhan cells were infected with a multiplicity of infection of 5.5% tissue culture infective doses per cell. Cycloheximide (50 mg/ml) (Sigma Chemicals, St. Louis, Mo.) or phosphonoacetic acid (PAA) (200 mg/ml) (Sigma) was added to some cultures. After being incubated for 2 h at 37°C, the cells were pelleted and the supernatant was replaced with 10 ml of fresh medium. The cycloheximide and PAA were added back to the appropriate cultures. The cultures were maintained for 8 h with and without cycloheximide and for 24 h with and without PAA. Mock-infected J-Jhan cells were incubated for 8 h with and without cycloheximide or for 24 h with and without PAA. Cells were harvested and washed in diethyl pyrocarbonate (Sigma)-treated phosphate-buffered saline. The pellets were frozen at -70°C until the RNA was prepared.

Plasmids. The construction of the pZVB70 plasmid containing the HHV-6A(GS) genomic BamHI B fragment (22.2 kb) in Bluescript vector (Stratagene, La Jolla, Calif.) has been described previously (14, 17, 18). The 3.8-kb subfragment (pSal I) containing the HIV LTR transactivating fragment was generated by the digestion of pZVB70 with SalI and cloning of the resulting fragments into Bluescript (14). The HIV-6 ORF B701 (U16) was amplified from the pZVB70 fragment (Fig. 1) by PCR using the 5' primer GGGTCGACATTATGAAGTCTGCACGTA and the 3' primer GGTAGAGATGGGCGCGGCTC. The PCR product of B701 was ligated in frame with the prokaryotic expression vectors pATH-2, pET-3b, and pGEMEX (Promega, Madison, Wis.) (14, 25, 30). The correct orientation of the cDNA ORFs; the nucleotide positions are numbered from right to left. The small hatched box indicates a 96-aa ORF.

FIG. 1. Comparison of HHV-6A(GS) and HHV-6A(U1102) genomes encoding ORFs U15 to U17 and the relationships of the identified cDNAs. (A) Schematic representation of HHV-6 unique long region (UL). IE-A and IE-B, locations of the IE regions (16, 25). The arrows indicate the 5'-to-3' orientation of the cDNA ORFs. (B and C) Relationships of the ORFs identified in HHV-6A(U1102) (B) and HHV-6A(GS) (C) to comparable regions, as well as their designations, intron locations. The polyadenylation signal used in each transcript is indicated by AAAAA. The arrow indicates the 5' end of the cDNA fragment (Fig. 1) by PCR using the 5' primer GGGTCGACATTATGAAGTCTGCACGTA and the 3' primer GGTAGAGATGGGCGCGGCTC. The PCR product of B701 was ligated in frame with the prokaryotic expression vectors pATH-2, pET-3b, and pGEMEX (Promega, Madison, Wis.) (14, 25, 30). The correct orientation of the cDNA ORFs; the nucleotide positions are numbered from right to left. The small hatched box indicates a 96-aa ORF. (D) Schematic representation of the virus-specific portion (E1E2) of the 1.9-kb cDNA and the included ORF. (E) Schematic representation of the 1.8-kb cDNA and the deduced ORFs.
Antibodies. The pATH-B701 plasmid in Escherichia coli strain RR-1 was induced with indoleacrylic acid to express the TrpE-HHV-6 B701 fusion protein (27). The B701 fusion protein band was cut out from the preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel, electroeluted in an Elutrap apparatus (Schleicher and Schuell, Keene, N.H.), and quantitated by bicinchoninic acid assay (27). BALB/c mice were immunized subcutaneously with 5 μg of B701 fusion protein in Freund’s complete adjuvant and then given a second dose in Freund’s incomplete adjuvant. Monoclonal antibodies were generated using standard procedures described previously (22). The spleen cells were fused with Sp2/O myeloma cells, and the hybridoma supernatants were screened by enzyme-linked immunosorbent assay with B701 protein immobilized on plates. Positive clones were cloned into the M13mp18 vector (Pharmacia). Overlapping deletion mutants of E. coli were separated by agarose gel electrophoresis, electroeluted, and ligated into SalI-digested pBluescript KS(−) and TCTTGCAACTCTGCGGCAGAC (3′) primers used for amplifying the various mRNAs are given in Table 1.

Southern blot analysis. Southern blotting was performed by standard methods using nylon membranes (31). DNA probes were radiolabeled with [32P]dCTP (NEN-DuPont) using a nick translation system (GIBCO/BRL, Gaithersburg, Md.) or with digoxigenin-dUTP using the Genius system (Boehringer Mannheim) according to the manufacturer’s instructions. For radiolabeled probes, specific hybridizing bands on the blots were visualized by autoradiography with XAR-5 film. For digoxigenin-labeled probes, specific hybridization reactions as described above. Labeled samples and molecular mass standards were exposed to XAR-5 film at −70°C for fluorography.

Riboprobe synthesis and Northern blot analysis. To synthesize the riboprobes, the cDNAs of ORFs U17/U16 and ORF U16 alone were cloned into the pGem-T vector. Control riboprobes were made from the HHV-6 early gene ORF 27 cDNA (5) and the GAPDH (glyceraldehyde-3-phosphate dehydrogenase) cDNA. Synthesis of the radiolabeled riboprobes was done in the T7/SP6 transcription system (Promega) according to the manufacturer’s instructions. The lysates were used for immunoprecipitation with monoclonal and rabbit polyclonal antibodies. Lysates containing equal amounts of trichloroacetic acid-precipitable radioactive counts were analyzed by SDS-PAGE. In vitro-translated products were boiled with sample buffer, and equal aliquots were used for each reaction. The primers used for amplifying the various mRNAs are given in Table 1.

In vitro transcription and translation. The cDNA inserts in the vectors containing S9 and T5 promoters were used for in vitro transcription experiments. Synthetic cDNA sense and antisense RNA transcripts using SP6 and T7 RNA polymerases and capping of RNAs at the 5′ ends were carried out using procedures described in the Riboprobe system instruction manual (Promega). RNA transcripts were translated in vitro using [35S]methionine (ICN, Irvine, Calif.) and rabbit reticulocyte lysate protein synthesis system (Promega), as described in the manufacturer’s instructions. Samples of in vitro-translated products were boiled with sample buffer, and equal trichloroacetic acid-precipitable radioactive counts were analyzed by SDS-PAGE. In vitro-translated products migrating between 0.1 and 0.5 M NaCl, 0.15 M NaCl, 1.0 M NaCl, 1.0 M NaCl, and 1.0 M NaCl were used for immunoprecipitation with monoclonal and rabbit polyclonal antibodies. Lysates containing equal amounts of trichloroacetic acid-precipitable counts were mixed with 10 μl of antibodies and 100 μl of protein A-Sepharose beads. Samples were mixed continuously at 4°C for 2 h. The immunoprecipitates were collected, washed, dissociated by boiling them in sample buffer, and analyzed by SDS-PAGE (2, 4).

Radio labeling procedures, radiolabeling, radiomunoprecipitation, and SDS-PAGE. Infected and uninfected cells was treated with DNase, and the complete removal of DNA was tested by PCR using primers designed to amplify an HHV-6A(GS) IE gene and the cellular actin gene. RT-PCR was then performed using a sense primer specific for the anchor and an antisense primer (ATCTCTGGATTGTTCCTTCCGTTG (5′ sense primer) and ATCTCTGGATTGTTCCTTCCGTTG (3′ antisense primer) FLEBBE-REHWALDT ET AL. J. VIROL. 11042 amplified by PCR products from both protocols were ligated into the pGEM-T vector with 3′ T overhangs. Clones were screened for virus-specific inserts by Southern blotting with a digoxigenin-labeled U17 gene as a probe, and positive clones were sequenced. The junction between the virus-specific sequence and the adapter sequence was cloned and the 3′ end was identified as the 5′ end of the messages. The nucleotide numbers from the 3.8-kb genomic sequence were used for numbering the cDNA sequences. Transcription start sites were considered authentic when more than one mature transcript was determined by the adapter sequence ligated on the RACE: cDNA (see Fig. 6C). PCR was carried out with the cDNA made from the Marathon kit using two virus-specific primers designed to amplify across the splice junctions. The 5′ sense primer (ATCTCTGGATTGTTCCTTCCGTTG) starts at position 1715, and the 3′ antisense primer (GACGCGTCTAATTCTGCT ATAT) starts at position 1715 on the 3.8-kb map (see Fig. 5). The PCR products were cloned into pGem-T and sequenced.

