JOURNAL OF VIROLOGY, June 1995, p. 3490–3500
0022-538X/95/$04.00+0

Vol. 69, No. 6

Identification and Characterization of a cDNA Derived from Multiple Splicing That Encodes Envelope Glycoprotein gp105 of Human Herpesvirus 6

BRIAN PFEIFFER,[1] BRIAN THOMSON,[2] AND BALA CHANDRAN[1]*

Department of Microbiology, Molecular Genetics and Immunology, The University of Kansas Medical Center, Kansas City, Kansas 66160,[1] and Department of Medicine, Addenbrooke's Hospital, Cambridge, United Kingdom[2]

Received 30 November 1994/Accepted 9 March 1995

The glycoprotein complex gp82-gp105 is a major virion envelope glycoprotein complex of human herpesvirus 6 variant A (HHV-6A) and consists of a number of related polypeptides. Monoclonal antibodies (MAbs) 2D4, 2D6, and 13D6 against this glycoprotein complex neutralized HHV-6A infectivity. We have previously reported the isolation, mapping, and characterization of a portion of the viral genomic DNA fragment encoding the gp82-gp105 complex and the identification of the neutralizing epitope (B. Pfeiffer, Z. N. Berneman, F. Neipel, C. K. Chang, S. Tirwatnapong, and B. Chandran, J. Virol. 67:4611–4620, 1993). This gene was further characterized by the identification of a 2.3-kb genomic fragment and by the identification of a 2.5-kb cDNA clone. The genomic sequence contains a short open reading frame (ORF) encoding the epitope recognized by the MAbs. The identified cDNA showed specificity for HHV-6 in Southern blot analysis with viral DNA. In Northern (RNA) blot analysis with total RNA from HHV-6A(GS)-infected cells, the cDNA insert specifically hybridized with several RNA species. Restriction mapping analysis localized this cDNA to the HHV-6A(U1102) genomic *Bam*HI G fragment, at the right end of the unique long segment of the genome and to the *Sal*I L and *Sal*I O fragments within the left and right terminal direct repeat regions, respectively. In vitro transcription and translation of the cDNA revealed a polypeptide of about 88.5 kDa which was glycosylated in the presence of microsomal membranes to a polypeptide of approximately 104.2 kDa. Both polypeptides were immunoprecipitated by MAb 2D6, verifying the identity of the cDNA as encoding the gp105 in the gp82-gp105 complex. Sequence analysis of the cDNA revealed a large ORF potentially encoding a 650-amino-acid protein with 11 potential N-linked glycosylation sites and 18 cysteine residues. A potential membrane-spanning domain is located only near the amino terminus of the putative protein, indicating that gp105 may be a class 2 glycoprotein. Comparison of the cDNA nucleotide sequence with sequences from HHV-6A(U1102) genomic *Bam*HI G and *Sal*I L fragments revealed that the gene encoding gp105 contains 12 exons, spanning over 20 kb of the viral genome, with intron 1 spanning about 8 kb of genomic DNA. The first exon of the cDNA mapped to the right and left terminal direct repeats, while the other exons mapped within the unique long segment of the genome. Exons 1 and 2 are untranslated, and the start site for the ORF and the epitope recognized by MAbs are in exon 3. This cDNA represents the first example of a human herpesvirus envelope glycoprotein that is encoded by a multiply spliced mRNA. The other protein components within the gp82-gp105 complex are most likely derived from the same gene through differential mRNA splicing.

Human herpesvirus 6 (HHV-6) is a recently identified T-lymphotropic herpesvirus, first isolated from the peripheral blood lymphocytes of patients with lymphoproliferative disorders and AIDS (37). It has since been isolated from patients with a variety of other disorders (reviewed in reference 33). Primary infection with HHV-6 occurs early in childhood, resulting in a 90% seropositivity by the age of 3 (33). The various known HHV-6 isolates can be categorized as either HHV-6 variant A (HHV-6A) or HHV-6B on the basis of in vitro growth properties, DNA restriction site polymorphism, and antigenic reactivity to a panel of HHV-6-specific monoclonal antibodies (MAbs) (1–4, 8, 13). HHV-6A isolates are grouped with the prototype HHV-6 strains GS and U1102, while HHV-6B isolates are represented by the prototype HHV-6 strain Z29. In children, HHV-6B is clearly associated with exanthem subitum, more commonly known as roseola (41, 45). The genome of HHV-6 is a linear, double-stranded DNA molecule with a unique long segment (U_L) flanked by left and right direct repeats (DR_L and DR_R) and is estimated to be approximately 170 kbp in length (20, 23, 29, 33), sufficient to code for more than 70 proteins. Regions of DNA sequence heterogeneity between HHV-6A and HHV-6B strains as well as intergroup sequence heterogeneity have been demonstrated (2–4, 11–13, 17, 20, 21). The intergroup heterogeneity is greatest and provides a basis for distinguishing the two variant groups. Sequence studies show that HHV-6 is closely related to human cytomegalovirus (CMV) and possesses conserved gene blocks within the U_L segment of the genome, which are collinear in order and orientation with the human CMV genes (9, 11, 12, 14–17, 20, 22, 25, 32, 44).

Herpesviruses encode at least 5 to 10 virion envelope-associated glycoproteins (36), and these glycoproteins are crucial to the viral replication cycle. They mediate several functions in the interaction between virus and host cells such as attachment to cellular receptors, fusion of membranes, and egress from the cell. Importantly, many of these envelope glycoproteins serve as major targets for the host immune system. Knowledge about HHV-6 glycoproteins is limited, and studies with human sera and MAbs have shown that HHV-6 encodes several glycoproteins (5, 6). Recent sequencing data show that HHV-6 has counterparts to glycoproteins gB, gH, and gL of other herpes-

* Corresponding author. Phone: (913) 588-7043. Fax: (913) 588-7295.

viruses. Studies with MAbs and rabbit antisera have identified these HHV-6 glycoproteins from the infected cells and in the virion envelopes (5, 6, 11, 15–17, 22, 28, 32). In addition to these conserved glycoproteins, using MAbs, we have identified a major envelope glycoprotein of HHV-6 with predominantly group-specific epitopes (5–8). These MAbs react with several glycosylated peptides of about 105, 92, 82, 62, 58, 38, 36, and 31 kDa (5–7) which are collectively designated the gp82-gp105 complex. MAbs 2D4, 2D6, and 13D6 against this glycoprotein complex neutralized HHV-6A infectivity (5–7). In the presence of 300 μg of phosphanoacetic acid per ml, which inhibits the synthesis of HHV-6 DNA, this glycoprotein complex was not detected, suggesting that it probably belongs to the late class of HHV-6 proteins (5–7). In Western blot (immunoblot) reactions, these three MAbs recognized several polypeptides of about 105, 92, 82, 62, and 58 kDa. Polypeptides ranging between 105 and 82 kDa were recognized by the MAbs from [^{3}H]glucosamine- and ^{125}I-labeled virions (5, 7). Pulse-chase experiments and assays with *N*-glyconase and tunicamycin suggest at least two major precursor polypeptides of about 58 and 84 kDa for the various polypeptides in the gp82-gp105 complex (5, 7). This observation suggests that the precursors for the gp82-gp105 complex may be the products of the same gene which are generated by proteolytic cleavage, by differential splicing, or by internal initiation of overlapping reading frames or the products of a family of closely related genes. To determine the genetic basis of relationships among the various polypeptides in the gp82-gp105 complex, MAb 2D6 was used to immunoscreen λgt11 recombinant plaques from an HHV-6A (GS) genomic library. A 624-bp gene fragment encoding a portion of the gp82-gp105 was identified, and the neutralizing epitope recognized by the three MAbs was mapped (34). This fragment showed HHV-6 specificity in Southern blot analysis and mapped to the right end of the U_L region of the HHV-6A(U1102) genome (34). In this report, we describe the identification and characterization of a 2.3-kb genomic fragment from the HHV-6A(GS) DNA and a 2.5-kb cDNA clone from the RNA of HHV-6A(GS)-infected cells. In vitro transcription and translation experiments suggest that the cDNA insert encodes gp105 in the HHV-6A gp82-gp105 complex. We have mapped the cDNA insert, identified the transcripts, and sequenced the cDNA insert. Sequence analysis revealed a single long open reading frame (ORF) potentially encoding a protein of 650 amino acids with 11 potential N-glycosylation sites. Comparison of the cDNA nucleotide sequence with the genomic sequence of HHV-6A(U1102) revealed that the cDNA originated from a highly spliced mRNA transcript. Sequence comparison revealed no homology to any known herpesvirus protein, and thus the gp82-gp105 complex appears to be HHV-6 specific.

MATERIALS AND METHODS

Cell cultures and virus infection. HHV-6A strains GS and U1102 were gifts from R. C. Gallo (National Cancer Institute, Bethesda, Md.) and from the late R. W. Honess (National Institute for Medical Research, London, England), respectively. HHV-6B(Z29) was a gift from P. Pellett (Centers for Disease Control and Prevention, Atlanta, Ga.). Suspension cultures of human T-cell lines HSB-2 (ATCC CCL 120.1 CCRF-HSB-2), Molt-3 (ATCC CRL 1552), and J-Jhan (13) were used for virus propagation. Cells were grown in RPMI 1640 medium (JRH Biosciences, Lenexa, Kans.) supplemented with 10% heat-inactivated fetal bovine serum and antibiotics. HSB-2 and J-Jhan cells were used for routine propagation of HHV-6A strains GS and U1102, while HHV-6B(Z29) was grown in Molt-3 and J-Jhan cells. Infection of cells and assay procedures have been described previously (6, 8, 9).

MAbs. The production and characterization of MAbs against HHV-6A(GS)-infected cells have been described previously (6). 2D6 is a neutralizing MAb specific for the glycoprotein complex gp82-gp105 of HHV-6A (5, 6, 16).

