JOURNAL OF VIROLOGY, Sept. 1989, p. 3792–3800
0022-538X/89/093792-09$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 63, No. 9

Identification and Characterization of a Human Cytomegalovirus Gene Coding for a Membrane Protein That Is Conserved among Human Herpesviruses

ROBERT LEHNER, HEIDI MEYER, AND MICHAEL MACH*

Institut für Klinische und Molekulare Virologie der Universität Erlangen-Nürnberg, Schlossgarten 4, 8520 Erlangen, Federal Republic of Germany

Received 21 February 1989/Accepted 18 May 1989

A rabbit antiserum was raised against envelope material from purified human cytomegalovirus strain AD169. The serum recognized polypeptides 200, 170, 160, 75, 58, and 45 kilodaltons in size. It was used to screen a cDNA library constructed from poly(A)$^+$ RNA from human cytomegalovirus-infected cells in the expression vector lambda gt11. A recombinant bacteriophage expressing cytomegalovirus-specific sequences was identified, and the corresponding gene was mapped to the *Hin*dIII R fragment. The gene is transcribed into a late 1.5-kilobase RNA. The nucleotide sequence of the coding region was determined. Computer analysis of the gene product revealed a polypeptide containing multiple potential membrane-spanning domains, representing a type of protein not identified in the envelope of herpesviruses before. The protein shows homology on the amino acid level to hypothetical proteins from reading frames BBRF3 of Epstein-Barr virus, UL10 of herpes simplex virus type 1, and ORF50 of varicella-zoster virus. By using an antiserum raised against procaryote-expressed parts of the cytomegalovirus membrane protein, a 45-kilodalton structural component of the virus was identified as the gene product.

Human cytomegalovirus (HCMV) has been associated with a wide variety of clinical syndromes. For example, it accounts for a significant number of developmentally disabled children (3). It also causes serious and often life-threatening complications in patients undergoing allograft transplantation (47). Additionally, it has been recognized as an important cause of secondary infection in patients with the acquired immune deficiency syndrome (54). Little is known about the immunological responses that confer protection, but most likely both the cellular and the humoral immune responses play important roles in limiting the consequences of HCMV infection. As in the other human herpesvirus infections, envelope proteins constitute important targets for this response. However, the HCMV antigens involved have not been completely characterized. So far, between three and eight electrophoretically separable glycosylated envelope proteins have been identified (5, 14, 20, 22, 31, 49, 50, 55, 56). In addition, three families of glycoprotein complexes have been described (24, 30). Of these polypeptides, proteins of 86 (51), 65 (8), 58 (7), and 47 to 52 kilodaltons (kDa) (23) are able to induce neutralizing antibodies in laboratory animals and may therefore represent targets of a protective humoral immune response.

Glycoproteins and their topology in the viral membrane have been investigated in detail in a variety of viral systems, such as the G protein of vesicular stomatitis virus (1), the hemagglutinin of influenza virus (61), and glycoprotein B of herpes simplex virus type 1 (HSV-1) (12). In general, these polypeptides carry the majority of the amino acid residues on the outside of the membrane; a large amount of sugar residues increases the molecular weight significantly. Little is known about the existence or the function of virus-encoded membrane proteins with a topology like that of bacteriorhodopsin, i.e., harboring multiple hydrophobic membrane-spanning domains and no significant glycosylation. The BNLF1 reading frame product of the Epstein-Barr virus (EBV), which has been termed LMP, contains six hydrophobic stretches separated by shorter hydrophilic motifs, suggesting multiple membrane-spanning domains (38, 48). LMP plays an important role in the immune response of EBV, since it is recognized by cytotoxic T cells and is most probably essential for the maintenance of the latent stage of the infection (59). However, it is expressed only early in the course of infection and has not been identified as a structural constituent of the virus. In HSV-1, major fusion-inducing mutations have been mapped to the UL53 gene (17, 44). The gene product has not been studied experimentally. However, predicted secondary structures of the hypothetical protein revealed several hydrophobic regions having characteristics normally associated with membrane proteins (17).

As a step in the characterization of membrane proteins of HCMV, we have generated an antiserum against detergent-solubilized envelope preparations. The serum was used to screen a cDNA library constructed in the expression vector lambda gt11. An HCMV gene was identified which codes for a 45-kDa structural protein containing multiple membrane-spanning domains. We also show that this protein has homologous counterparts in other human herpesviruses.