Transfection procedures and CAT assays. The HHV-6 ORFs 1, 2, and 3 (HVT-Tat) were cloned into the eukaryotic expression vector pCMV-CAT (Invitrogen, San Diego, Calif.). Transfection of CV-1 monkey kidney cells, HeLa cells, and CEM human lymphoid cells was carried out using the DEAE-dextran method (GATAGTCTCTTTTCTGATATAAC). The cDNAs were purified, and single-stranded anchor oligonucleotides in the kit were ligated to the 5′ ends. PCR was then performed using a sense primer specific for the anchor and an antisense primer designed to amplify an HHV-6A(GS)-infected J-Jhan cells that were approximately 30 or 80% infected. The P1 primer mentioned above was used for the PCR. The double-stranded cDNA PCR products from both protocols were ligated into the pGEM-T vector with 3′ T overhangs. Clones were screened for virus-specific inserts by Southern blotting with a digoxigenin-labeled U17 gene as a probe, and positive clones were sequenced. The junction between the virus-specific sequence and the adapter sequence was cloned and the 3′ end was identified as the 5′ end of the messages. The nucleotide numbers from the 3.8-kb genomic sequence were used for numbering the cDNA sequences. Transcription start sites were considered authentic when more than one mature transcript was determined by the adapter sequence ligated on the RACE: cDNA (see Fig. 6C). PCR was carried out with the cDNA made from the Marathon kit using two virus-specific primers designed to amplify across the splice junctions. The 5′ sense primer (ATCTCTGGATTGTTCCTTCCGTTG) starts at position 1715, and the 3′ antisense primer (GACGCGTCTAATTCTGCT ATAT) starts at position 1715 on the 3.8-kb map (see Fig. 5). The PCR products were cloned into pGem-T and sequenced.

Riboprobe synthesis and Northern blot analysis. To synthesize the riboprobes, the cDNAs of ORFs U17/U16 and ORF U16 alone were cloned into the pGem-T vector. Control riboprobes were made from the HHV-6 early gene ORF 27 cDNA (5) and the GAPDH (glyceraldehyde-3-phosphate dehydrogenase) cDNA. Synthesis of the radiolabeled riboprobes was done in the T7/SP6 transcription system (Promega) according to the manufacturer’s instructions. For radiolabeled probes, specific hybridization reactions as described above. Labeled samples and molecular mass standards were exposed to XAR-5 film at −70°C for fluorography. The riboprobes were labeled by incorporating [32P]dCTP (NEN-DuPont) in the riboprobe reaction mixture. More than 107 cpm of each riboprobe was added to hybridization solutions for each blot. Northern blotting was performed using standard procedures (30). HHV-6A(GS)-infected cells were collected at 96 h postinfection, when about 50% of the cells were expressing viral antigens. Poly(A)+-selected RNA was run on 1 or 1.5% agarose gels denatured with formaldehyde. The blots were hybridized with the HHV-6-specific probe first, stripped by boiling them in 0.1 M SSC for 15 min, and then hybridized with the HHV-6a(GS) probe. Hybridization was performed in 50% formamide, 5× SSC, 0.1% SDS, 0.5 M NaCl, and 0.001 M sodium citrate, and reprobed with the GAPDH-specific probe.

RT-PCR. RNA from infected and uninfected cells was treated with DNase, and the complete removal of DNA was tested by PCR using primers designed to amplify an HHV-6A(GS) IE gene and the cellular actin gene. RT-PCR was then performed using a sense primer specific for the anchor and an antisense primer (ATCTCTGGATTGTTCCTTCCGTTG (5′ sense primer) and ATCTCTGGATTGTTCCTTCCGTTG (3′ antisense primer) FLEBBE-REHWALDT ET AL. J. VIROL. 11042 amplified by PCR products from both protocols were ligated into the pGEM-T vector. Control riboprobes were made from the HHV-6 early gene ORF 27 cDNA (5) and the GAPDH (glyceraldehyde-3-phosphate dehydrogenase) cDNA. Synthesis of the radiolabeled riboprobes was done in the T7/SP6 transcription system (Promega) according to the manufacturer’s instructions. For radiolabeled probes, specific hybridization reactions as described above. Labeled samples and molecular mass standards (Sigma) were electrophoresed in parallel lanes. The gels were stained, destained, infused with 1 M maleic acid, dried on filter paper, and placed in contact with XAR-5 film at −70°C for fluorography.
RESULTS

Comparison of genomic sequences of the HHV-6A(GS) 3.8-kb Sa/I transactivating fragment and the homologous region in HHV-6A(U1102). The transactivating 3.8-kb Sa/I subfragment of the 22-kb BamHI genomic fragment (Fig. 1A) of HHV-6A(GS) was sequenced, and the results were compared with the sequence of the same region of HHV-6A strain U1102 (15, 24) (Fig. 1B and C). Although the sequences of the two strains are homologous in this region, there are significant differences, and for that reason, the base pair coordinates given refer to the 3.8-kb sequence derived in this study. Three leftward ORFs designated GS-ORF A to C are included within the 3.8-kb DNA (Fig. 1C), with ORFs A and B in reading frame 3 and ORF C in reading frame 2. The HHV-6(GS) ORF A initiates at nucleotide position 702 and ends at nucleotide 978, with the first methionine codon at nucleotide 720. The predicted protein is 86 aa in length. HHV-6A (U1102) ORF U17, formerly EFLF1 (15, 25), corresponds to GS-ORF A, but U1102-U17 encodes a larger protein of 133 aa (Fig. 1B). Except for one amino acid, the 88-aa product of GS-ORF A is identical to the first 88-aa stretch of the product of U1102-ORF U17. GS-ORF B starts at nucleotide 1032, ends at nucleotide 1805, and potentially encodes a protein of 258 aa. The first methionine codon in GS-ORF B is at nucleotide 1377 (Fig. 1C). The HHV-6A (U1102) ORF U16, formerly EFLF2 (15, 25) corresponds to GS-ORF B and encodes a protein of 258 aa. The predicted protein of GS-ORF B differs from the U1102-U16 protein by two amino acids. GS-ORF C initiates at nucleotide position 2000 and ends at nucleotide 2398, with a first methionine codon at nucleotide 2069. The predicted protein is 110 aa long and does not show any significant homology with other known herpesvirus sequences. The HHV-6A (U1102) ORF U15, formerly EFLF3 (15, 25), corresponds to GS-ORF C and encodes a protein of 110 aa. It is 100% identical to GS-ORF C. Within the 3.8-kb Sa/I genomic fragment, seven potential TATA sequences have been identified at nucleotide positions 95, 335, 640, 790, 1249, 1355, and 2005 (Fig. 1A). Two potential polyadenylation signal sequences are at nucleotide positions 1869 and 2892.

Transcripts encoding HHV-6A ORFs U16 and -17 are spliced. In the HHV-6A(GS) ORF A (U17) and ORF B (U16) sequence, a potential splice donor site [T/(GT)AAGT] at position 948 and an acceptor site [A/(AG)G] at position 1032 were proposed, based on the sequence homology to HCMV. To determine whether the HHV-6A(GS) ORFs A and B are joined by splicing of the transcript, HHV-6A(GS)-infected total cell RNA was reverse transcribed with random hexamers, and the resulting cDNAs were subjected to PCR using specific primers (F and R) designed to direct amplification across the predicted splice junction (Fig. 2A, primer set 1, and Table 1). Amplifications of 570- and 486-bp products, representing unspliced and spliced RNA, respectively, were predicted. The PCR products were Southern blotted and probed with a digoxigenin-labeled oligonucleotide probe (O) that is specific for ORF B (U16) and internal to the primers used for the cDNA amplification (Fig. 2A, and Table 1). RT-PCR was also performed using primers designed to amplify a 458-bp fragment.
portion of the cDNA was referred to as E1E2. As predicted, the 1.1-kb virus-specific portion of the 1.9-kb cDNA contained both U17 and U16 ORFs (Fig. 1D) spliced together into one large ORF by the removal of the 84-bp genomic intron sequence. The resulting ORF was predicted to encode a protein 335 aa long with a molecular mass of 38 kDa. There were four base pair differences in the virus-specific cDNA compared with the 3.8-kb genomic sequence, but the reading frame was not affected. The polyadenylation signal immediately 3' to the U16 ORF, nucleotide position 1869 (Fig. 1), was used to terminate this transcript. Approximately 800 bp at the 5' end of the 1.9-kb cDNA insert sequence did not share any homology with the HHV-6A(GS) 3.8-kb sequence or with the complete genomic sequences of HHV-6A(U1102) in the data bank (data not shown). However, this sequence was 97% homologous with the human mitochondrial gene coxIII. This indicated that the 1.1-kb portion including U17 and U16 was the only HHV-6A (GS)-specific portion of the 1.9-kb cDNA. The coxIII gene sequences identified in the 1.9-kb cDNA included an ORF, but it was in the orientation opposite to that of the virus-specific ORF. The coxIII portion was probably the result of a cloning artifact that occurred during the generation of the cDNA library.