Construction of genomic libraries in pBluescript. DNA was isolated from sucrose gradient-purified HHV-6A(GS) virions (9, 23, 34). Briefly, purified virus particles (1 ml) were lysed at 37°C in lysis buffer (1% sodium dodecyl sulfate [SDS], 200 μg of proteinase K, 5 mM EDTA [pH 8.0]) and then heated to 65°C for 10 min. After phenol-chloroform extraction, the DNA was precipitated, resuspended, and treated with RNase A prior to cesium chloride purification. Purified DNA was then digested with restriction enzymes and ligated to pBluescript (Stratagene, La Jolla, Calif.) which had been digested with the same restriction enzyme. Ligated material was then transformed into competent *Escherichia coli* NM522 host cells.

Screening of genomic libraries by colony hybridization. Transformed NM522 cells were plated onto Nytran membranes overlaying Terrific broth plates containing 100 μg of ml ampicillin per ml. After 10 h of growth, the membranes were duplicated and grown for an additional 5 to 6 h at 37°C. Duplicate membranes were then removed and placed onto Terrific broth plates containing ampicillin (100 μg/ml) and chloramphenicol (0.15 μg/ml). Colony DNA was bound to membranes by standard methods (38). Membranes were then hybridized to [α-^{32}P]dCTP (Dupont-NEN, Wilmington, Del.)-labeled 82G or 82G del insert probe (see below) overnight at 50°C in hybridization buffer [2× piperazine-*N*,*N*′-bis(ethanesulfonic acid) (PIPES; pH 6.5), 50% formamide, 0.5% SDS, salmon sperm DNA, and approximately 1×10^7 to 2×10^7 cpm of probe per 50 ml of buffer). Membranes were then washed three times at 55°C in 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS. Dried membranes were exposed to XAR5 film (Kodak) at −70°C.

Dideoxy sequencing. For genomic clones in pBluescript, single-stranded sequencing was used. Briefly, clones were transformed into *E. coli* XL1-Blue competent cells, and deletion mutants were constructed by using exonuclease III as specified for the Erase-A-Base system (Promega, Madison, Wis.). Single-stranded DNA was isolated from the pBluescript clones by using the VCSM13 helper phage as recommended by the manufacturer (Promega). The single-stranded DNA was sequenced by Sequenase version 2.0 method (U.S. Biochemical Cleveland, Ohio) according to manufacturer's protocol, using either a T3 or T7 promoter primer. All sequences were run on 6% urea-polyacrylamide gels. For cDNA clones in pGEMEX-1, double-stranded sequencing was performed. Briefly, cDNA deletion mutants were constructed by using the Erase-a-Base system and were transformed into *E. coli* JM109(DE3) competent cells. DNA from deletion mutants was purified through Plasmid Quik minicolumns (Stratagene) and sequenced at 65°C by using a *Bst* polymerase sequencing kit (Bio-Rad, Hercules, Calif.) according to the manufacturer's protocol with either a T3 or Sp6 promoter primer.

Amplification of genomic sequence. Primers for amplification by PCR were designed on the basis of nucleotide sequences derived from sequencing data of 82G and pB2.3 genomic clones. The 5′ primer (P1) sequence was 5′-GGCCTC GAGATGGCGATAACTGTTCCGGT-3′. The 3′ reverse primer (P2) sequence was 5′-TGAGAGCTCCGAACATGATGCTGCGATA-3′. HHV-6A(GS) purified virion DNA was amplified by using the P1 and P2 genomic primers in 1× *Taq* polymerase buffer (200 μM each deoxynucleoside triphosphate, 2 mM $MgCl_2$, 0.4 μM each primer, 2.5 U of *Taq* DNA polymerase) (Promega). Conditions for step cycle were 94°C for 1 min (denaturing), 50°C for 1 min (annealing), and 72°C for 2 min (extension). The cycle was repeated 35 times followed by one 10-min extension at 72°C, using a model 110s TempCycler II (Coy Corporation, Grass Lake, Mich.). The PCR product was purified from an agarose gel by using a Geneclean II kit (Bio 101, Inc., La Jolla, Calif.).

Screening of the cDNA library. A λgt11 cDNA library of HHV-6A(GS)-infected cell mRNA (9) was screened by using standard plaque hybridization techniques (9, 34, 38) with [α-^{32}P]dCTP-labeled 1.3-kb genomic PCR product as the probe. Nytran membranes were hybridized overnight at 55°C in hybridization buffer (1× PIPES [pH 6.5], 50% formamide, 0.5% SDS, salmon sperm DNA, and 2.5×10^7 cpm of radiolabeled probe in a total volume of 35 ml) and then washed in 0.1% SDS–0.1× SSC at 55°C. After drying, membranes were exposed to XAR5 (Kodak) film at −70°C. The positive clone, λ105, was isolated by standard methods (38) and digested with *Eco*RI prior to ligation with *Eco*RI-digested prokaryotic expression vector pGEMEX-1 (Promega). This cDNA subclone was designated pG105.

Southern blot analysis. The purification of HHV-6A(GS) genomic DNA from purified virion particles and the isolation of DNA from infected and uninfected cells have been described previously (9, 19, 34). A panel of cosmid clones spanning the genome of HHV-6A(U1102) (32) was obtained from F. Neipel, Erlangen, Germany. The *Sal*I L genomic fragment from HHV-6A(U1102) in pBluescript vector (42) was a generous gift from L. J. Rosenthal, Georgetown University School of Medicine, Washington, D.C. Southern blotting was performed on Nytran membranes by standard methods (38, 40) with 1 μg of viral genomic DNA or <100 ng of either cosmid or plasmid DNA per lane. Blots were probed at 60°C in hybridization buffer (5× Denhardt's reagent, 6× SSPE [1× SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA {pH 7.7}], 0.5% SDS, 50% formamide, salmon sperm DNA, and 10^7 cpm of radiolabeled probe in 20 ml of buffer) and washed under stringent conditions (55°C in 0.1× SSC).

In vitro transcription from a plasmid template. Approximately 11 μg of cDNA insert containing plasmid pG105 was linearized by using restriction enzyme *Sac*I and purified with Geneclean (Bio 101). The linearized plasmid was blunt ended by using the Klenow fragment of DNA polymerase I. Capped RNA transcripts were synthesized in vitro from the linearized plasmid template by using a Riboprobe kit (Promega). Approximately 10 μg of the template DNA

was used in the transcription reaction [1× transcription buffer, 10 mM dithiothreitol, 10 μg of bovine serum albumin, 120 U of RNasin, 500 μM ATP, 500 μM CTP, 500 μM UTP, 25 μM GTP, 500 μM $m^7G(5')ppp(5')G$, 80 U of Sp6 RNA polymerase] at 33°C for 2 h. After transcription, the DNA template was removed by incubation at 37°C for 30 min with the addition of 7 U of RNase-free DNase and subjected to phenol-chloroform extraction. The in vitro-transcribed RNA was precipitated for 1 h at −70°C with the addition of 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ethanol. The RNA pellet was resuspended in RNase-free Tris-EDTA buffer and stored at −70°C.

Analysis of in vitro-synthesized protein product from pG105. Approximately 1.5 μg in vitro-transcribed RNA was translated in vitro by using rabbit reticulocyte lysate as recommended by the manufacturer (Promega). Three micrograms of in vitro-transcribed RNA was used in glycosylation reaction mixtures containing canine microsomal membranes (Promega). Radioimmunoprecipitation using protein A-agarose was done as described before (6). Equal counts of [^{35}S]methionine-labeled protein products were run on an SDS–9% polyacrylamide gel, fluorographed, and exposed to XAR5 film (Kodak) at −70°C (6).

Synthesis of the radiolabeled RNA probe. The radiolabeled RNA probe was synthesized in vitro by using a Riboprobe kit (Promega). Approximately 2 μg of plasmid pG105 was linearized by using restriction enzyme *Xho*I, phenol-chloroform extracted, precipitated with 0.1 volume of 3 M sodium acetate (pH 7) and 2.5 volumes of ethanol at −70°C, and resuspended in reaction buffer (1× transcription buffer, 500 μM ATP, 500 μM UTP, 500 μM GTP, 12 μM CTP, 5 μl of [α-^{32}P]CTP [specific activity, 3,000 Ci/mmol at 10 mCi/ml; Dupont-NEN], 0.6 μl of RNasin, 10 mM dithiothreitol, 20 U of T3 RNA polymerase). In vitro transcription was carried out at 35°C for 2 h. The DNA template was removed by treating the reaction mixture with 2 U of RNase-free DNase, extracted with phenol-chloroform, and precipitated at −70°C with the addition of 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ethanol. The RNA pellet was resuspended in RNase-free Tris-EDTA buffer and precipitated again. The purified, radiolabeled RNA pellet was resuspended and added to 5 ml of prehybridization buffer for Northern (RNA) blot analysis.

Northern blot analysis. Total cellular RNA was isolated from HHV-6A(GS)-infected J-Jhan cells and uninfected J-Jhan cells as described previously (9, 10, 34). Northern blotting was performed on Nytran membranes by using standard methods (38) with 5 μg of total RNA. Blotted membranes were hybridized overnight at 55°C in RNA hybridization buffer (5× Denhardt's reagent, 5× SSC, 50 mM phosphate buffer [pH 6.5], 0.5% SDS, 50% formamide, 0.25 ml of salmon sperm DNA [10 mg/ml], and 1 μg of radiolabeled in vitro-transcribed RNA in a 10-ml volume). Radiolabeled RNA was prepared from the linearized plasmid pG105. After hybridization, membranes were washed twice for 5 min each time at room temperature in 2× SSC, once for 5 min at 72°C in 0.1× SSC–0.15% SDS, and once for 5 min at room temperature in 0.1× SSC–0.15% SDS. Membranes were exposed to Kodak XAR5 film at −70°C.

Nucleotide sequence accession number. The nucleotide sequence reported here has been deposited in the GenBank database and assigned accession numbers U23466 and U23467.