MATERIALS AND METHODS

Virus purification and preparation of envelope material. Human foreskin fibroblasts (HFF cells) were infected with the HCMV strain AD169 at a multiplicity of infection of 1 to 5 PFU per cell. Virus material was pelleted from tissue culture supernatant 6 days after infection (1 h, 24,000 rpm, 4°C) by centrifugation in a Beckman SW27 rotor. The virus particles were suspended in 0.04 M sodium phosphate buffer (pH 7.4) and purified by centrifugation in a glycerol-tartrate gradient by the method of Talbot and Almeida (57). The fractions containing virions and noninfectious enveloped particles (NIEPs) were combined. Dense bodies were handled separately. HCMV particles were collected from tissue

* Corresponding author.

culture supernatants by centrifugation (1 h, 24,000 rpm, 4°C) in a Beckman SW27 rotor. Envelope material was solubilized by treatment with Nonidet P-40 (final concentration, 2%) and sodium deoxycholate (final concentration, 0.5%) for 1 h at 4°C. Insoluble material was removed by centrifugation in a Sorvall SM24 rotor (1 h, 16,000 rpm, 4°C).

Production of polyspecific antisera. Bacterium-derived fusion proteins were separated on preparative polyacrylamide gels, and the respective bands were cut out after being stained with Coomassie brilliant blue. The proteins were extracted from the gel matrix with 20 mM ammonium carbonate (pH 9.5). Antisera were raised against detergent-solubilized membrane components of HCMV as well as against bacterial fusion proteins. For a single injection, 50 to 60 μg of total protein was used. Polypeptides were denatured with sodium dodecyl sulfate (0.2%), heated to 65°C for 5 min, and mixed with an equal volume of Freund adjuvant by sonication. Complete adjuvant was used for the initial injection, and incomplete adjuvant was used for the booster injections. Initially, rabbits were injected subcutaneously at six to eight sites on the back; booster injections were given intramuscularly at 2- to 3-week intervals beginning 4 weeks after the first immunization.

Protein gel electrophoresis and Western immunoblot analysis. Separation of proteins on polyacrylamide gels, transfer to nitrocellulose membranes, and detection by specific antisera were performed by the methods of Laemmli (35) and Burnette (10).

Construction and screening of the cDNA library. The construction and screening of the lambda gt11 library have been described in detail by Jahn et al. (27), Mach et al. (39), and Young and Davis (64).

DNA cloning and Southern blot analysis. All DNA cloning and hybridization techniques were performed by the methods of Maniatis et al. (40).

Radiolabeling of specific DNA fragments. The [α-^{32}P]dATP-labeled DNA probes were produced by second-strand synthesis on a single-stranded DNA template (25). Single-stranded DNA from recombinant M13mp19 was prepared as described previously (53, 63). One microgram of DNA was hybridized with 5 ng of universal primer (17-mer; Boehringer GmbH, Mannheim, Federal Republic of Germany) by heating for 1 h to 60°C. The reaction mixture was allowed to cool to room temperature, and DNA synthesis was initiated by the addition of dCTP, dGTP, and dTTP (final concentration of each, 0.2 mM); 30 μCi of [α-^{32}P]dATP; and 2 to 3 units of Klenow polymerase in the presence of 0.1 mM $MgCl_2$ and 10 nM dithiothreitol. After 25 min at room temperature, the reaction was stopped by adding EDTA, and the DNA was separated from nonincorporated nucleotides by centrifugation through a Sephadex G-50 column (40).

DNA sequencing. The chain termination method of Sanger et al. using didesoxynucleotides (52) was employed for DNA sequencing. A T7 polymerase sequencing kit purchased from Pharmacia (Freiburg, Federal Republic of Germany) was used, and the manufacturer's instructions were followed precisely. The nucleotide sequence was determined by using two plasmids (pIMP-SX, containing a 580-base-pair (bp) *Xba*I-*Sph*I fragment, and pIMP520, containing a 520-bp *Pvu*II-*Sma*I fragment). The plasmids represent the 5′ and 3′ ends of the 1.7-kilobase (kb) *Xba*I-*Sma*I genomic fragment, respectively. In addition, nine synthetic oligonucleotides (17-mers) from various areas of the entire *Xba*I-*Sma*I fragment were used to complete the nucleotide sequence. The nucleotides were prepared with a Cyclone DNA synthesizer (Biosearch, Inc., San Rafael, Calif.) by using the phosphoamidite method (41). Both strands of the DNA were sequenced at least three times.