Identification of a 1.8-kb cDNA including the HHV-6 U16 and U15 ORFs. DNA sequence analysis of the 1.8-kb cDNA revealed that the insert consisted of 1,734 bp and included three ORFs (Fig. 1E). Comparison with the genomic 3.8-kb SalI sequence showed only four base pair differences, which did not affect the reading frames, and demonstrated that this cDNA lacked ORF U17. The ORF at the 5' end of the cDNA started at nucleotide position 1020 and ended at nucleotide position 1805, and the first methionine codon was located at nucleotide position 1377. This ORF was in reading frame 1 and corresponded to the HHV-6A(GS) ORF B, or HHV-6A (U1102) U15 (formerly EFLF3), in the genomic sequence (16, 25) (Fig. 1B, C, and E). The first methionine codon for U15 was at position 2069 and was in a good context for a translational start site. Towards the 3' end of the cDNA (at bp 1511), 160 bp of the genomic sequence was spliced out, and the splice donor and acceptor consensus sequences were identified at the boundaries (from positions 2531 to 2690 of the 3.8-kb SalI sequence). The third ORF generated by this splicing is predicted to encode a protein of 96 aa, but the first methionine codon is not in the appropriate context for a translational start site. The polyadenylation signal identified at the 3' end of this cDNA corresponds to the polyadenylation signal found at position 2892 of the 3.8-kb SalI genomic DNA fragment.

In vitro expression of cDNA inserts containing HHV-6 ORFs U16/U15 and U17/U16. To determine the specificity of the proteins encoded by the 1.8- and 1.9-kb cDNAs, we raised rabbit polyclonal antisera and generated murine monoclonal antibodies specific for the 148-aa carboxyl terminus of the U16 (B701) protein that is common to both cDNA products. These antibodies specifically reacted with the B701 protein expressed in bacteria, as demonstrated by the Western blot shown in Fig. 3A, lane 1. Antibodies recognized a protein with a molecular weight of approximately 20,000 (20K), as well as higher-molecular-weight proteins, probably representing dimeric and multimeric forms of the B701 protein. These antibodies were subsequently used in radioimmunoprecipitation and Western blot reactions with proteins expressed from the cDNA inserts. In the 1.8-kb cDNA, the first ORF, U16, encodes a protein 261

FIG. 2. RT-PCR showing evidence for splicing in the transcripts encoding HHV-6A ORFs U16 and U17. (A) Schematic locations of primer sets and oligonucleotides used in RT-PCR and expected sizes of DNA products amplified from spliced and unspliced transcripts. The ORFs U17 and U16 correspond to Fig. 1 ORFs A and B, respectively. The arrowhead labeled F designates the location of the forward primer, and the arrowhead labeled R designates the location of the reverse primer. The line labeled O indicates the location of the probe used to detect the specific PCR products. (B) Southern blot of RT-PCR products from HHV-6A(GS)-infected cell RNA using primer set 1 (Fig. 2B, U17+U16) and U16 were joined by splicing of the transcript. A single band of 458 bp was detected in the Southern blots of RT-PCR and corresponded to the HHV-6A(GS) ORF B, or HHV-6A (U1102) U15 (formerly EFLF3), in the genomic sequence (16, 25) (Fig. 1B, C, and E). The first methionine codon for U15 was at position 2069 and was in a good context for a translational start site. Towards the 3' end of the cDNA (at bp 1511), 160 bp of the genomic sequence was spliced out, and the splice donor and acceptor consensus sequences were identified at the boundaries (from positions 2531 to 2690 of the 3.8-kb SalI sequence). The third ORF generated by this splicing is predicted to encode a protein of 96 aa, but the first methionine codon is not in the appropriate context for a translational start site. The polyadenylation signal identified at the 3' end of this cDNA corresponds to the polyadenylation signal found at position 2892 of the 3.8-kb SalI genomic DNA fragment.

In vitro expression of cDNA inserts containing HHV-6 ORFs U16/U15 and U17/U16. To determine the specificity of the proteins encoded by the 1.8- and 1.9-kb cDNAs, we raised rabbit polyclonal antisera and generated murine monoclonal antibodies specific for the 148-aa carboxyl terminus of the U16 (B701) protein that is common to both cDNA products. These antibodies specifically reacted with the B701 protein expressed in bacteria, as demonstrated by the Western blot shown in Fig. 3A, lane 1. Antibodies recognized a protein with a molecular weight of approximately 20,000 (20K), as well as higher-molecular-weight proteins, probably representing dimeric and multimeric forms of the B701 protein. These antibodies were subsequently used in radioimmunoprecipitation and Western blot reactions with proteins expressed from the cDNA inserts. In the 1.8-kb cDNA, the first ORF, U16, encodes a protein 261
peptides were specifically immunoprecipitated by two different antibodies in the Western blot reaction with the B701 protein expressed in the pET-3b vector. The 20K protein and the higher-molecular-weight dimeric and/or multimeric forms of B701 protein recognized are indicated by the arrowheads. (B) SDS-PAGE analysis of in vitro-expressed polypeptides from the 1.8-kb cDNA insert. In vitro-synthesized mRNA, transcribed from the cDNA insert in the pGEM-T plasmid, was translated in vitro using rabbit reticulocyte lysate. Lanes 1, 4, and 7, translation reaction components without the addition of RNA. Lanes 2, 6, and 9, translation from RNA transcribed with Sp6 RNA polymerase. Lanes 3, 5, and 8, translation from RNA transcribed with T7 RNA polymerase. Lanes 1 to 3, total translated products. Lanes 4 to 6, immunoprecipitations of translated products using a monoclonal antibody against the B701 (U16) protein. Lanes 7 to 9, immunoprecipitations of translation products using polyclonal rabbit antisera raised against the B701 protein. Samples were analyzed on 12% acrylamide cross-linked with bisacrylamide, and standard molecular mass markers were included in parallel lanes. The numbers indicate approximate molecular masses (in kilodaltons) of the prominent polypeptides identified.

FIG. 3. (A) Reactivity of monoclonal antibody against the B701 (U16) protein. Lane 1, reactivities of antibodies in the Western blot reaction with the B701 protein encoded by the E1E2 insert encoding the 38K protein was ligated in frame with the T7 gene 10 in the pGEMEX vector, and expression of the fusion protein was induced with IPTG (isopropyl-β-D-thiogalactopyranoside). The expected size of the gene 10-E1E2 fusion protein was approximately 66K. In the Western blot reactions, the B701 (U16) monoclonal antibodies recognized a protein of about 66K in the cell pellets from induced bacteria carrying pGEMEX-E1E2 (Fig. 4B, lanes 3 and 5) but not in the pellets from bacteria carrying only pGEMEX (Fig. 4B, lanes 1 and 2). With increased induction time, the larger protein disappeared and there was an increase in the smaller polypeptides recognized by the anti-B701 antibody (Fig. 4B, lane 5), suggesting the fusion protein product was unstable and easily degraded.