RESULTS

Isolation and cloning of HHV-6A(GS) genomic clones pB2.3 and pB1.6 encoding the gp82-gp105 complex. We have previously identified and characterized a 624-bp genomic insert from HHV-6A(GS) DNA, designated 82G, which represents a portion of the gene encoding the gp82-gp105 envelope glycoprotein complex (34). This insert contains a 501-nucleotide ORF with an open 5′ end, stop codons near the 3′ end, and a polyadenylation signal sequence. We initially concluded that the 82G insert contains the 3′ portion of the gene encoding this glycoprotein complex. Mapping of the 82G insert on the HHV-6A(U1102) genome localized the 624-bp fragment within the 2.1- and 9.5-kb *Hin*dIII W and F genomic fragments, respectively (23, 29, 34); in the HHV-6A(GS) genome, these fragments are 1.5 and 10 kb in length, respectively (23, 34). These mapping data are shown schematically in Fig. 1A. The open 5′ end of the ORF in the 82G insert is located in the *Hin*dIII F fragment. To clone the genomic sequence 5′ to the ORF, genomic libraries in pBluescript were constructed by using enzyme restricted genomic DNA from HHV-6A(GS). These libraries were screened by colony hybridization with the 82G insert probe, and three positive clones were identified only from the *Hin*dIII library. To determine the 5′ gene fragment, we also used the 82G del probe, which is a clone of the 82G insert that has been deleted from the 3′ end and does not contain the *Hin*dIII restriction site found in the 82G insert

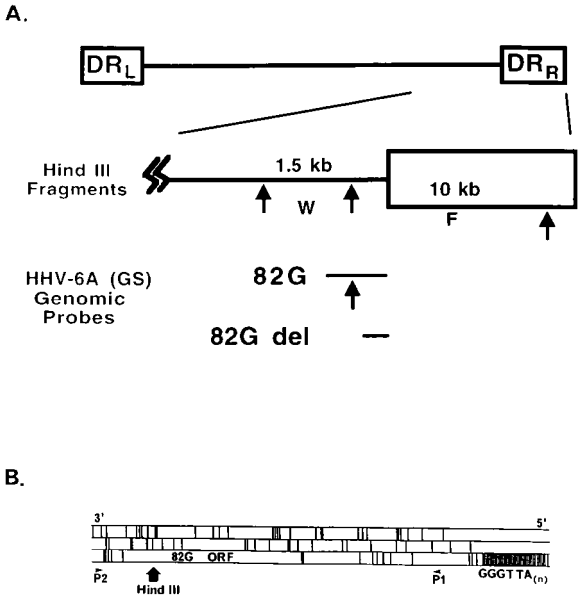

FIG. 1. Genome position and ORF analysis of a gene sequence containing the entire 82G ORF. (A) Map location of the gene encoding the 82G ORF in the HHV-6 genome. The diagram shows the locations of *Hin*dIII fragments W and F on the HHV-6A(U1102) genome. The sizes (in kilobases) of the HHV-6A(GS) *Hin*dIII W and F fragments are indicated. The locations of genomic DNA probes 82G and 82G del are also indicated. The arrows represent the locations of *Hin*dIII restriction sites. (B) ORF analysis of the HHV-6A(GS) gene encoding the neutralizing epitope of the gp82-gp105 complex. The entire sequence of 2.1 kb was compiled from sequencing data for the 82G, pB2.3, and pB1.6 inserts (Fig. 2). The sequence orientation and ORF analysis of the compiled sequence are shown. The telomeric repeat region is indicated as GGGTTA$_{(n)}$. The arrow indicates the *Hin*dIII restriction site that separates the genomic *Hin*dIII fragments W (on the left) and F (on the right). P1 and P2 indicate the locations of the primers used for PCR amplification.

(Fig. 1A). Thus, any *Hin*dIII clones which hybridize to only 82G should contain the 3′ portion of the gene, and any *Hin*dIII clones which hybridize to both probes should contain the 5′ portion of the gene. Upon rescreening of the three colonies with the 82G del probe, only two colonies remained positive. The inserts released from these colonies by *Hin*dIII digestion were 2.3 and 1.6 kb long and were designated pB2.3 and pB1.6, respectively. The mapping data described previously indicate that the *Hin*dIII F fragment containing the 5′ portion of the gene should be about 10 kb long, and the shorter length of the obtained inserts suggests that some sort of deletion and/or rearrangement of genomic sequences had taken place during the cloning of this fragment.

DNA sequence analysis of the HHV-6A(GS) genomic clone pB2.3. Sequencing of the pB2.3 clone revealed that the 5′ of the 82G ORF extended an additional 204 bp, after which there was a series of stop codons (Fig. 1B and 2). The overall length of the ORF within this genomic sequence is 235 amino acids, therefore extending the 167-amino-acid 82G ORF (34) another 68 amino acids. Upstream of the start of the ORF, a series of GGGTTA repeats was identified (Fig. 1B and 2). These repeat sequences in HHV-6 have previously been reported to exist near the termini of both genomic DRs as well as the DR-U_L junctions (20, 27, 29, 43, 44). As in HHV-6A(U1102) and HHV-6B(Z29), the HHV-6A(GS) right U_L-DR junction contains repeats of the telomeric homolog sequence (GGGTTA) referred to as DR2 and several hexameric DR3 repeats (43) as well as the Pac 1 consensus sequence C_n-G_n-N_n-G_n, which is hypothesized to be involved in

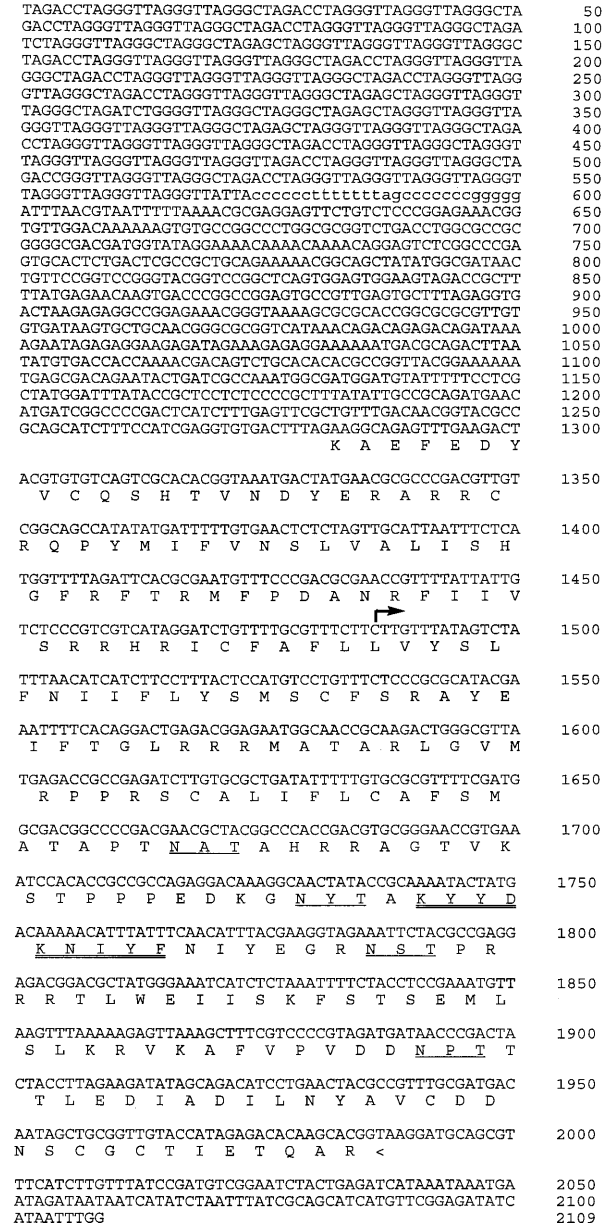

FIG. 2. Nucleic acid sequence of the HHV-6A(GS) gene encoding part of the gp82-gp105 complex and the deduced ORF. Nucleotide positions are numbered on the right, and amino acid positions are numbered on the left. The telomeric repeat region is located between nucleotide positions 8 to 571. Lowercase nucleotide lettering indicates the position of the putative Pac 1 site. The bent arrow indicates the beginning and orientation of the previously identified (34) 82G genomic fragment DNA sequence. Putative N-glycosylation sites (N-X-T/S) are underlined, and the ORF termination codon is indicated (<). The amino acid position of the neutralizing epitope recognized by MAb 2D6 is indicated by a double underline.

replication and packaging of viral genome probably via the rolling-circle model (43). The sequence 5′ to the GGGTTA repeat portion of the pB2.3 insert was identical to the 3′ *Hin*dIII fragment of the 82G insert (data not shown). The presence of this sequence further supports the belief that a deletion on rearrangement event had taken place in this clone. The ORF analysis (Fig. 1B) and DNA sequence (Fig. 2) of the pB2.3 insert shown are modified to include the entire 82G 3′ sequence but do not include the sequence 5′ to the GGGTTA repeats. Sequencing of the pB1.6 insert has shown that it contains sequence that is 3′ coterminal with the pB2.3 sequence and was used to confirm the DNA sequence of this region.

Amplification of a 1.3-kb genomic DNA PCR product. From the sequence data, it was not possible to predict whether the stop codon at the 5′ end of the 82G ORF was valid or whether it was added to the gene through a rearrangement event which might have occurred during the cloning of the genomic fragment. To resolve this question, PCR amplification of genomic DNA was performed with primers designed from the genomic clones pB2.3 and p82G. The 5′ primer (P1), designed from sequences upstream of the 82G ORF within pB2.3, and the 3′ primer (P2), designed from sequences downstream of the 82G ORF within the 624-bp 82G insert, were used to amplify a segment of HHV-6A(GS)-purified virion DNA (Fig. 1B). Since these primer sequences were separated by approximately 1.3 kb in the combined genomic clone sequences, amplification of a PCR product of similar size from purified virion DNA would suggest that the sequences between these primers are contiguous in the viral genome. A 1.3-kb DNA fragment was indeed amplified by PCR from the virion DNA by using primers P1 and P2 (data not shown), suggesting that the 5′ stop codon within the 82G ORF is correctly positioned. Since the unglycosylated precursor polypeptides of the gp82-gp105 complex were about 58 to 84 kDa in size (5, 7), we reasoned that the putative 235-amino-acid protein encoded by the genomic 82G ORF is not likely to be large enough to encode the precursors for the gp82-gp105 complex and that the RNA transcript produced from this gene is probably spliced. Further characterization of this gene necessitated the isolation of cDNA clones.