Bacterial expression of fusion proteins. A part of the coding region identified by the recombinant bacteriophage 92env was expressed as a β-galactosidase fusion protein in the bacterial vector system pBD2IC20H (9). This plasmid contains multiple cloning sites following the gene coding for the first 375 amino acids of *Escherichia coli* β-galactosidase. A 520-bp *Pvu*I-*Sma*I HCMV fragment was inserted into the *Eco*RV restriction site of the multilinker. Recombinant plasmids were transfected into *E. coli* BMH7118. Induction of fusion proteins was initiated by the addition of isopropyl-β-D-thiogalactopyranoside (final concentration, 1 mM) to bacterial cultures when the cell density (A_{600}) was approximately 1.0. Synthesis of fusion proteins was allowed to proceed for 3 to 4 h. Cells were harvested by centrifugation and lysed in sodium dodecyl sulfate gel sample buffer, and protein extracts were analyzed by polyacrylamide gel electrophoresis and Western blot analysis.

RNA extraction and Northern (RNA) blot analysis. RNA preparation and Northern blotting were performed as described previously (46).

Computer analyses. The computer analyses of the nucleotide and amino acid sequences were done on a VAX11/788 computer by using the software package UWGCG (University of Wisconsin Genetics Computer Group) (18). The hydropathy analyses (34) were done with the program "Peptide-structure," with a sliding window set to 11 amino acid residues. For calculating homology of proteins, the program "Gaps" was employed (gap weight, 3.0; length weight, 0.15).

RESULTS

Characterization of the HCMV envelope preparation. Tissue culture supernatant from HCMV strain AD169-infected fibroblasts contains three different forms of particles: infectious virions, dense bodies, and NIEPs. With the exception of the 36-kDa assembly protein, which is a constituent of NIEPs only, virions and NIEPs contain the same structural proteins (26). The dense bodies represent enveloped particles that do not contain a capsid (2, 29, 42) and consist mainly of the 65-kDa phosphorylated matrix protein (pp65) (26). The different particle forms can be separated in a positive-density negative-viscosity glycerol-tartrate gradient (57).

Initially, we used three particle fractions for the preparation of envelope material. One contained the virions plus NIEPs (V/N fraction), the second contained the dense bodies, and the third contained all three particle forms obtained from tissue culture supernatant by high-speed centrifugation (CMV fraction). The particles were extracted with detergent, and the insoluble material was removed by centrifugation. Both the soluble and the insoluble materials were analyzed on polyacrylamide gels and Western blots. Treatment of the different particle forms with detergent resulted in soluble fractions which were devoid of abundant matrix proteins, such as pp65 or pp150, as judged from polyacrylamide gel electrophoresis followed by silver staining (data not shown). In order to characterize the distribution of proteins between the fractions more precisely, we carried out Western blot analyses with various antibodies. A monoclonal antibody specific for the 150-kDa major capsid protein (a kind gift from W. Britt, the University of Alabama at Birmingham, Birmingham, Ala.) recognized polypeptides