Identification of HHV-6 U17/U16 gene products in infected cells. To identify the HHV-6 U16/U17 gene products in the infected cells, uninfected and HHV-6-infected cells collected after 3 days postinfection were labeled for 20 h with [35S]methionine and used in immunoprecipitation assays. Rabbit polyclonal antibodies against part of the U16 (B701) protein...
immunoprecipitated HHV-6(GS)-specific polypeptides with approximate molecular weights of 51, 45, 38, and 34K (Fig. 5A, lane 4). These polypeptides were not recognized from uninfected cells (Fig. 5A, lane 3) and were not recognized by preimmune serum (Fig. 5A, lanes 1 and 2). Similar-size polypeptides from HHV-6-infected cells were also recognized by the monoclonal antibody against U16 (B701) (Fig. 5B, lane 2).

Transcripts originating from HHV-6 U17/U16 genes are differentially spliced and are initiated from two potential transcriptional start sites. To determine whether the 1.8-kb cDNA encoding ORFs U16 and U15 represents a complete or incomplete transcript and to determine the transcriptional start site for the U17/U16-encoded transcripts, 5' RACE was done using mRNA from HHV-6A(GS)-infected cells. The sequencing analysis of the 3.8-kb SalI genomic fragment from HHV-6A (GS) identified numerous TATA sequences, located upstream or within the coding regions of ORFs 17 and 16 (Fig. 6), that could potentially serve as transcription initiation sites. Figure 6 shows a schematic representation of ORFs U16 and U17, the 5' untranslated region, the potential TATA boxes, and the locations of the P1 and P2 primers used for 5' RACE. Sequencing of several cDNAs generated by 5' RACE showed that the cDNAs represented both spliced and unspliced messages (Fig. 6A to C). In the spliced transcripts represented in Fig. 6A, an 84-bp intron sequence was removed in frame between ORFs U17 and U16, generating the combined U17/U16 ORF previously identified in the 1.1-kb cDNA, with the first methionine codon at the 5' end of U17. The 5' ends of two of these spliced transcripts initiated at nucleotide position 680, which is 39 bp downstream of the TATA-1 sequence at nucleotide positions 637 to 641 in the SalI 3.8-kb genomic fragment (Fig. 6A). Since this TATA-1 sequence is 83 bp upstream of the first methionine in ORF U17 (Fig. 6A), this suggested that the TATA box could function as the promoter site for these messages (Fig. 6A). A third spliced transcript and an unspliced transcript, which initiated 54 bp downstream of the TATA-1
A second cDNA was found at approximately position 175. (C) Schematic representation of the SalI site, TTGCAG/TTT, at nucleotide position 802 of the 3.8-kb donor site, TTG/GTATGT, at position 294 and an acceptor site, M, at position 680. The 5' end of the cDNA was located at approximately position 175; an 84-bp intron between the 5' end of the cDNA and the 3' end of the cDNA was located at approximately position 175; an 84-bp intron between U17 and U16 and a 508-bp intron, which included part of U17 and the upstream untranslated region, were removed. The ORF is indicated by a thick line, and the first methionine codon, indicated by M, is at the end of U17. The hatched box underneath the transcript represents the protein translated from the transcript. The bent line indicates intron locations.

The bent line indicates intron locations.

**FIG. 6.** Identification of 5' ends of cDNAs. A schematic representation of the ORFs U16 and U17 from HHV-6A(GS) is shown at the top. The numbers correspond to the sequence numbers of the 3.8-kb SalI genomic fragment. The bent arrows indicate the positions of potential TATA sites for transcription initiation. Transcription of this region occurs in the leftward orientation, and the AAA at the left side of U16 indicates the position of the polyadenylation signal used by the U17/U16 spliced transcript. P1 and P2 indicate the positions of the nested antisense oligonucleotide primers used for the 5' RACE procedures. (A and B) Schematic representations of the cDNAs that encode the U17/U16 splice product. The vertical lines mark the transcriptional start sites, and the numbers above correspond to the 3.8-kb map. The 84-bp intron was removed from both transcripts, generating one continuous ORF represented by a hatched box underneath each transcript. The protein products encoded by these cDNAs are identical, with the first methionine codon at position 720 in U17 (indicated by M). (A) Schematic representation of the 5' RACE products that terminated near the TATA-1 promoter at position 640. The 5' end of the cDNA was located at approximately position 680. (B) Schematic representation of the 5' RACE products that terminated near the TATA-2 promoter at position 95. The 5' end of the cDNA was located at approximately position 175. (C) Schematic representation of a 5' RACE product that terminated near the TATA-2 promoter. The 5' end of the cDNA was at approximately position 175; an 84-bp intron between U17 and U16 and a 508-bp intron, which included part of U17 and the upstream untranslated region, were removed. The ORF is indicated by a thick line, and the first methionine codon, indicated by M, is at the end of U17. The hatched box underneath the transcript represents the protein translated from the transcript. The bent line indicates intron locations.

The site, were also identified, and they could represent incompletely extended cDNAs generated by the 5' RACE technique. Another spliced transcript identified by 5' RACE initiated at nucleotide position 175, which is downstream of a TATA sequence at position 95 and is 625 bp upstream of the first methionine codon in ORF U17 (TATA-2 [Fig. 6B]). This transcript also had the U17 and U16 ORFs spliced together in frame, resulting in a single ORF identical to the ORF identified in Fig. 6A. These results suggested that the transcripts including the U17/U16 ORF were initiated from two potential TATA box sequences. An additional transcript, initiating at nucleotide position 175, was identified and also probably used the TATA-2 site at position 95 (TATA-2 [Fig. 6C]). However, this cDNA represented a transcript with a unique splicing pattern. Sequence analysis of the transcript identified a splice donor site, TTG/GTATGT, at position 294 and an acceptor site, TTGCAG/TTT, at nucleotide position 802 of the 3.8-kb SalI genomic fragment (Fig. 6C). The acceptor site was 80 bp downstream of the first methionine codon in ORF U17. The cDNA sequence continued through the rest of U17 and encountered the identical splice site found between U17 and U16 in the spliced cDNAs identified above (Fig. 6A and B). This spliced transcript included an ORF with the first methionine codon found at the last codon of U17 before the splice junction. The reading frame for this transcript is the same as that for the U17/U16 transcript, but because of the splicing pattern, only U16 would be expressed. This transcript was designated U16+, since the splicing pattern allows the entire ORF U16 plus one methionine codon from ORF U17 to be expressed rather than just the codons for the carboxyl end, as was predicted for B701 (14). The size of the predicted U16+ protein is 29K. To confirm the splicing patterns detected here, the 5' RACE cDNAs were subjected to PCR using a primer specific for the viral sequence near the 5' end of the U16+ cDNA (near position 175). The 3' antisense primer was specific for a site within ORF U16, 150 bp upstream of the 3' end of U16. Sequencing of cDNAs revealed both spliced and unspliced transcripts. A PCR product with a splicing pattern identical to that of the cDNA shown in Fig. 6C was sequenced, confirming that the multiply spliced transcript was present in the cDNA library generated from the 5' RACE Marathon kit.

**Analysis of transcripts generated from the HHV-6(GS) U17/U16 genes.** The 5' RACE results indicated heterogeneity in the transcripts and suggested potential differential regulation of the transcripts from the HHV-6 U17/U16 region. Northern blot reactions were performed to determine the extent of the heterogeneity and to evaluate the kinetics of expression of these transcripts. Radiolabeled antisense riboprobes from the HHV-6 U16, U17/U16, and U27 (early-late) genes and the cellular GAPDH genes were used to probe poly(A)-selected RNA from HHV-6A(GS)-infected and uninfected J-Jhan cells. The U16 riboprobe hybridized with multiple transcripts of 3.9, 2.8, 2.5, and 1.5 to 1.6 kb from RNA purified from infected J-Jhan cells (Fig. 7A, lane 1), and no specific binding was detected with RNA from uninfected J-Jhan cells (Fig. 7A, lane 2). Transcripts of similar sizes were also detected in RNA from infected cells by the radiolabeled U17/U16 riboprobe (Fig. 7B, lanes 1 and 2). Northern blot analysis using an antisense riboprobe generated from the U17 ORF alone was not successful, due to nonspecific binding of the probe to residual rRNA, both in uninfected and infected cells (data not shown). However, there were no transcripts unique to the U17/U16 probe, suggesting that transcripts encoded by the U17 gene were not expressed independently of the U16 gene. As a control for virus-specific transcripts, we used the HHV-6 early gene U27 (5), and the transcripts recognized by the U27 riboprobe (Fig. 7C) were of sizes comparable to those reported previously (5). Though similar amounts of RNA samples and radioactive probes were used in these Northern blot reactions, the U16 and U17 transcripts were visualized only after 72 h of exposure of the autoradiographs while the U27 transcripts were detected by 4-h exposure of the autoradiograph. This suggested that the relative abundances of the U16- and U17-specific transcripts were quite low compared with the expression of the U27-specific transcripts. After the HHV-6-specific transcription pattern was determined, these blots were stripped and rehybridized with a riboprobe specific for GAPDH (Fig. 7). The GAPDH riboprobe hybridized to one species of RNA from both infected and uninfected J-Jhan cells, indicating that the various sizes of viral transcripts detected were not the result of degradation of the poly(A)-selected RNA.