Identification of a 2.5-kb cDNA clone encoding the HHV-6A(GS) gp82-gp105 complex and its HHV-6 specificity. To identify cDNAs encoding the gp82-gp105 complex, a cDNA library of HHV-6A(GS)-infected cells constructed in λgt11 was screened by plaque hybridization using the [α-^{32}P]dCTP-labeled 1.3-kb PCR product probe generated as described above. After four screenings, three recombinant phages with insert sizes of 1.8, 2.7, and 2.5 kb were identified. Each of these clones also hybridized in Southern blot analysis with the 82G insert, verifying the presence of the neutralizing epitope-coding region in each of the phage inserts (data not shown). The insert from these phages were excised by using *Eco*RI and subcloned into *Eco*RI-digested cloning vectors, including the prokaryotic expression vector pGEMEX-1. However, we were able to clone only the 2.5-kb phage insert, and further characterization of this insert in pGEMEX-1, designated pG105, is reported here. In Western blot analysis, no fusion protein was recognized by MAb 2D6 from the lysate of pG105-transformed bacteria (data not shown). To determine the viral specificity of the cDNA insert and to map the location of the gene encoding the 2.5-kb cDNA on the viral genome, [α-^{32}P]dCTP-labeled 2.5-kb insert was used in Southern blot analysis with HHV-6 DNA. The pG105 insert hybridized specifically with 4.5- and >12-kb *Eco*RI DNA fragments from HHV-6A(GS) purified virions and with a >12-kb HHV-6A(U1102) DNA fragment extracted from infected cells (Fig. 3A, lanes 8 and 9). Upon longer exposure, the insert also showed hybridization with 4.5-kb and >12-kb *Eco*RI-digested HHV-6B(Z29) DNA fragments extracted from infected cells (Fig. 3A, lane 11; Fig. 3B). Differences in the hybridization patterns with the three HHV-6 strains further demonstrate the previously reported restriction site heterogeneity (2, 3, 12, 13, 17, 20, 21). The insert did not hybridize with *Eco*RI-digested DNA from herpes simplex virus types 1 and 2 (HSV-1 and -2), CMV, or Epstein-Barr virus (EBV) or with uninfected host cell DNA (Fig. 3A, lanes 1 to 7

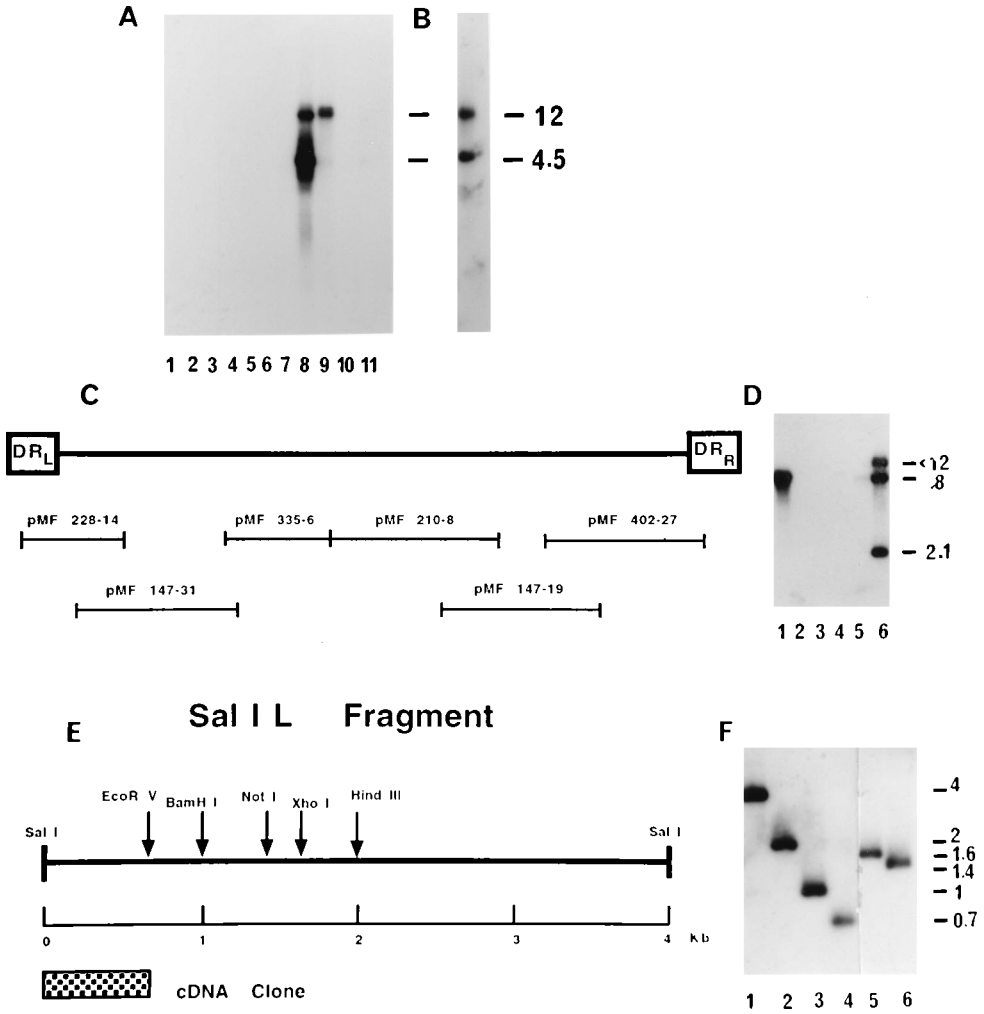

FIG. 3. HHV-6 specificity of the 2.5-kb cDNA insert. (A) Southern blot analysis with the radiolabeled 2.5-kb cDNA insert under high-stringency conditions with uninfected and herpesvirus-infected cell DNAs digested with *Eco*RI. The DNAs used were BHK-21 cell DNA (lane 1), HSV-1 DNA (lane 2), HSV-2 DNA (lane 3), EBV DNA from Raji cells (lane 4), EBV DNA from MCUV cells (lane 5), human CMV DNA (lane 6), HSB-2 cell DNA (lane 7), HHV-6A(GS) DNA from purified virions (lane 8), HHV-6A(U1102) DNA from infected J-Jhan cells (lane 9), uninfected J-Jhan cell DNA (lane 10), and HHV-6B(Z29) DNA from infected Molt-3 cells (lane 11). The sizes (in kilobases) of the HHV-6-specific DNA fragments hybridized are indicated. Standard lambda marker DNAs of known sizes were included in parallel lanes. (B) A longer exposure of lane 11 which contains the HHV-6B(Z29) DNA from infected Molt-3 cells. (C) Mapping of the cDNA, using overlapping HHV-6 cosmid clones. The representation of the HHV-6 genome shows the U_L segment and the terminal DRs. The names of the overlapping HHV-6A(U1102) cosmid clones (32) and their locations on the HHV-6 genome are shown. (D) Southern blot hybridization of the HHV-6A(U1102) cosmid clone DNAs digested with *Hin*dIII and hybridized with the radiolabeled 2.5-kb cDNA insert under high-stringency conditions. The cosmid clones used were pMF228-14 (lane 1), pMF147-31 (lane 2), pMF335-6 (lane 3), pMF210-8 (lane 4), pMF147-19 (lane 5), and pMF402-27 (lane 6). The sizes (in kilobases) of the hybridized DNA fragments are indicated. Standard lambda marker DNAs of known sizes were included in parallel lanes. (E) Mapping of the 2.5-kb cDNA to the HHV-6 DR_L *Sal*I L genomic DNA fragment. The HHV-6 *Sal*I L genomic DNA fragment is shown with an approximate scale in kilobases. The arrows represent sites for the restriction endonucleases. (F) Southern blot analysis of the *Sal*I L fragment is shown. The HHV-6A(U1102) *Sal*I L fragment contained in the plasmid vector pBluescript was digested with *Sal*I (to release the insert) and *Hin*dIII, *Bam*HI, *Eco*RV, *Xho*I, and *Not*I (lanes 2 to 6, respectively). The DNA was analyzed by Southern blot hybridization with the radiolabeled 2.5-kb cDNA insert under high-stringency conditions. Lane 1, plasmid DNA digested with *Sal*I alone. The sizes (in kilobases) of the hybridized DNA fragments are indicated. The cDNA insert mapped to the region of the *Sal*I L fragment indicated by the checkered box.

and 10). These data clearly demonstrate the HHV-6 specificity of the 2.5-kb cDNA insert.

Mapping of the 2.5-kb cDNA insert encoding the HHV-6A(GS) gp82-gp105 complex. When tested with overlapping cosmid clones containing the entire genome of HHV-6A(U1102) (32), the pG105 cDNA insert hybridized with two clones, pMF228-14 and pMF402-27 (Fig. 3C and D). Hybridization of the cDNA insert to two cosmid clones spaced 100 kb apart on the viral genome can be attributed to overlapping sequences within the terminal DR regions. The insert in the pMF228-14 clone was digested with *Not*I to release the insert and then with different restriction enzymes and subsequently hybridized with the labeled 2.5-kb insert. The results of this analysis mapped the pG105 cDNA insert to a 1.4-kb *Sal*I-*Not*I fragment (data not shown). Fine mapping of this region was further accomplished by hybridizing the [α-^{32}P]dCTP-labeled 2.5-kb cDNA insert with the restriction enzyme-digested 4.0-kb cloned *Sal*I L fragment from HHV-6A(U1102). These data localized the cDNA insert to the first 700 bp of the *Sal*I L fragment of HHV-6A(U1102) (Fig. 3E and F). Since the *Sal*I L fragment is located within DR_L, it is also present as a nearly direct copy in DR_R as a *Sal*I O fragment (29, 43).