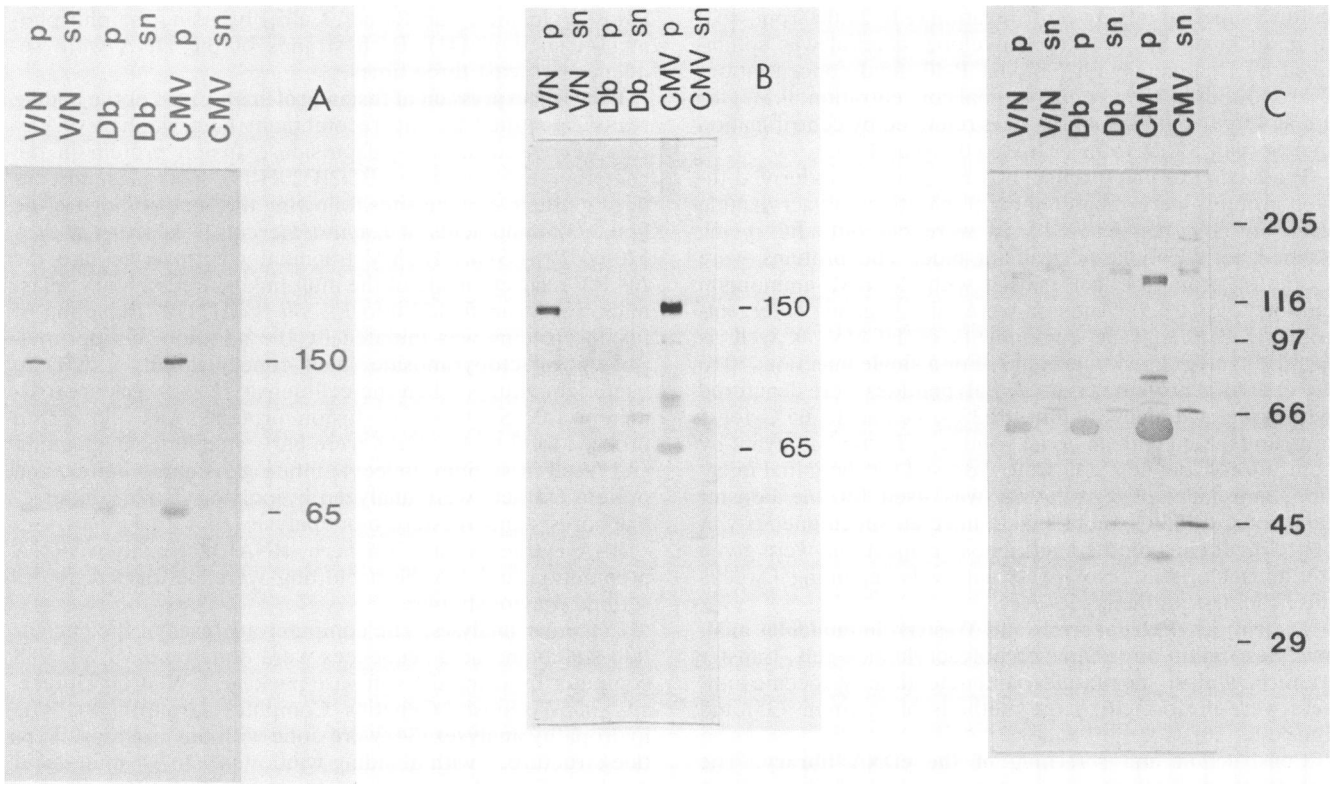

FIG. 1. Western blot analysis of detergent-soluble (sn) and insoluble (p) HCMV proteins with various antisera. Polypeptides of gradient-purified virions plus NIEPs (V/N), dense bodies (Db), and HCMV particles (CMV) were separated on a 10% polyacrylamide gel, transferred to nitrocellulose, and incubated with a monoclonal antibody (diluted 1:100) against the 150-kDa major capsid protein (A), a polyvalent antiserum (diluted 1:250) against the 150-kDa basic phosphoprotein (B), and a serum raised against NIEP proteins (diluted 1:500) (C). Protein A coupled to horseradish peroxidase and 1,4-chloronaphthol staining were used as a detection system.

in the insoluble V/N and CMV particle fractions (Fig. 1A). The absence of the capsid protein in the soluble virions/NIEPs and CMV particle fractions indicated that the capsids were left intact by the extraction procedure. No reaction was observed against either dense body fraction. The absence of capsid proteins in dense bodies is consistent with the observation that this form of particle does not contain capsids. The reactivity of the antibody with the 65-kDa protein was considered nonspecific since it can sometimes also be observed when unrelated monoclonal antibodies or monospecific sera are used in Western blots with HCMV antigen. In addition, proteins of 130 and 65 kDa have been shown to contain Fc-receptor-binding activity of HCMV (62). By using a polyclonal rabbit antiserum directed against the 150-kDa basic phosphoprotein (27), a polypeptide believed to be located within the matrix of the virus (22), polypeptides were detected only in the insoluble fractions (Fig. 1B). No reactive polypeptide was detectable in the soluble V/N and CMV particle fractions, indicating that no significant leakage of soluble proteins from the matrix had occurred. An antiserum raised against NIEPs showed that the polypeptides in the different fractions were mutually exclusive (Fig. 1C). The 65-kDa phosphoprotein, for example, a polypeptide localized in the matrix of the virus (37), was not found in the soluble fractions. On the other hand, the antiserum detected a polypeptide of 160 kDa only in the soluble fractions. This protein does not represent the major capsid protein or the basic phosphoprotein since these proteins cannot be separated on standard 10% polyacrylamide gels (28) and in addition are not present in the soluble fraction (Fig. 1A and B). Altogether, the polyvalent NIEP antiserum recognized in the detergent-soluble fractions predominantly proteins with approximate molecular sizes of 200, 170, 160, 70, and 45 kDa.