In order to determine if the various HHV-6 U17/U16 transcripts were expressed as IE, early, or late gene products, Northern blot analyses with poly(A)-selected RNA from HHV-6A(GS)-infected J-Jhan cells at 8, 24, and 48 h post-infection were carried out (Fig. 7, lanes 4, 5, and 7). Infected cells were also incubated with the protein synthesis inhibitor cycloheximide for 8 h (Fig. 7, lanes 3) or with the viral DNA synthesis inhibitor PAA for 24 h (Fig. 7, lanes 6). This infection procedure was used in order to synchronize the infection in all the cells of the individual cultures. Infected cells with a cell-free HHV-6 supernatant is not as efficient as allowing cell contact between infected and uninfected cells. Less than 10% of the infected cells were expressing viral antigens by 48 h
postinfection with cell-free HHV-6 supernatant. Consequently, the level of expression of the U17/U16 transcripts was quite low in these experiments and was difficult to detect by Northern blotting. Though the riboprobes used had high specific activities, the time required to detect the bound probe by autoradiography was 3 weeks. This contrasts sharply with the 72-h exposure time required for the Northern blots using RNA derived from cultures in which 50% of the cells were infected (Fig. 7, lanes 1 and 2). The U16 antisense riboprobe detected a single species of about 2.3 kb in RNA collected from cells 8 h after infection (Fig. 7A, lane 4), but it was not detected in cells infected for 8 h in the presence of cycloheximide (Fig. 7A, lane 3). This suggested that the 2.3-kb RNA species is an HHV-6 early gene product. The U16 riboprobe hybridized with several transcripts of about 1.5 to 3.9 kb from cells collected 24 h after infection with HHV-6A(GS) (Fig. 7A, lane 5). Except for the 1.6-kb transcript, the expression of the transcripts was unaffected by PAA (Fig. 7A, lane 6). The expression of only the 1.6-kb band appeared to be affected by PAA during the infection (Fig. 7A, lane 6), suggesting that this RNA is an HHV-6 late gene product and that the others belong to an early class of transcripts (Fig. 7A, lane 6). The estimated sizes of the RNAs detected on these blots vary slightly from the sizes measured in Fig. 7A, lane 1. This is most likely due to the different concentration of agarose used for the gels (1.5% agarose for the gels used in the Northern blot for Fig. 7, lanes 3 to 7, and 1% agarose for the gels used in Fig. 7, lanes 1 and 2).

The antisense riboprobe from U17/U16 hybridized with transcripts similar in size to those recognized by the U16 riboprobe (Fig. 7B, lanes 3 to 7). The 2.3-kb transcript was detected 8 h after infection but not in RNA from cells infected for 8 h in the presence of cycloheximide (Fig. 7B, lanes 3 and 4). The presence of PAA during infection eliminated expression of the 1.6-kb transcript but did not affect the expression of the other U17/U16-encoded transcripts (Fig. 7B, lane 6). There was an additional RNA species, approximately 4.7 kb in size, detected by the U17/U16 transcript that was not recognized by the U16 riboprobe (Fig. 7B, lanes 4 to 7). A band of similar size was also detected by the U17 riboprobe in RNA from uninfected cells (data not shown), and it may represent the U17 portion of the U17/U16 riboprobe nonspecifically binding to a host cell RNA. The absence of this band in Fig. 7A and B, lanes 1 and 2, is most likely due to the shorter exposure time required to visualize these HHV-6 transcripts in RNA collected from cells when a higher percentage of the cells (50%) were infected. The nonspecific binding was apparently weak, but it appeared with extended exposure of the autoradiographs.

The HHV-6 ORF U27 was previously shown to be expressed as an early-late gene product and was included here as a control. The transcripts detected by the U27 antisense riboprobe were much more abundant (Fig. 7C, lanes 3 to 7), and the expression pattern corresponded to the pattern previously published (5, 42). No transcripts were generated 8 h after infection (Fig. 7C, lanes 3 and 4), but by 24 h after infection, six different species were detected (Fig. 7C, lane 5). The presence of PAA diminished the expression of these transcripts slightly (Fig. 7C, lane 6) but did not completely inhibit it. This is the expression pattern that identified U27 as an early and late gene (5, 42). All the blots of RNA from the time course infection were stripped and reprobed with the GAPDH antisense riboprobe (Fig. 7C, lanes 3 to 7). The 2.3-kb transcript was detected on these blots vary slightly from the sizes measured in Fig. 7C, lane 6), suggesting that this RNA is an HHV-6 late gene product. The U16 riboprobe hybridized with several transcripts of about 1.5 to 3.9 kb from cells collected 24 h after infection with HHV-6A(GS) (Fig. 7A, lane 5). Except for the 1.6-kb transcript, the expression of the transcripts was unaffected by PAA (Fig. 7A, lane 6). The expression of only the 1.6-kb band appeared to be affected by PAA during the infection (Fig. 7A, lane 6), suggesting that this RNA is an HHV-6 late gene product and that the others belong to an early class of transcripts (Fig. 7A, lane 6). The estimated sizes of the RNAs detected on these blots vary slightly from the sizes measured in Fig. 7A, lane 1. This is most likely due to the different concentration of agarose used for the gels (1.5% agarose for the gels used in the Northern blot for Fig. 7, lanes 3 to 7, and 1% agarose for the gels used in Fig. 7, lanes 1 and 2).

The antisense riboprobe from U17/U16 hybridized with transcripts similar in size to those recognized by the U16 riboprobe (Fig. 7B, lanes 3 to 7). The 2.3-kb transcript was detected 8 h after infection but not in RNA from cells infected for 8 h in the presence of cycloheximide (Fig. 7B, lanes 3 and 4). The presence of PAA during infection eliminated expression of the 1.6-kb transcript but did not affect the expression of the other U17/U16-encoded transcripts (Fig. 7B, lane 6). There was an additional RNA species, approximately 4.7 kb in size, detected by the U17/U16 transcript that was not recognized by the U16 riboprobe (Fig. 7B, lanes 4 to 7). A band of similar size was also detected by the U17 riboprobe in RNA from uninfected cells (data not shown), and it may represent the U17 portion of the U17/U16 riboprobe nonspecifically binding to a host cell RNA. The absence of this band in Fig. 7A and B, lanes 1 and 2, is most likely due to the shorter exposure time required to visualize these HHV-6 transcripts in RNA collected from cells when a higher percentage of the cells (50%) were infected. The nonspecific binding was apparently weak, but it appeared with extended exposure of the autoradiographs.

The HHV-6 ORF U27 was previously shown to be expressed as an early-late gene product and was included here as a control. The transcripts detected by the U27 antisense riboprobe were much more abundant (Fig. 7C, lanes 3 to 7), and the expression pattern corresponded to the pattern previously published (5, 42). No transcripts were generated 8 h after infection (Fig. 7C, lanes 3 and 4), but by 24 h after infection, six different species were detected (Fig. 7C, lane 5). The presence of PAA diminished the expression of these transcripts slightly (Fig. 7C, lane 6) but did not completely inhibit it. This is the expression pattern that identified U27 as an early and late gene (5, 42). All the blots of RNA from the time course infection were stripped and reprobed with the GAPDH antisense riboprobe. A single band was detected after a 24-hour exposure, indicating that the RNA was not degraded (Fig. 7, lanes 3 to 7). These results suggested that most of the transcripts originating from the HHV-6A(GS) U17/U16 genes are expressed as early transcripts and one transcript (1.6 kb) is expressed as a late transcript.