The pG105 cDNA insert was also analyzed by Southern blot hybridization with restriction enzyme-digested viral DNA ex-

Fragment	Reported Size (Kb)	Observed Size (Kb)
Sal I G	6.7	6.5
Sal I L	4.1	4.1
Sal I O	3.5	3.3
Hind III C	12	12
Hind III F	9.4	9.5
Hind III G	9.0	--
Hind III W	2.05	2.1
EcoR I A	14.5	>12
EcoR I C	13.5	>12
BamH I G	9.5	10
BamH I O	2.5	2.6
Pst I A	12.4	12
Pst I E	9.8	10
Pst I N	5.6	6.0
Xho I A	33	>>12
Xho I O	3.2	3.2
Not I	80	>>12
Not I	7.7	7.5

FIG. 4. Mapping of the 2.5-kb cDNA insert to the HHV-6 genome. The cDNA insert mapped to the HHV-6A(U1102) genome in the regions indicated by the checkered boxes. The observed sizes (in kilobases) of the HHV-6(U1102) genomic fragments that hybridized to the cDNA insert were calculated. Genomic fragment letter designations and reported sizes were taken from reference 29.

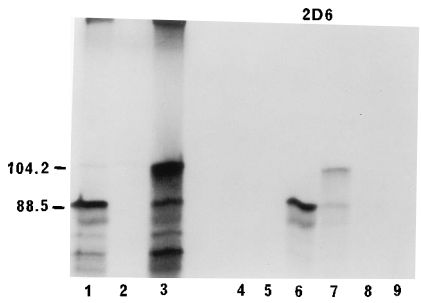

FIG. 5. Analysis of in vitro-synthesized polypeptides from the 2.5-kb cDNA insert. In vitro-synthesized mRNA, transcribed from the cDNA insert of plasmid pG105, was translated in vitro by using rabbit reticulocyte lysate without (lanes 1, 4, 6, and 8) and in the presence of (lanes 3, 5, 7, and 9) canine microsomal membranes. Lane 2 contained the translation reaction components without the addition of RNA. The in vitro-translated polypeptides were immunoprecipitated with protein A-agarose alone (lanes 4 and 5) or with the addition of MAb 2D6 (lanes 6 and 7) or MAb 6A5 (lanes 8 and 9). Samples were analyzed on 9% acrylamide cross-linked with *N-N'*-diallyltartardiamide. Standard molecular size markers were included in parallel lanes. Approximate molecular masses (in kilodaltons) of the polypeptides are indicated and were calculated from sizes of the standard protein markers.

tracted from HHV-6A(U1102)-infected cells and with purified virion DNA from HHV-6A(GS), and intragroup heterogeneity was seen in these hybridization studies (data not shown). The hybridized fragment lengths of HHV-6A(U1102) used for mapping purposes are summarized in Fig. 4. These data, along with previous mapping data, localized the cDNA insert at the right end of the U_L segment of the viral genome within the *Bam*HI G fragment (CB11 [29, 43]) and at the left and right DRs within the *Sal*I L and O fragments, respectively (Fig. 4). Since the cDNA insert mapped to a region within the DRs and at the right end of the U_L segment of the virus genome, but not to the contiguous sequence connecting these regions, it appeared that the cDNA is derived from a spliced mRNA.

In vitro transcription and translation of the 2.5-kb cDNA insert. To determine whether the 2.5-kb cDNA indeed encodes a protein in the gp82-gp105 complex of HHV-6A, in vitro transcription and translation experiments were done. The RNA transcribed from pG105 produced a polypeptide with an apparent molecular mass of 88.5 kDa (Fig. 5, lane 1). With the addition of canine microsomal membranes to the translation reaction mixture, this polypeptide shifted in mobility to a polypeptide of about 104.2 kDa, indicating that glycosylation had taken place (lane 3). Treatment of these in vitro-translated products with *N*-glyconase decreased the apparent molecular mass of the 104.2-kDa polypeptide to about 88 kDa, indicating removal of N-linked oligosaccharides (data not shown). To determine the specificity of the in vitro-synthesized polypeptides, the translated products were immunoprecipitated with MAb 2D6, which recognizes the gp82-gp105 complex from HHV-6A-infected cells. From the in vitro-translated products of pG105, MAb 2D6 specifically immunoprecipitated a polypeptide of about 88.5 kDa and its 104.2-kDa N-glycosylated form (lanes 6 and 7). No polypeptides were immunoprecipitated with protein A-agarose alone (lanes 4 and 5) or with MAb 6A5, which recognizes the HHV-6 gp116-gp64-gp54 complex (lanes 8 and 9). These results clearly demonstrate that the pG105 insert encodes a glycoprotein which contains the neutralizing epitope recognized by MAb 2D6. This cDNA, then, most likely encodes the gp105 polypeptide component within the HHV-6 gp82-gp105 complex.

Northern blot analysis with the 2.5-kb cDNA insert. To define further the viral specificity of 2.5-kb pG105 insert, equal quantities of total RNA or poly(A)$^+$-selected cytoplasmic RNA from HHV-6A(GS)-infected cells and uninfected cells were analyzed by Northern blot hybridization with an antisense RNA probe transcribed from the cDNA insert. Results with total RNA are shown. An in vitro-transcribed [α-^{32}P]CTP-labeled RNA probe was generated from the linearized plasmid pG105 and hybridized with the Northern blotted HHV-6A RNA. No hybridization was detected with Northern blot of total RNA from uninfected cells (Fig. 6, lanes 1 and 3). Under conditions of high stringency, this RNA probe hybridized specifically to at least six RNAs, one of which appeared to consist of an RNA doublet. The calculated sizes of the RNA bands were 3.9, 3.1/2.9, 2.1, 1.6, 1.1, and 0.6 kb (lanes 2 and 4). The bands of strongest intensity were the 3.9-kb band and the 3.1/2.9-kb RNA doublet. This doublet could be distinguished only at low exposure with a probe of high specific activity. Similar results were seen in Northern blot hybridization with poly(A)$^+$-selected cytoplasmic RNA (data not shown). The glyceraldehyde-3-phosphate dehydrogenase gene (G3PDH) is constitutively expressed in the cell. Both infected and uninfected cell RNA hybridized to a G3PDH RNA probe, and this hybridization served as an verification for the integrity of the RNA samples (lanes 5 and 6). These results further confirm the HHV-6 specificity of the pG105 insert.

DNA sequence analysis of the cDNA insert pG105. DNA sequence analysis of plasmid pG105 revealed a cDNA insert consisting of 2,512 bp with an overall mean G+C composition of 45%. The cDNA sequence was further verified by restriction digestion analysis of the predicted restriction enzyme sites and restriction fragment lengths. The complete cDNA sequence and predicted primary amino acid sequence are shown in Fig. 7. Computer analysis of the cDNA sequence revealed a single

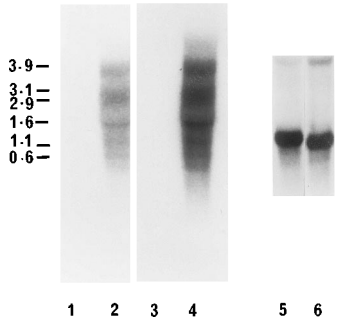

FIG. 6. Northern blot analysis of HHV-6-infected and uninfected cell RNA. Total unfractionated RNA was isolated from uninfected J-Jhan cells (lanes 1 and 3) and HHV-6A(GS)-infected HSB-2 cells (lanes 2 and 4) and analyzed by Northern blot hybridization with a radiolabeled, in vitro-transcribed RNA probe synthesized from the 2.5-kb cDNA insert containing plasmid pG105 (lanes 1 to 4). Lanes 1 and 2 show a lower exposure of the autoradiograph. Lanes 5 and 6 show Northern blot hybridization with a radiolabeled G3PDH RNA probe with RNA from uninfected (lane 5) and infected (lane 6) cells. The sizes (in kilobases) of the hybridized RNAs are indicated. Standard RNA marker molecules of known sizes were included in parallel lanes.

ORF encoding a potential protein of 650 amino acids (Fig. 7 and 8A). The 353 nucleotides at the 5′ end of the cDNA and the 170 nucleotides at the 3′ end are untranslated in this ORF. The initiation methionine (ATG) at nucleotide position 393 conforms to Kozak's rules of translational start sequences (24), having an A at position −3 and a G at position +4. The ORF stop codon (TAA) was located at nucleotide position 2343 (Fig. 7). There is a potential polyadenylation signal sequence (ATTAAA) at the end of the cDNA ORF located between nucleotides 2340 and 2347. The calculated molecular mass of the putative nascent translated polypeptide was 74.5 kDa, with a pI of 8.55. There were 11 potential sites for the addition of N-linked oligosaccharides and 18 cysteine residues spread throughout the predicted primary amino acid sequence. Hydropathic analysis of the cDNA-translated ORF indicated a hydrophobic domain near the amino terminus (Fig. 8B), and analysis with the PC/Gene program (IntelliGenetics Inc., Mountain View, Calif.) predicted amino acids 14 to 30 in this region to be a transmembrane domain that is noncleavable by signal sequence peptidases. Immediately following this region is a large hydrophilic domain which contains the epitope amino acid sequence that is recognized by the neutralizing MAb 2D6 (34). The epitope also lies within an amino acid region with a high antigenic index and a good probability of surface exposure (Fig. 8B). The potential membrane-spanning domain is located only near the amino terminus of the putative protein, which suggests that gp105 may be a class 2 glycoprotein. Preliminary studies with rabbit antibodies against a peptide in the carboxy terminus of the ORF reacted with the surfaces of HHV-6A-infected cells, and supporting the notion that gp105 is a class 2 glycoprotein (data not shown).

No significant homology was detected in either nucleotide or amino acid sequence when the cDNA sequence was compared with other reported herpesvirus sequences or any other non-herpesvirus sequence in the data banks. The approximate molecular mass of the in vitro-translated product from the cDNA (88.5 kDa) is slightly larger than the calculated molecular mass of the putative protein encoded by the cDNA (74.5 kDa); this difference could be due to the anomalous mobility of the protein in SDS-polyacrylamide gel electrophoresis. However, immunoprecipitation of the nonglycosylated 88.5-kDa polypeptide and the glycosylated 104.2-kDa polypeptide from the in vitro-translated products by MAb 2D6 verifies the identity of this cDNA as a message encoding the gp105 in the gp82-gp105 complex.