The presence of a characterized envelope protein in the soluble fraction was tested with a monospecific rabbit antiserum against the major glycoprotein gp58 (39) (Fig. 2). This protein has been shown to be located in the envelope of the virus (36). In the dense body preparation, gp58 was detectable in both the soluble and, to a lesser extent, insoluble material, whereas in the V/N preparation the gp58 was present predominantly in the soluble fraction. The distribution pattern of proteins localized either in the envelope or in inner structures of the virus was taken as evidence that the extraction method was capable of separating envelope proteins from the remainder of the virus. Because of the cleaner separation obtained with the V/N preparation, we chose these extracts for the generation of an antiserum specific for HCMV envelope proteins.

Generation of an anti-envelope serum. A rabbit was immunized with envelope preparations containing 50 to 60 μg of total protein. After three booster injections, the reactivity of the antiserum was tested in Western blots by using extracellular HCMV particles as the antigen. No reactivity against HCMV polypeptides was observed with the preimmune serum (data not shown). The immune serum reacted predominantly with polypeptides of 200, 170, 160, 75, 58, and 45 kDa (Fig. 3). In addition, a ladder of polypeptides in the molecular size range of 160 to 58 kDa was detected by the antiserum. This ladder was observed consistently with dif-

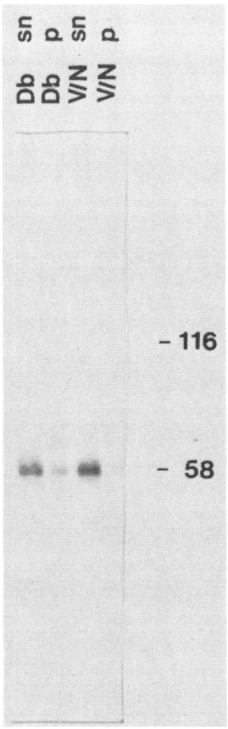

FIG. 2. Immunological characterization of the envelope preparations. The insoluble (p) and soluble (sn) fractions of the detergent-extracted material from virions plus NIEPs (V/N) and dense bodies (Db) were subjected to electrophoresis on 10% polyacrylamide gels. After electroblotting onto nitrocellulose membranes, the proteins were assayed with an antiserum raised against the major glycoprotein gp58 (diluted 1:100). Molecular sizes are given in kilodaltons.

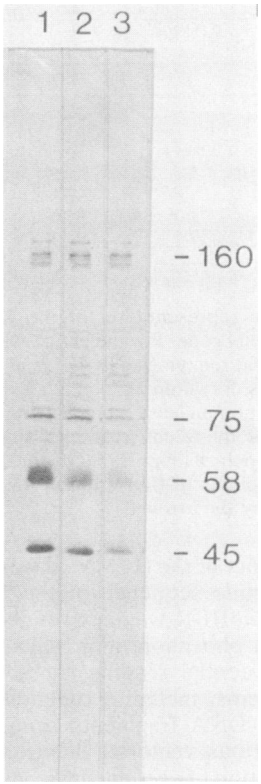

FIG. 3. Specificity of the anti-envelope serum. Proteins from HCMV particles were separated on 10% polyacrylamide gels, transferred to nitrocellulose membranes, and incubated with different dilutions of anti-envelope serum. Lanes: 1, 1:500 dilution; 2, 1:1,000 dilution; 3, 1:2,000 dilution. All sizes are given in kilodaltons.

ferent preparations of virions, dense bodies, and NIEPs. We do not have an explanation for this phenomenon; however, we cannot rule out the possibility that it represents degradation products of a high-molecular-weight protein. The recognition pattern of the anti-envelope serum is similar to that seen in Western blots of the detergent-soluble fractions with the anti-NIEP serum (Fig. 1C); i.e., polypeptides of 200, 170, 160, 75, and 45 kDa did react. In addition, the anti-envelope serum strongly recognized the major glycoprotein gp58. An identical recognition pattern was observed when the soluble fraction was used as the antigen. No reaction was seen with proteins in the insoluble fraction (data not shown).