**RT-PCR identified both IE and late transcripts from the HHV-6A(GS) U17/U16 IE-B region.** The percentage of infected cells in the Northern blot time course experiment was low, resulting in low expression of the HHV-6A(GS) U17/U16-specific transcripts. Since cycloheximide can decrease expression of some IE genes (36, 37), evaluating the expression of HHV-6 U17/U16 with a more sensitive method was necessary. RT-PCR was carried out with RNA from HHV-6A(GS)-infected J-Jhan cells collected at different time points. Three different primer pairs (Table 1) were used to detect U17- and/ or U16-containing transcripts. Figure 8C shows a schematic diagram of the primers used to detect B701 (U16) and the sizes of transcripts and one transcript (1.6 kb) is expressed as a late transcript.

**RT-PCR identified both IE and late transcripts from the HHV-6A(GS) U17/U16 IE-B region.** The percentage of infected cells in the Northern blot time course experiment was low, resulting in low expression of the HHV-6A(GS) U17/U16-specific transcripts. Since cycloheximide can decrease expression of some IE genes (36, 37), evaluating the expression of HHV-6 U17/U16 with a more sensitive method was necessary. RT-PCR was carried out with RNA from HHV-6A(GS)-infected J-Jhan cells collected at different time points. Three different primer pairs (Table 1) were used to detect U17- and/ or U16-containing transcripts. Figure 8C shows a schematic diagram of the primers used to detect B701 (U16) and the sizes of transcripts and one transcript (1.6 kb) is expressed as a late transcript.
of the expected PCR products. Since the 3' antisense PCR primer for the B701 protein was at the 3' end of the ORF, detecting it by PCR from the cDNA primed with the random hexamers was difficult. We therefore modified the RT-PCR procedure slightly and used the B701-specific 3' antisense primer to prime the cDNA synthesis by RT. Subsequently, the same antisense primer and the 5' sense primer (Fig. 8C, primer pair I, and Table 1) were used for the PCR. The B701-specific PCR products were detected by Southern blot analysis using a digoxigenin-labeled oligonucleotide probe hybridizing to a site internal to the primers (Table 1). The transcript including ORF B701 (458-bp PCR product) was detected 24 h after infection (Fig. 8A, lane 6), and expression was not affected by the presence of PAA (Fig. 8A, lane 7). The expression of this transcript was maintained throughout the observed 72 h of infection (Fig. 8A, lanes 8 and 9). We reasoned that the sensitivity of the RT-PCR would be increased by using an antisense primer for the PCR that was internal to the primer used to reverse transcribe the cDNA. The cDNA generated using the initial antisense primer was subjected to PCR using a 3' antisense primer internal to the RT primer and the same 5' sense primer, amplifying a 344-bp PCR product (Fig. 8C, primer pair N, and Table 1). The products were analyzed by Southern blotting, and as can be seen in Fig. 8B, lanes 4 and 5, both spliced (369-bp) and unspliced (453-bp) PCR products from these transcripts were detected in RNA from cells infected for 8 h with and without cycloheximide. This suggested that the U17/U16 transcript was expressed as an IE viral gene and that the expression was maintained throughout early and late stages of infection. RNA from the uninfected control cells was negative for U17/U16-specific transcripts (Fig. 9A and B, lanes 1 to 3), demonstrating the specificity of these procedures.

RT-PCR was also used to verify the U16+ cDNA identified in the 5' RACE experiments (Fig. 7) and to examine the transcripts originating from the upstream promoter (TATA-2 [Fig. 7B and C]). Primers were designed to amplify PCR products across both splice junctions identified in Fig. 6C, and the expected product sizes are shown in Fig. 10C. The outside primers (primer pair 1) amplified several products of different sizes from RNAs taken at various times after infection. The RNA isolated from cells infected with HHV-6A(GS) for 8 h in the presence of cycloheximide contained a 1,088-bp unspliced transcript (Fig. 10A, lane 4) which was detected in small amounts without cycloheximide (Fig. 10A, lane 5). In RNA isolated from cells 24 h after infection, unspliced and multiply spliced products of 1,088, 1,004, 580, and 496 bp were detected (Fig. 10A, lane 6), but in the presence of PAA, only the 1,088-bp product representing the unspliced transcript was detected (Fig. 10A, lane 7). The multiply spliced transcripts were detected 48 and 72 h after infection (Fig. 10A, lanes 8 and 9).
When nested-PCR amplification was done (primer pair N), the unspliced transcript was detected in RNA isolated from cells infected for 8 h with and without cycloheximide (Fig. 10B, lane 4 and 5), but otherwise no additional PCR products were identified (compare Fig. 10A and B). This suggested that the spliced U16\textsuperscript{1} transcript was a late gene product and only the unspliced transcript was expressed under IE conditions. RNA from uninfected controls was negative for the virus-specific transcripts (Fig. 10A and B, lanes 1 to 3). As can be seen in Fig. 10C, the digoxigenin-labeled oligonucleotide probe used in the Southern blot for these transcripts was internal to the primers used for the initial PCR amplification but overlapped the 3’ antisense primer used in the nested-PCR amplification. The nested amplification demonstrated that the products were specific. Southern blotting of these same PCR products was done with a digoxigenin-labeled oligonucleotide probe internal to both sets of PCR primers, and the same products were detected (data not shown).

The quality of the RNA used in these experiments was verified by performing three different control RT-PCRs. It has previously been shown that transcripts derived from HHV-6A U89 genes were expressed as IE gene products and were the

---

**Fig. 9.** Detection of U17/U16-specific transcripts by RT-PCR. Southern blots of U17/U16-specific RT-PCR products detected by a single round of PCR amplification (A) and a second nested round of PCR amplification (B) are shown. (A and B) Details of the identification of RNA samples used and the lanes with each sample’s products are given in the legend to Fig. 7. (C) Schematic representation of ORFs U16 and U17 and expected sizes of PCR products. The arrows mark the locations of the primers used for PCR in panel A (I) and the nested PCR in panel B (N). The line with circles internal to both sets of primers marks the location of the oligonucleotide used to detect the specific transcripts. The thick black lines below indicate the two transcripts predicted, spliced and unspliced, and the predicted sizes of the products are shown to the right. The number in parentheses is the size of the nested product. These numbers correspond to the sizes of the bands marked by arrows in panels A and B.

**Fig. 10.** Detection of U16\textsuperscript{1}-specific transcripts by RT-PCR. Southern blots of U16\textsuperscript{1}-specific RT-PCR products by a single round of PCR amplification (A) and a second nested round of PCR amplification (B) are shown. (A and B) Details of the identification of RNA samples used and the lanes with each sample’s products are given in the legend to Fig. 7. (C) Schematic representation of U16 and U17 and the upstream noncoding region. The arrows mark the two sets of PCR primers used, initial primers (I) used in panel A and the nested primers (N) used in panel B. The line with circles marks the location of the oligonucleotide used to detect the specific products on Southern blots. The thick lines underneath represent the four possible transcripts, and the sizes of the PCR products are shown to the left. The number in parentheses is the size of the nested-PCR product.
most abundantly expressed IE genes in HHV-6 (32). RT-PCR was done using the PCR primers previously shown to amplify the U89-containing transcripts. The primers were designed to amplify across a splice junction, and hence, both spliced and unspliced products were detected. As can be seen in Fig. 11A, the 647-bp spliced and 754-bp unspliced U89 products were detected 8 h after infection, and the transcripts were maintained throughout 72 h of infection, even in the presence of PAA (Fig. 11A, lanes 5 to 9). The PCR products from the initial round were subjected to nested PCR, and under these conditions, the spliced (536-bp) and unspliced (643-bp) U89 products were detected in RNA from cells infected for 8 h in the presence of cycloheximide (Fig. 11A, lanes 4 and 5). The low level of expression of the U89 transcript in the presence of cycloheximide was probably the result of the inefficient infection of cells by the cell-free virus, as reported previously (32). The U89 transcripts were apparently expressed at higher levels than the U17/U16-containing transcripts, since the U89 transcript was detected in RNA from 8-h infections on the initial PCR and the U17/U16 transcripts were not detected from this RNA unless a nested amplification was done (Fig. 9A and B, lane 5).