The cDNA that encodes gp105 represents a highly spliced mRNA transcript. The cDNA nucleotide sequence was compared with the nucleotide sequences of HHV-6A(U1102) genomic fragments *Bam*HI-G and *Sal*I-L to determine sequence alignment and homology. Comparison of the first 252 nucleotides of the cDNA with the complementary strand of the reported *Sal*I-L sequence (42) revealed a genomic DNA region with 95.2% identity (Table 1). Comparison of the remainder of the cDNA sequence with the complementary strand of the reported *Bam*HI-G sequence (29, 43) revealed that several genomic DNA regions aligned with the cDNA with high nucleotide identity and were separated by genomic stretches without identity (Table 1). The boundaries of these high-identity regions were inspected for the presence of consensus splice site nucleotide sequences (31, 39). Splice donor and splice acceptor consensus sequences were identified at the boundaries of each potential intron except intron 2, which had a nonconsensus splice acceptor site (Table 2). The genomic sequence of HHV-6A(GS) that corresponds to this region, however, contains a one-nucleotide substitution which restores consensus to the acceptor site (Table 2). Other genomic nucleotide differences between HHV-6A strains GS and U1102 were identified in the intron 2 donor site and the intron 3 acceptor site and are indicated in Table 2. Overall, there is 97% sequence homology between cDNA nucleotides 1 to 2363 and the HHV-6A(U1102) genomic sequences. Following nucleotide position 2363 is a string of 19 adenines which we believe to represent the true poly(A) tail of the gp105 mRNA, since the DNA sequence following this tail (nucleotides 2383 to 2512) exhibits no homology with any reported HHV-6 DNA sequence or with any known herpesvirus DNA sequence. It is possible that this portion of the cDNA was unrelated to the gp105 mRNA and that its addition was an artifact of cDNA synthesis.

Analysis of individual exon homologies is presented in Table 1. In all, 12 exons were identified in the cDNA sequence. The transcript is in a leftward reading frame. Exons 1 and 2 are untranslated, and the epitope recognized by MAbs is in exon 3. A schematic summary of the alignment of the cDNA with the HHV-6A(U1102) genome and the splice sites which were most likely used in the generation of the mRNA encoding gp105 is presented in Fig. 9. The sequence of the 2.5-kb cDNA from nucleotides 1 to 2363 showed 97% identity with the genomic sequences of HHV-6A(U1102). Alignment of the cDNA with the U1102 genome (Fig. 9) revealed that the first exon of the gene mapped to DR_R, while the other exons mapped within the U_L component. Exons 2 to 4 map nearest the region at the U_L/DR_R junction which has been shown to be a heterogeneous region among all HHV-6 strains (20, 29). Consequently, the calculated nucleotide identity between exons 1 to 4 and genomic sequences was lower (95%) than the calculated identity between exons 5 to 12 and genomic sequences (99%). Thus, the cDNA sequence covers over 20 kb of genomic sequence with intron 1 spanning about 8 kb of genomic DNA.

DISCUSSION

Three immunologically distinct groups of HHV-6 virion glycoproteins from infected cells have been defined by immunoprecipitation with MAbs (5, 6). The first two groups, gp116-gp64-gp54 and gp100-gp80-gp32, represent HHV-6 homologs of glycoproteins gB and gH/gL, respectively, which are conserved throughout the herpesvirus family (5, 11, 15–17, 22, 28). The third virion glycoprotein group was identified in HHV-

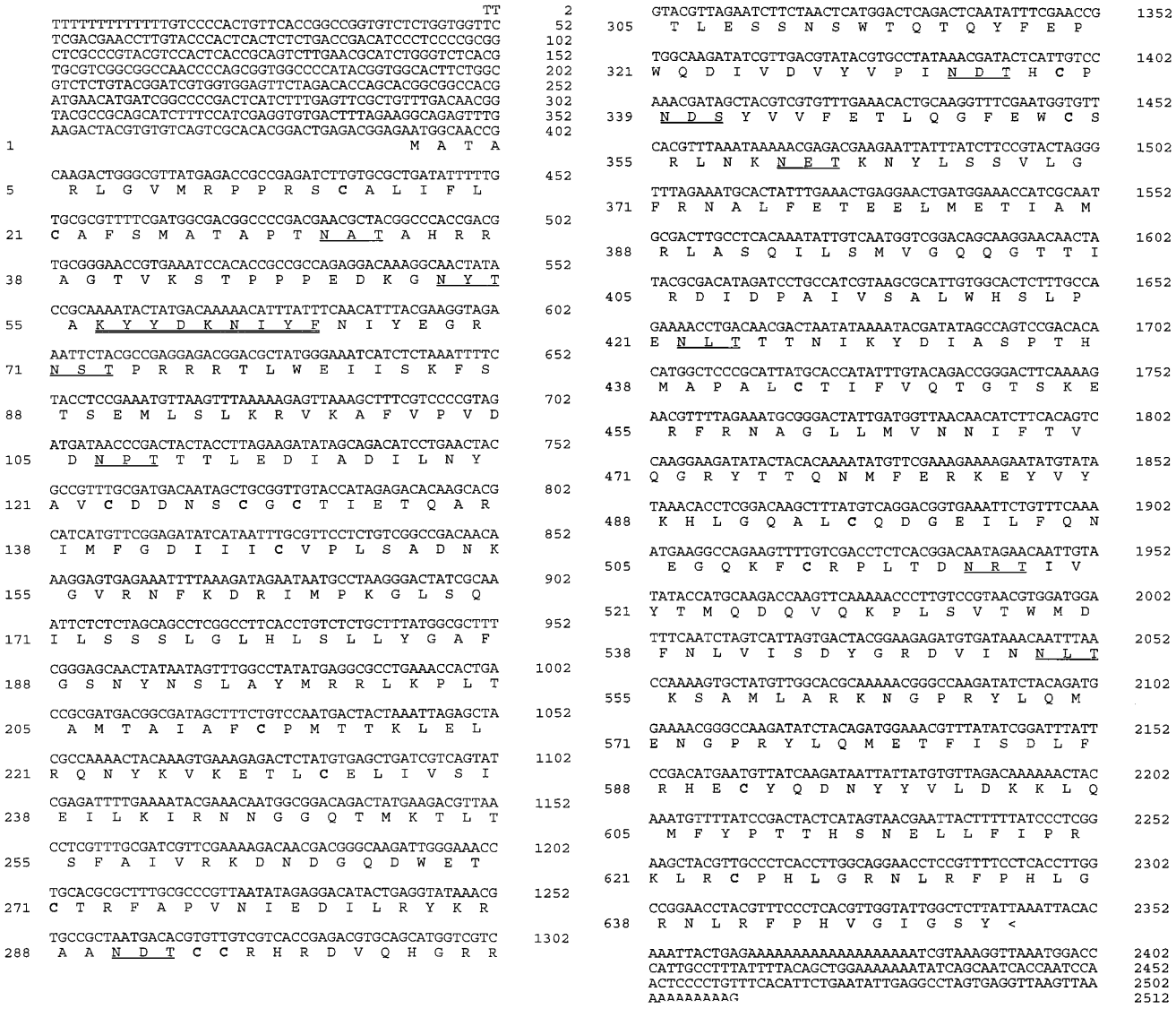

FIG. 7. Nucleic acid sequence of the 2.5-kb cDNA insert and the deduced ORF. Amino acid positions are numbered on the left, and nucleotide positions are numbered on the right. The ORF termination codon (TAA) is indicated (<), and the putative polyadenylation signal sequence (ATTAAA) overlaps this termination codon. Putative N-glycosylation sites (N-X-T/S) are underlined, and cysteine residues are indicated in boldface. The amino acid sequence of the neutralizing epitope from HHV-6A gp105 that is recognized by MAb 2D6 is indicated by a double underline.

6A-infected cells by MAbs 2D6, 2D4, 13D6, 3B5, and 2G3. Each of these MAbs immunoprecipitated several glycosylated polypeptides, and these glycoproteins were collectively designated the gp82-gp105 complex (5–7). Virion envelope association of this glycoprotein complex was demonstrated by immunoprecipitation of gp105, gp92, and gp82 from ^{125}I-labeled virions by using MAb 2D6, neutralization of virus infectivity by MAbs, and immune electron microscopy (5, 7). Isolation and characterization of the genomic clones 82G (34) and pB2.3 provided the evidence that gp82-gp105 may be encoded on a spliced mRNA. The complete ORF encoding the neutralizing epitope was confined to a 705-bp DNA in the p2.3 genomic clone. We reasoned that an ORF of this size was unlikely to encode a glycoprotein of 82 to 105 kDa and therefore may represent an exon of a larger gene. The detailed characterization of a 2.5-kb cDNA presented here shows that this cDNA encodes the gp105 protein in the HHV-6 gp82-gp105 complex and that this transcript is derived from complex multiple splicing.

We have previously reported the hybridization of a gp82-gp105 gene fragment (82G) to at least three RNA species in Northern blot analysis (34). Hybridization of the single-stranded RNA probe derived from the 2.5-kb cDNA with at least six separate RNA species from the total RNA or from poly(A)$^+$-selected RNA from infected cells suggests a complex pattern of transcription from the gp82-gp105 gene. Since the 2.5-kb cDNA is derived from the HHV-6 genomic region spanning more than 20 kb and since differential splicing appears to be involved in generating the proteins in the gp82-gp105 complex, it is not surprising that the single-stranded RNA probe derived from the cDNA hybridized with at least six messages from infected cells. Evidence such as the reactivity of MAbs with several polypeptides of about 105, 92, 82, 62, and 58 kDa in Western blot reactions indicating shared antigenic features,