Isolation of a lambda gt11 recombinant expressing an HCMV envelope protein. A cDNA library constructed from RNA of HCMV-infected fibroblasts was screened with the anti-envelope serum. The library has been used before to isolate recombinants for various other HCMV structural polypeptides (27, 39, 46). A total of 400,000 recombinant gt11 phage were tested. Eight positive signals were obtained; one recombinant (designated 92env) was purified and characterized in detail. *E. coli* Y1089 cells were infected with 92env phage, and the synthesis of fusion proteins was induced by the addition of isopropyl-β-D-thiogalactopyranoside. In cells carrying the 92env phage, an abundant polypeptide that was slightly larger than β-galactosidase was detectable upon polyacrylamide gel electrophoresis of protein extracts (data not shown). In cells infected with lambda gt11 phage, a 116-kDa protein corresponding to β-galactosidase was synthesized. The 92env fusion protein was recognized specifically by the anti-envelope serum in Western blots. Neither polypeptides from lambda gt11-infected cells nor fusion proteins from unrelated recombinants reacted, indicating that 92env was expressing HCMV-specific sequences (data not shown). To further substantiate this, DNA from 92env was prepared and the cDNA insert of approximately 210 bp was recloned into M13mp19. The resulting recombinant was designated m92env. DNA from this clone was radiolabeled with [α-^{32}P]dATP by second-strand synthesis and hybridized in a dot-spot assay to cosmid clones representing the whole HCMV genome (21). The cosmid pCM1058, containing the *Hin*dIII fragments T, R, and E, hybridized to the cDNA. In a more detailed Southern blot analysis, the m92env complementary sequences were confined to a 1,700-bp fragment at the center of *Hin*dIII-R (Fig. 4), between the *Sma*I site at 2.0 kb and the *Xba*I site at 3.7 kb. The nucleotide sequence of the m92env cDNA was determined. It is composed of 144 nucleotides of HCMV-specific DNA plus a stretch of 70 to 80 AT base pairs, most probably representing the poly(A) tail of the respective mRNA. A stop codon (TAA) was found 28 nucleotides downstream of the β-galactosidase HCMV fusion point. The fact that only nine amino acids of HCMV-specific sequences were expressed in 92env explains the small difference in size between the fusion protein and β-galactosidase produced from lambda gt11-infected cells. It is also an indication of high antigenicity of this part of the molecule, since it was sufficient to be recognized by the anti-envelope serum. Of the nine HCMV-specific amino acids, seven are hydrophilic (see Fig. 5).

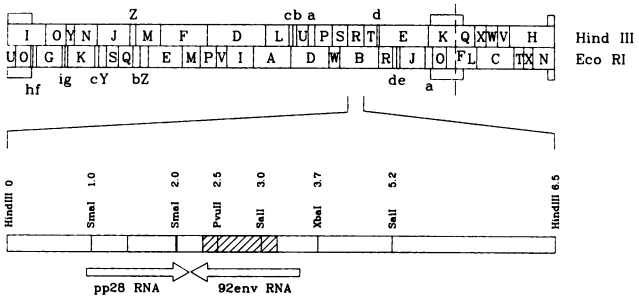

FIG. 4. Schematic representation of the HCMV genome and localization of the 92env gene. For the entire genome, the restriction sites are given for the enzymes *Hin*dIII and *Eco*RI. The 6.5-kb *Hin*dIII R fragment is drawn on an expanded scale, including a few relevant restriction endonuclease sites. The hatched area indicates the coding region for the 92env reading frame; the stippled area corresponds to the open reading frame of the 28-kDa phosphoprotein (pp28). The directions of transcription for the pp28 and 92env genes are indicated by the arrows.

Nucleotide sequence of the HCMV genomic region encoding 92env. The nucleotide sequence of the 1.7-kb *Xba*I-*Sma*I segment within *Hin*dIII-R was determined (Fig. 5). During our work with the phosphoprotein pp28, which is located adjacent to the potential reading frame (Fig. 4), we had encountered problems, including deletions and recombinations, when small DNA fragments from this region were subcloned into various vectors. Therefore, we chose as a strategy a combination of sequencing available plasmids and double-stranded sequencing with synthetic primers. The 1.7-kb *Sma*I-*Xba*I fragment contained an open reading frame of 1,116 bp. At the 5′ end, between nucleotides 1355 and 1382, this frame showed 100% homology with the cDNA. The stretch of adenosine residues that was present in the cDNA sequence was not present in the genomic DNA. This indicates that the ATTAAA motif between nucleotides 1474 and 1479 in fact represents the poly(A) signal for the 92env reading frame (6). According to the cDNA sequence, the poly(A) tail is added to the RNA at position 1599. Interestingly, the poly(A) signal for the 28-kDa phosphoprotein is located at the identical position, but on the opposite DNA strand (Fig. 5). The AUG codon at position 267 follows the known rules for effective translational start sites (33). There is no convincing sequence motif for TATA and CAAT boxes, respectively, within 300 bp upstream of that AUG. The primary translation product of the 92env frame is a protein 372 amino acids in length with a calculated molecular size of 42,860 Da. It should be mentioned that a second open reading frame is located within this DNA fragment. Starting at nucleotide 557 and ending at nucleotide 1243, this frame has the capacity to code for a protein 27 kDa in size. It is, however, most probably not translated into a polypeptide since computer analyses of the codon frequency and rare codon usage showed that it is not representative of HCMV genes (data not shown).