A second control RT-PCR was done to demonstrate the expression of an HHV-6 late gene product, U100, previously shown to be multiply spliced (26). Primers designed to amplify across two splice junctions were used to detect the gene products in RT-PCR, and the results are shown in Fig. 11B. The 657-bp U100-specific product was detected 48 and 72 h after infection (Fig. 11B, lanes 8 and 9) but not at earlier time points. A faint band was sometimes detected in RNA from cells infected for 24 h without PAA (data not shown), but no product was ever detected 24 h after infection when PAA was present throughout the infection (Fig. 11B, lane 7). This verified that U100 is expressed as a late gene product. Finally, RT-PCR was done to measure the cellular transcript amounts. As can be seen in Fig. 11C, all the RNA samples were positive for the actin transcript, demonstrating that the RNAs from the uninfected and infected cells were not degraded and were present in roughly equivalent amounts. Figure 12 shows schematically the summary of the transcripts detected from RT-PCR.

Transactivation of HIV LTR CAT by ORFs U17/U16 and U16+. To determine the transactivating ability of the ORFs U17/U16 and U16+, the HIV-CAT expression vector and a eukaryotic expression vector containing either U17/U16, U16+, or B701 were cotransfected into CV-1 or CEM cells and the CAT expression was measured (Fig. 13). HHV-6 A(GS) infects the human T-cell line CEM (20), and the transactivating activities of B701 in HSB-2 T cells and CV-1 fibroblasts were shown to be comparable (14). Both the U17/U16 and U16+ ORFs transactivated the HIV LTR (Fig. 13), increasing CAT expression about four- to sevenfold. This level of activation was similar to the transactivation observed previously when B701 was shown to transactivate the HIV LTR (14).

DISCUSSION

Analysis of the DNA sequence of the genome of the HHV-6A strain U1102 (16, 25) demonstrates the close relationship of HHV-6 with HCMV. These two betaherpesviruses have a colinear arrangement of genes in several different regions of their genomes. In HHV-6, one of these, designated the IE-B gene block, contains an ORF (U16) which has been previously shown by us to activate the HIV LTR (14, 17). It was predicted that the HHV-6A(U1102) ORF U16 would be expressed with ORF U17 as a spliced gene product (15, 25). Several observations argue in favor of this prediction. The HHV-6 U17/U16 splice product bears positional homology to the HCMV UL36 IE gene. The UL36 gene is part of the UL36-UL38 gene family from HCMV (6, 12, 15), and the products of this family are transcriptional activators (8, 33, 36). The predicted sequences of the U17/U16 and UL36 proteins both contain several US22 protein-specific motifs first described by Kouzarides et al. (19). The US22 motifs are unique to the betahepesviruses and are found in several IE gene products of HCMV. In the HHV-6 U17/U16 and HCMV UL36 predicted proteins, these homologous motifs are colinear in their arrangement and are brought together by comparable splicing events in the two genes (6, 15, 25).

We have identified a family of transcripts generated from the IE-B region of the HHV-6A(GS) genome. The identifica-
tion of HHV-6 U17/U16 (E1E2) from the 1.9-kb cDNA demonstrates that HHV-6 expresses the U17 and U16 ORFs as the predicted spliced gene product. The splicing event removes an intron of 84 bp, bringing together U17 and U16 into one ORF. The protein predicted from the nucleotide sequence was estimated to be around 38K, and in vitro transcription and translation of this cDNA confirmed the size. The protein generated in vitro and by expression in a prokaryotic expression vector was recognized by polyclonal antisera and by a monoclonal antibody generated against the B701 (U16) protein.

Our Northern blot data confirm that multiple transcripts are at least partially encoded by the U17/U16 region of the genome and are generated from the leftward strand of the HHV-6 genomic DNA. Given the number and sizes of transcripts detected on the Northern blots by U16 and U17/U16 riboprobes, it seems quite likely that additional overlapping transcripts, initiated from alternative promoters, have yet to be identified. While we are not able to conclude which TATA sequence was used to initiate transcription of the 1.8-kb transcript identified, our results show that the 1.8-kb transcript is incomplete. If the complete 1.8-kb transcript initiated from TATA-1 or TATA-2 (Fig. 5), its size would be approximately 2.1 or 2.5 kb, respectively. It is also possible that the 1.8-kb cDNA is part of a much larger transcript initiating from a promoter far upstream of the promoters identified here. The predicted sizes of transcripts expressing the U17/U16 splice product range from roughly 1.2 to 1.7 kb, depending on the precise transcriptional initiation site and the size of the poly(A) tail. The U16+ transcript would also be at least 1.2 kb fully spliced if the 3' end is at the 3' end of U16. However, the 3' end of this transcript is not known. Allowing for variation in transcript size due to the lengths of the poly(A) tails, these sizes can all be correlated with the estimated sizes of the transcripts identified on the Northern blots. The largest transcript, approximately 3.9 kb, seen on Northern blots does not fit with the cDNAs identified to date. The HCMV UL37 transcript is a large transcript of approximately 3.4 kb (8, 36, 37, 38). This transcript initiates from a promoter also used to transcribe HCMV UL37X1, but because of differential splicing, UL37 terminates much farther downstream and includes a large noncoding region. This noncoding region overlaps the HCMV UL36 cDNA, and hence, a UL36-specific probe also detects the large UL37 transcripts in Northern blot analysis. It is possible that a similar situation also occurs in HHV-6.

The level of expression of the entire family of HHV-6 U17/U16 transcripts was low when compared with those of other IE (U89) and early (U27) genes from HHV-6. Since cycloheximide can decrease expression of some IE genes (36, 37), we used a more sensitive RT-PCR technique. Our results show that transcription of the U17/U16-containing transcripts was initiated at one of two different TATA sequences upstream of U17 (Fig. 6 and 12). These transcripts were differentially spliced and differentially expressed. At least some of these transcripts were expressed as IE transcripts, and their expression continued throughout infection. One transcript, U16+, was described as a late gene, as its expression was dependent on DNA replication. This transcript demonstrated a unique and unexpected splicing pattern that would allow the expression of the complete U16 ORF independent of U17 (Fig. 11).

Using RT-PCR with RNA from HHV-6A(U1102)-infected cells, Mirandola et al. (22) classified the U16/U17 transcript as an early transcript, since it was absent in cells treated with inhibitors of protein synthesis. The virus subtype, host cells, method of infection, and RT-PCR protocol described in our studies were all different from the methods described by Mirandola et al. (22) and could account for the differences in the results. RNA samples in the study by Mirandola et al. (22) were said to be free from contaminating DNA, but the detection of an unspliced gp82/105 RT-PCR product at IE time points would indicate the presence of DNA. The presence of DNA in RT-PCRs can interfere with the function of the RT and result in PCR amplification of erroneous transcripts. In our study, all RNA samples were subjected to nested rounds of PCR without prior reverse transcription to be certain all DNA was removed. Furthermore, the quantity of RNA that was subjected to this nested PCR was equivalent to the quantity used in the complete RT-PCR. In the Mirandola et al. (22) studies, only 200 ng of RNA was subjected to PCR, yet 1 μg was used for the RT-PCR. Additionally, the cDNA generated from the reverse transcription was precipitated before PCR. This could have had the affect of concentrating any genomic DNA present.

The result presented by Mirandola et al. (22) supporting the conclusion that the U17/U16 transcript was an IE transcript in HHV-6B but an early transcript in HHV-6A is not entirely

FIG. 13. Activation of HIV LTR by U17/U16 and U16+ ORFs. CV-1 cells (A) and CEM cells (B) were transfected with 2.5 μg of HIV-CAT (A and B, lanes 1) and cotransfected with pRC/RSV vector (V) alone (A, lane 5; B, lane 4) or with different concentrations of plasmids containing the various HHV-6 genes (A, lanes 2, 3, 4, and 6 to 9; B, lanes 2, 3, and 5) and the HIV-Tat gene (A, lane 2; B, lane 6) in the pRC/RSV vector. At 36 h after transfection, the cells were harvested and CAT activities were measured by TLC. Cm, unacetylated [14C]chloramphenicol. 1-, 3-, and 5-AC, acetylated forms of [14C]chloramphenicol. The percentage of CAT activity was calculated by dividing the radioactive counts present in the acetylated chloramphenicol spots by the total radioactivity present in both the acetylated and unacylated chloramphenicol spots. Fold activation was calculated as the increase in CAT activity over the background levels for HIV-CAT with the pRC/RSV vector.
convinced. Perhaps the RT-PCR performed by Miranda et al. (22) was not sensitive enough to detect the low-level expression. Our results indicate that HHV-6A(GS) U17/U16 transcripts are expressed in the presence of protein synthesis inhibitors but at a lower level, similar to the HHV-6B results presented (22). One might conclude that the decrease in expression of U17/U16 that we observed is due to a “leaky transcript,” but an alternative explanation is that there are two pathways regulating expression of this transcript, one requiring the synthesis of another viral protein while the other does not.