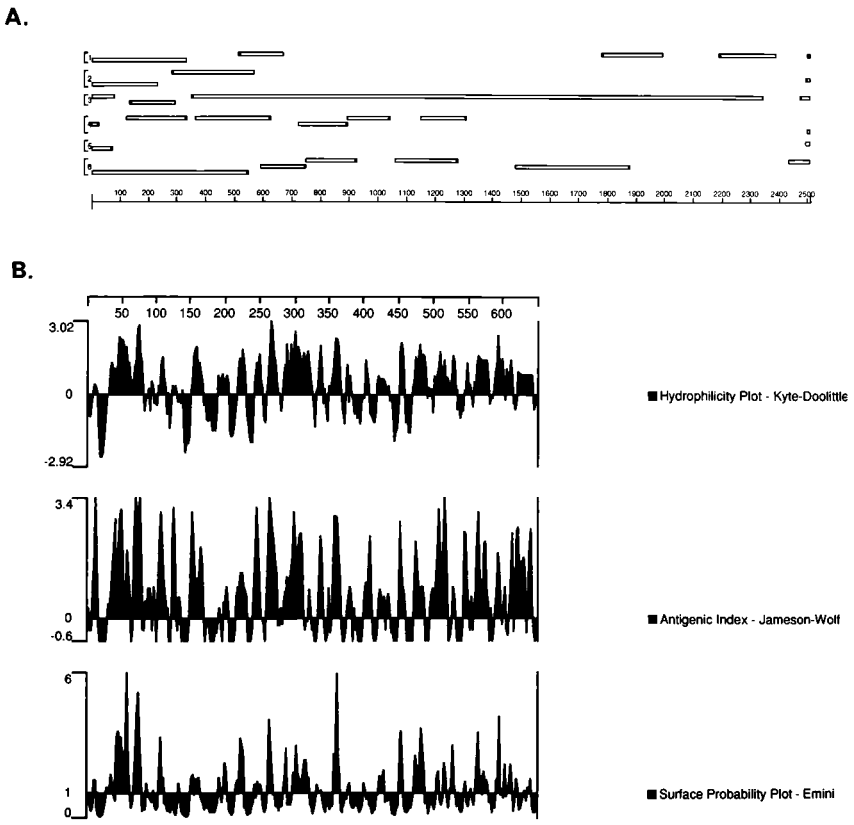

FIG. 8. ORF and polypeptide analysis of the cDNA insert. (A) ORF analysis of the cDNA sequence. ORFs 1 to 3 are oriented left to right, and ORFs 4 to 6 are oriented right to left. An approximate scale (in kilobases) is shown. All ORFs greater than 150 nucleotides in length are shown. (B) Analysis of the deduced polypeptide sequence from the cDNA ORF 3. The hydropathic analysis shows hydrophilicity as positive numbered peaks above the horizontal axis and hydrophobicity as negative numbered peaks below the horizontal axis. Antigenicity and surface probability of the deduced polypeptide regions are represented as peaks above the horizontal axis of the antigenic index and surface probability plots, respectively. An approximate scale indicating the amino acid position is shown.

V8-digested peptide maps showing closely related polypeptides forming this complex (7), and hybridization of several RNAs with a cDNA encoding the neutralizing epitope suggests that the components within the gp82-gp105 complex may be encoded on separate mRNAs originating from the same gene and are probably derived by multiple differential splicing. We have also isolated two other HHV-6(GS) cDNA clones of about 1.8 and 2.7 kb which also hybridized with the 82G probe containing the neutralizing epitope. Since the cDNA library was generated by oligo(dT) priming, most of the clones presumably represent the 3′ end of mRNAs. Preliminary PCRs using primers to amplify the 253-bp exon 1 leader sequence and with primers to amplify the 385-bp in the exon 3 region containing the neutralizing epitope amplified fragments of similar sizes from the 2.7- and 1.8-kb cDNA clones. These data indicate that the 2.5-, 2.7-, and 1.8-kb cDNAs share a common 5′ end and thus must have come from related transcripts. Studies investigating other cDNA clones encoded by this region of the genome including the isolated 1.8- and 2.5-kb cDNAs are under way. Examination of the *Sal*I-L sequences revealed a TATA element about 500 bp upstream from the 253-bp exon 1 leader sequence in the cDNAs (data not shown). Whether this region serves as the promoter element for the transcripts encoding the gp82-gp105 complex needs to be examined.

In our ongoing studies with HHV-6B(Z29), we have identified and sequenced 2.4- and 1.9-kb HHV-6B(Z29) cDNAs. Comparison with the HHV-6B(Z29) and HHV-6A genomic sequences demonstrated that the 1.9-kb cDNA is derived by multiple splicing of six introns and utilizes a polyadenylation sequence at the end of exon 7. The ORF encoded in the 2.4-kb cDNA is identical to the ORF encoded by the 1.9-kb cDNA

TABLE 1. Exon positions within the gene that encodes gp105 in the cDNA and HHV-6A(U1102) sequences.

Exon	Position (nucleotide no.)		% Identity
	cDNA	Genomic[a]	
1	1–252	3438–3684	95.2
2	253–378	3706–3831	91.3
3	379–802	4073–4494	95.3
4	803–1084	4596–4877	96.8
5	1085–1420	5286–5621	98.5
6	1421–1498	5739–5816	100.0
7	1499–1606	5913–6020	99.1
8	1607–1720	6112–6223	98.2
9	1721–1881	6310–6420	97.5
10	1882–1891	6980–6989	100.0
11	1892–2020	7140–7268	100.0
12	2021–2363[b]	7379–7725	99.7

[a] All positions refer to the complementary DNA strand of the HHV-6A(U1102) *Bam*HI G fragment except exon 1, in which case the first cDNA exon has identity with the complementary strand of the HHV-6A(U1102) *Sal*I L/O genomic DNA fragment at the indicated nucleotide positions.

[b] cDNA nucleotides 2364 to 2512 in exon 12 have no identity with any known herpesvirus DNA sequence and likely represent an artifact of cloning.

TABLE 2. Potential intron splice junctions of the cDNA gene that encodes gp105[a]

Intron	Donor site	Acceptor site
1	ACG/(GT)AGGT	TTATATTGCCGC(AG)/A
2	ACG/(GT)AC^{A}AG^{T}	CGAAATTTTCAC(AA^{G})/G
3	ACG/(GT)AAGG	T^{C}TAATTTATCGC(AG)/C
4	ATG/(GT)GAGT	TTTATTCACTCA(AG)/T
5	CGT/(GT)AAGT	TCGTTTTTTGTC(AG)/G
6	ACT/(GT)AAGT	TATGTTTCTAAC(AG)/A
7	ACG/(GT)GAGT	CCGGTTATGCAC(AG)/C
8	ATG/(GT)GAGC	TTCTTAATTTGC(AG)/C
9	AGG/(GT)GGGC	TTCGACCTGCCT(AG)/A
10	AAT/(GT)AAGT	TGATATTTGTTC(AG)/T
11	TAG/(GT)ATTA	GTTTTTTTTTTA(AG)/T
Consensus sequence	NAG/(GT)AAGT	TNTTTTTTTTNC(AG)/G

[a] The nucleotide sequence at each site was taken from the genomic sequence of HHV-6A(U1102). Nucleotides which were found to be different in HHV-6A(GS) at these sites are indicated as superscripts. Parentheses indicate nucleotides within the splice junctions which show 100% identity with the consensus splice sequence. N designates any nucleotide.

except that the 3′ end of the cDNA is longer than in the 1.9-kb cDNA and uses a polyadenylation signal 1.0 kb downstream with an additional exon near the polyadenylation site (exon 8). The splice donor acceptor sites and the lengths of the first six exons of the 1.9- and 2.4-kb cDNAs are identical. Comparison with the 2.5-kb HHV-6A(GS) cDNA encoding gp105 shows that the splice donor and acceptor sites and the lengths of the first six exons of the 1.9- and 2.4-kb HHV-6B cDNAs and exons 3 to 8 of the HHV-6A 2.5-kb cDNA are identical. There are additional exons in the HHV-6A cDNA sequences 3′ to exon 7, and the ORF in these exons is open and continues with the preceding exons. In contrast, even though the 2.4-kb HHV-6B cDNA ends at the same site as does the HHV-6A 2.5-kb cDNA, the ORF stops 1 kb upstream. Taken together, these data strongly support our notion that differential splicing may play a role in the generation of the various mRNA species that encode the proteins in the gp82-gp105 complex of HHV-6A and HHV-6B. Further studies are in progress to define the relationships of the various HHV-6A and HHV-6B cDNA clones encoding the proteins in the gp82-gp105 complex. These studies will also define the molecular basis for the variations of the gp82-gp105 complex between the two groups and the regulation of expression of genes encoding HHV-6A and HHV-6B gp82-gp105.

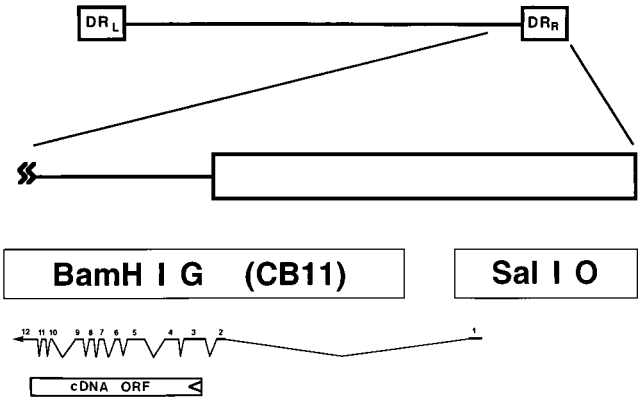

FIG. 9. The gene that encodes gp105. The HHV-6 genome is shown with emphasis on the right terminal region. The mapped locations of the HHV-6A(U1102) *Bam*HI G and *Sal*I O genomic fragments are shown. The thin irregular line below the genomic fragments represents the mapped location of the gene that encodes gp105. The arrowhead indicates the 5′-to-3′ orientation of the gp105 ORF. Exons of the gene are numbered, and introns are shown between the exons. The single long ORF begins within exon 3 and terminates in exon 12, as indicated by the labeled box.