Identification of the 45-kDa structural protein as the translation product of the 92env reading frame. In order to determine the viral protein encoded by the 92env reading frame, a monospecific antiserum against a procaryote-expressed part of the polypeptide was generated. A 520-bp *Pvu*II-*Sma*I fragment (Fig. 4) encoding the last 72 amino acids of the 92env frame was inserted into the vector pBD2IC20H. In this plasmid, a fusion protein can be synthesized that contains 375 amino acids of β-galactosidase and the foreign sequence. The insertion of the HCMV DNA

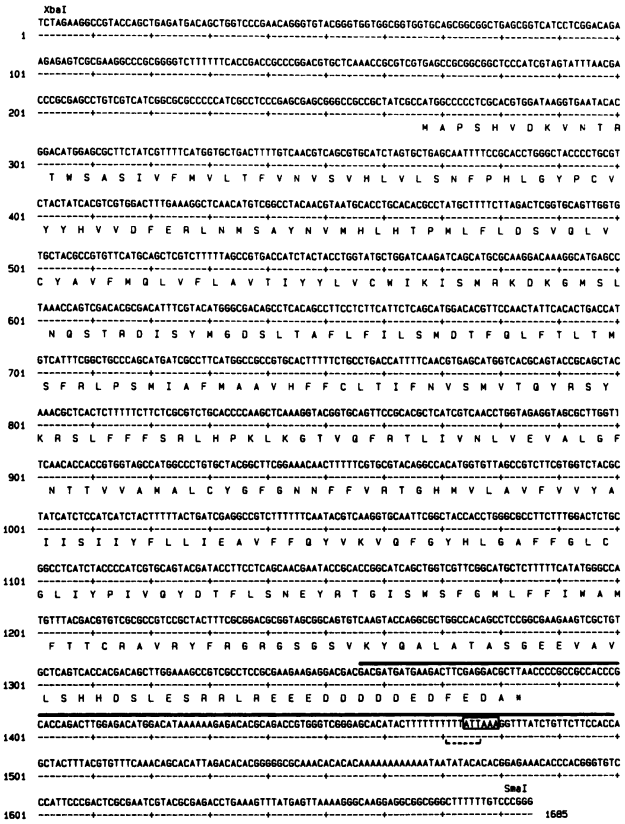

FIG. 5. Nucleotide sequence of the 1.7-kb *Xba*I-*Sma*I fragment coding for the 92env protein. The area of homology between the cDNA and the genomic nucleotide sequence is indicated by the boldface line (nucleotides 1356 to 1499). The boxed nucleotides ATTAAA represent the poly(A) signal of the 92env gene; the dashed line indicates the AATAAA poly(A) signal of the 28-kDa phosphoprotein (pp28) on the complementary strand. The amino acid sequence of the 92env translation product is given in the one-letter code below the nucleotide sequence.

led to the formation of a polypeptide which was recognized specifically by the anti-envelope serum (data not shown). The fusion protein was used to immunize rabbits. After four booster injections, the specificities of a preimmune serum and the immune sera were tested in Western blots with extracellular HCMV particles as the antigen. The immune serum specifically recognized a viral polypeptide 45 kDa in size (Fig. 6). The weaker reactions with polypeptides of 65 and 30 kDa were considered nonspecific since they were also observed with the preimmune serum. This result confirms that the 45-kDa structural protein reacting with the anti-envelope serum is encoded by the 92env reading frame.