The regulation of expression of IE genes in HHV-6 and other herpesviruses is complex. The major IE complex in HHV-6 is encoded by ORFs U86, U89, and U90, and two overlapping transcripts, the 3.5-kb IE-1 and the 4.7-kb IE-2 gene products, are expressed as the result of differential splicing (32). HCMV UL36 is part of the UL36-UL38 gene family that expresses at least four different but overlapping transcripts from three different transcription initiation signals (36, 37, 38), and these transcripts are also differentially regulated. HCMV UL36 is regulated by a promoter that is immediately upstream of the first exon of this gene, and two other promoters that are found approximately 2.5 and 3.0 kb upstream of UL36 exon 1 regulate expression of UL37 and UL37X1, and UL38. Since UL37 and UL37X1 are expressed from the same promoter, the differential expression suggests additional mechanisms are involved in the regulation. Some of these mechanisms could include post-transcriptional means of regulation, such as splicing- and translation-coupled degradation of the transcript (8, 36, 37, 38).

Our studies demonstrate that while the HHV-6 IE-B region and the HCMV UL36-UL38 gene family have some similarities, there are also several differences. Two transcriptional start sites have been identified upstream of HHV-6 U17, which differs from the situation in the homologous HCMV gene. The U16+ cDNA isolated by 5’ RACE studies is also unique to HHV-6. Unlike HCMV, there are two possible polyadenylation signals found in this region. Similar to HCMV UL36, the first one is located immediately 3’ of U16 and was the termination signal used to generate the U17/U16 transcript identified. This other is located about 800 bp downstream of the 3’ end of U16 and was identified as the termination signal in the 1.8-kb cDNA. Since this transcript also encodes HHV-6 U15, a gene unique to HHV-6, it is possible that U15 expression is regulated by one of the two promoters identified. The promoter immediately upstream of the first methionine of U15 could also regulate expression of ORF U15, but it was not investigated here. The HHV-6 U18 and U19 genes are homologous to the HCMV UL37 and UL38 genes, respectively (15, 25), and are part of the HHV-6 IE-B gene block. Since U18 and U19 are not part of the 22-kb BamHI genomic fragment that was originally identified as a transactivating fragment (14, 17), these ORFs were not included in these studies.

Temporal regulation of transcript expression is a hallmark of herpesviruses. The IE, or α, transcripts are detected within a few hours after initiation of infection, and de novo protein synthesis is not required for their expression. Specific viral protein synthesis is required for the expression of the early, or β, transcripts, and the late, or γ, transcripts are expressed when viral DNA synthesis begins. Our results show that expression of the U17/U16 gene products is tightly regulated and complex. The U17/U16 spliced transcript is expressed immediately at a low level and then increases until 24 h after infection and stays high throughout the rest of the replication cycle of the virus (Fig. 8A and B). This is similar to the expression of the HCMV UL36 gene (8, 36, 37). While this gene is also expressed as an IE gene, the presence of protein synthesis inhibitors during the culture period reduces the level of UL36 expression but does not eliminate it. The expression pattern of HHV-6 U17/U16 could be the result of expressing the same transcript from two promoters. It is possible that the 5’ RACE product we detected and show schematically in Fig. 6A is expressed under IE conditions while the product shown in Fig. 6B is expressed under late conditions. Initiation from the TATA-2 promoter (Fig. 6B and C) produced multiply spliced transcripts, and the RT-PCR data show that one is the approximate size of U17/U16. Thus, the TATA-1 promoter would express U17/U16 immediately after infection without the need for viral protein synthesis, and when the synthesis of some unidentified viral gene regulator begins, the TATA-2 promoter is activated. This would allow the maintenance of expression of U17/U16 throughout the HHV-6 replication cycle. However, we cannot rule out the possibility that the 5’ RACE product identified in Fig. 6B is actually an intermediate or incompletely spliced transcript of U16+ (Fig. 6C). Our data also strongly suggest that generation of the spliced transcripts regulated by TATA-2 is dependent on DNA replication (Fig. 10A and B). Furthermore, the regulation of expression of the U16+ gene product may involve splicing and protein synthesis, making U16+ a late gene product. The unspliced HHV-6 RT-PCR product detected in Fig. 10 was seen at IE times in the presence of cycloheximide (Fig. 10A), suggesting that the inhibition of protein synthesis was allowing the transcript to accumulate. It is conceivable that this PCR measured a larger transcript originating from a promoter upstream of U17 but continued through the U17/U16 region of the genome, similar to HCMV UL37.

Our present transactivation studies and other studies (14, 16, 17) show that the U17/U16 region of the HHV-6 genome encodes products capable of activating transcription under the control of the HIV promoter (14, 17, 25). Some of the members of the HCMV US22 family, such as HCMV IRS1 and TRS1 (28, 34), and HCMV UL36 (8) have been shown to function as gene regulators. IRS1 and TRS1 have been shown to work in a cooperative fashion with HCMV IE1 and IE2 for regulating expression of HCMV UL44 (34) and the HCMV major IE gene promoter (28). Similarly, transactivation mediated by UL36 alone was reported to be minimal, but when UL36 was included in cotransfection experiments with other HCMV IE genes, the synergistic response was much greater than the transactivation mediated by the IE gene alone (8). Other studies have shown that HCMV UL36 and possibly UL37 are essential for viral DNA replication (32). In fact, an antisense oligonucleotide specific for the splice donor site common to both transcripts will inhibit the growth of the virus almost completely. These observations suggest that the gene regulation function of the UL36-UL38 gene family may be secondary to the DNA replication function. Given the similarities between UL36 and U17/U16, the possibility that U17/U16 might be involved in DNA replication and/or may function in a cooperative manner with other HHV-6 IE genes to regulate viral gene expression needs to be investigated.

The function of the HHV-6 U16+ ORF should also be investigated. If the transcript terminates at the poly(A) signal found at the 3’ end of the U16 ORF, the protein encoded by this cDNA would be identical to the carboxyl terminus of the U17/U16 protein. The structural homology to other US22 family members is found within the U16 ORF, so U16+ would also be a US22 family member. The protein encoded by U16+ would be about 29 kDa and, like U17/U16, has several potential phosphorylation sites and glycosylation sites. The expression of truncated forms of regulatory proteins is seen in other gene families of HCMV. IRS1 is expressed at two different proteins, the smaller one generated from the carboxyl-terminal domain of the larger (28). The smaller protein, IRS1c, is
expressed throughout infection, but predominantly at early and late times after infection, and was shown to antagonize the activation functions of both TRS1 and IRS1. The IE2 gene of HCMV is also expressed as multiple products resulting from differential splicing. One of these, IE2^{338aa}, is the carboxy-terminal ORF of IE2^{579aa}. This smaller protein has been shown to inhibit the transactivating function of IE2^{579aa}. However, IE2^{338aa} also has its own transactivating function (reviewed in reference 23). It is interesting to speculate that, like IRS1 and IE2^{338aa}, U16+ may be a repressor of the U17/U16 functions in this virus and that this repressor will be necessary to evaluate the functions of proteins encoded by the transcripts generated from the U17/U16 region of the HHV-6 genome.

ACKNOWLEDGMENTS

This study was supported in part by Public Health Service grants AI34502 to B.C., CA 62810 to C.W., and AI3056 to C.W. and B.C. and by the University of Kansas Medical Center Endowment Fund grant 711295 to B.C.

REFERENCES