The 2.5-kb cDNA sequence covers over 20 kb of HHV-6 genomic sequence, with intron 1 spanning about 8 kb of genomic DNA. There are examples in other herpesviruses for differential splicing of mRNA from a gene that encodes two or more separate yet related polypeptides. Two envelope glycoproteins of EBV are known to be encoded on RNAs which map to the same DNA fragment. The message for gp220 is derived from an internal, in-frame splicing event within the gp350 ORF (18, 26). Differential splicing of transcripts from the major immediate-early region of human CMV has also been reported (30, 35). However, no herpesvirus envelope glycoprotein, other than the gp105 reported here, has been reported to be encoded on an mRNA generated from multiple splicing. Comparisons of amino acid sequences show that all human herpesviruses have counterparts to HSV-1 glycoprotein gB, gH, and gL, suggesting that they may mediate some conserved functions. In addition, these viruses also possess other glycoproteins which may be responsible for their unique biological characters. For example, EBV has a restricted tropism and binds to CR2 receptor on B cells (18, 26). This specific binding is mediated by glycoproteins gp350 and gp220, for which no counterparts have been identified in other human herpesviruses. Even the conserved glycoproteins may serve different functions best suited for the particular virus, which is illustrated by the fact that unlike gB of HSV, varicella-zoster virus, and CMV, the gB of EBV is not detected on the virion envelopes or on the infected cell membranes (18, 26). Several features such as the multiple species of closely related polypeptides forming a virion envelope glycoprotein and highly spliced messages potentially encoding the polypeptides distinguish the HHV-6 gp82-gp105 complex from the glycoproteins of other human herpesviruses. Absence of hybridization with other herpesvirus DNAs and the absence of sequence homology with other herpesvirus genes indicate that the gp82-gp105 complex must be unique to HHV-6. The ability of MAbs to neutralize the infectivity of HHV-6 suggests that this complex must play a specific role in the biology of HHV-6; further studies are in progress to define these roles.

ACKNOWLEDGMENTS

This study was supported in part by Public Health Service grants AI 24224, AI 30356, AI 33502, and BRSG-SO7RR05373-28 to B.C. B.P. is a recipient of an M.D./Ph.D scholarship from the Life and Health Insurance Medical Research Fund.

REFERENCES

1. **Ablashi, D. V., H. Agut, Z. Berneman, G. Campadelli-Fiume, D. Carrigan, L. Cecerini-Nelli, B. Chandran, S. Chou, H. Collandre, R. Cone, T. Dambough, S. Dewhurst, D. DiLuca, L. Foa-Tomasi, B. Fleckenstein, N. Frenkel, R. Gallo, U. Gompels, C. Hall, M. Jones, G. Lawrence, M. Martin, L. Montagnier, F. Neipel, J. Nicholas, P. Pellett, A. Razzaque, G. Torrelli, B. Thomson, S. Salahuddin, L. Wyatt, and K. Yamanishi.** 1993. Human herpesvirus-6 strain groups: a nomenclature. Arch. Virol. **129:**363–366.
2. **Ablashi, D. V., N. Balachandran, S. F. Josephs, C. L. Hung, G. R. F. Krueger, B. Kramarsky, S. Z. Salahuddin, and R. C. Gallo.** 1991. Genomic polymorphism, growth properties and immunologic variations in human herpesvirus-6 (HHV-6) isolates. Virology **184:**545–542.
3. **Aubin, J.-T., H. Agut, H. Collandre, K. Yamanishi, B. Chandran, L. Montagnier, and J.-M. Huraux.** 1993. Antigenic and genetic differentiation of the two putative types of human herpes virus 6. J. Virol. Methods **41:**223–234.
4. **Aubin, J.-T., H. Collandre, D. Candotti, D. Ingrand, C. Rouzioux, M. Burgard, S. Richard, J.-M. Huraux, and H. Agut.** 1991. Several groups among human herpesvirus 6 strains can be distinguished by Southern blotting and polymerase chain reaction. J. Clin. Microbiol. **29:**367–372.

5. **Balachandran, N.** 1992. Proteins of human herpesvirus-6, p. 97–120. *In* D. V. Ablashi, G. R. F. Krueger, and S. Z. Salahuddin (ed.), Human herpesvirus-6: epidemiology, molecular biology, and clinical pathology, vol. 4, 1st ed. Elsevier Science Publishers, Amsterdam.
6. **Balachandran, N., R. E. Amelese, W. W. Zhou, and C. K. Chang.** 1989. Identification of proteins specific for human herpesvirus 6-infected human T cells. J. Virol. **63:**2835–2840.
7. **Chandran, B., and Q. Gao.** Characterization of envelope glycoprotein gp82-gp105 of HHV-6. Submitted for publication.
8. **Chandran, B., S. Tirawatnapong, B. Pfeiffer, and D. V. Ablashi.** 1992. Antigenic relationships among human herpesvirus-6 isolates. J. Med. Virol. **37:**247–254.
9. **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.
10. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
11. **Chou, S. W., and G. I. Marousek.** 1992. Homology of the envelope glycoprotein B of human herpesvirus-6 and cytomegalovirus. Virology **191:**523–528.
12. **Chou, S. W., and G. I. Marousek.** 1994. Analysis of interstrain variation in a putative immediate-early region of human herpesvirus 6 DNA and definition of variant-specific sequences. Virology **198:**370–376.
13. **Dewhurst, S., B. Chandran, K. McIntyre, K. Schnabel, and C. B. Hall.** 1992. Phenotypic and genetic polymorphisms among HHV-6 isolates from North American infants. Virology **190:**490–493.
14. **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.
15. **Ellinger, K., F. Neipel, L. Foa-Tomasi, G. Campadelli-Fiume, and B. Fleckenstein.** 1993. The glycoprotein B homologue of human herpesvirus 6. J. Gen. Virol. **74:**495–500.
16. **Qian, G., C. Wood, and B. Chandran.** 1993. Identification and characterization of glycoprotein gH of human herpesvirus-6. Virology **194:**380–386.
17. **Gompels, U. A., A. L. Carss, N. Sun, and J. Arrand.** 1992. Infectivity determinants encoded in a conserved gene block of human herpesvirus-6. DNA Sequence **3:**25–39.
18. **Gong, M., and E. Kieff.** 1990. Intracellular trafficking of two major Epstein-Barr virus glycoproteins, gp350/220 and gp110. J. Virol. **64:**1507–1516.
19. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
20. **Inoue, N., T. R. Dambaugh, and P. E. Pellett.** 1994. Review: molecular biology of human herpesviruses 6A and 6B. Infect. Agents Dis. **2:**343–360.
21. **Jarrett, R. F., A. Gallagher, S. Gledhill, M. D. Jones, I. Teo, and B. E. Griffin.** 1989. Variation in restriction map of HHV-6 genome. Lancet **i:**448–449.
22. **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.
23. **Josephs, S. F., D. V. Ablashi, S. Z. Salahuddin, B. Krmarsky, B. R. Franza, Jr., P. Pellet, A. Buchbinder, S. Memon, F. Wong-Staal, and R. C. Gallo.** 1988. Molecular studies of HHV-6. J. Virol. Methods **21:**179–190.
24. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15:**8125–8148.
25. **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.
26. **Liebowitz, D., and E. Kieff.** 1993. Epstein-Barr virus, p. 107–172. *In* B. Roizman, R. J. Whitley, and C. Lopez (ed.), The human herpesviruses. Raven Press, Ltd., New York.
27. **Lindquester, G. J., and P. E. Pellett.** 1991. Properties of the human herpesvirus-6 strain Z29 genome: G+C content, length, and presence of variable-length directly repeated terminal sequence elements. Virology **182:**102–110.
28. **Liu, D. X., U. A. Gompels, J. Nicholas, and C. Lelliott.** 1993. Identification and expression of the human herpesvirus 6 glycoprotein H and interaction with an accessory 40K glycoprotein. J. Gen. Virol. **74:**1847–1857.
29. **Martin, M. E., 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.
30. **Mocarski, S. D.** 1993. Cytomegalovirus biology and replication, p. 173–226. *In* B. Roizman, R. J. Whitley, and C. Lopez (ed.), The human herpesviruses. Raven Press, Ltd., New York.
31. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10:**459–472.
32. **Neipel, F., K. Ellinger, and B. Fleckenstein.** 1991. The unique region of the human herpesvirus type 6 genome is essentially collinear to the UL segment of human cytomegalovirus. J. Gen. Virol. **72:**2293–2297.
33. **Pellett, P. E., J. B. Black, and M. Yamamoto.** 1992. Human herpesvirus 6: the virus and the search for its role as a human pathogen. Adv. Virus Res. **41:**1–52.
34. **Pfeiffer, B., Z. N. Berneman, F. Neipel, C. K. Chang, S. Tirwatnapong, and B. Chandran.** 1993. Identification and mapping of the gene encoding the glycoprotein complex gp82-gp105 of human herpesvirus 6 and mapping of the neutralizing epitope recognized by monoclonal antibodies. J. Virol. **67:**4611–4620.
35. **Rawlinson, W. D., and B. G. Barrell.** 1993. Spliced transcripts of human cytomegalovirus. J. Virol. **67:**5502–5513.
36. **Roizman, B., and A. E. Sears.** 1993. Herpes simplex viruses and their replication, p. 11–68. *In* B. Roizman, R. J. Whitley, and C. Lopez (ed.), The human herpesviruses. Raven Press, Ltd., New York.
37. **Salahuddin, S. Z., D. V. Ablashi, P. D. Markham, S. F. Josephs, S. Sturzenegger, M. Kaplan, G. Halligan, P. Biberfeld, F. Wong-Staal, B. Kramarsky, and R. C. Gallo.** 1986. Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. Science **234:**596–601.
38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
39. **Senapathy, P., M. B. Shapiro, and N. L. Harris.** 1990. Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Methods Enzymol. **183:**252–278.
40. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
41. **Suga, S., T. Yoshikawa, Y. Asano, T. Yazaki, and S. Hirata.** 1989. Human herpesvirus-6 infection (exanthem subitum) without rash. Pediatrics **83:**1003–1006.
42. **Thompson, J., S. Choudhury, F. Kashanchi, J. Doniger, Z. Berneman, N. Frenkel, and L. J. Rosenthal.** 1994. A transforming fragment within the direct repeat region of human herpesvirus type 6 that transactivates HIV-1. Oncogene **9:**1167–1175.
43. **Thomson, B. J., S. Dewhurst, and D. Gray.** 1994. Structure and heterogeneity of the *a* sequences of human herpesvirus 6 strain variants U1102 and Z29 and identification of the human telomeric repeat sequences at the genomic termini. J. Virol. **68:**3007–3014.
44. **Thomson, B. J., and R. W. Honess.** 1992. The right end of the unique region of the genome of human herpesvirus 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.
45. **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.