Analysis of the 92env transcript and time of appearance. The size of the 92env-specific transcript was determined in Northern blot analyses. Total cellular RNA from AD169-infected fibroblasts was isolated, separated on 1.5% denaturing agarose gels, and hybridized with [α-^{32}P]dATP-labeled DNA fragments spanning the entire *Hin*dIII-R region. Two major transcripts of 1.3 and 1.5 kb were observed in RNA preparations isolated late in the course of the infection. As previously shown, the 1.3-kb RNA, detected by the probes pHM2 and pHM9, respectively, corresponds to the pp28 transcript (46). The probe pHM9 hybridized to both the 1.3- and the 1.5-kb RNAs (Fig. 7). The plasmid pHM10

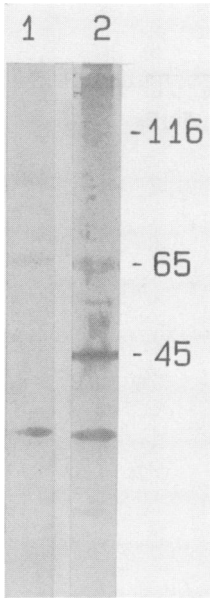

FIG. 6. Confirmation of the 45-kDa envelope protein as the translation product of the 92env reading frame. A rabbit antiserum raised against a recombinant protein containing the last 72 amino acids of the 92env gene product was analyzed in a Western blot with HCMV particles as antigens. Lanes: 1, preimmune serum (diluted 1:100); 2, anti-92env serum (diluted 1:200). Sizes are given in kilodaltons.

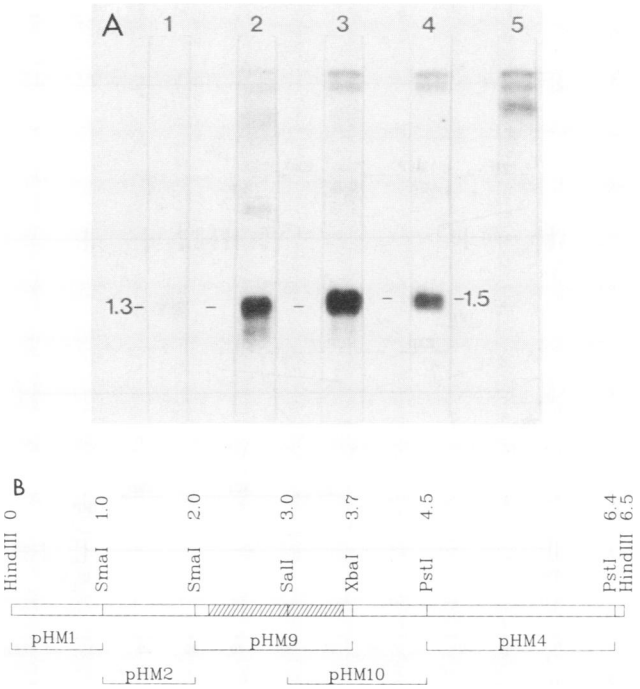

FIG. 7. The 92env reading frame is transcribed into a 1.5-kb late RNA. (A) Total RNA was isolated from HCMV-infected fibroblasts 72 h after infection. The RNA was separated on a 1.5% agarose gel, transferred to nitrocellulose membranes, and hybridized with various plasmids representing the entire *Hin*dIII R fragment. Lanes 1 to 5 correspond to the plasmids pHM1, pHM2, pHM9, pHM10, and pHM4, respectively. (B) The 92env open reading frame is indicated by the hatched area. Sizes are given in kilobases.

detected only the 1.5-kb RNA. Taken together with the nucleotide sequence data of the 92env cDNA, this result suggests that the pp28 RNA and the 92env transcript, respectively, terminate in the same area between nucleotides 1400 and 1500 and that the 92env reading frame is transcribed into an abundant 1.5-kb RNA late in the course of the infection. Northern blot analyses with strand-specific probes confirmed that both transcripts have an opposite orientation (data not shown). No such signal was observed with RNA preparations from immediate-early or early times of infection (data not shown).

Computer analysis of the 92env protein and comparison with homologous polypeptides of other human herpesviruses. Hydrophobicity analysis by the method of Kyte and Doolittle (34) revealed stretches of hydrophilic amino acids at both termini of the protein. At the amino terminus, the first 12 amino acids are mainly hydrophilic, and the carboxy-terminal part consists of approximately 40 highly hydrophilic residues. Between these hydrophilic regions, the 92env protein contains several clusters of hydrophobic amino acids framed by short stretches of hydrophilic residues. Eight hydrophobic domains of the 92env protein consist of at least 20 mainly nonpolar amino acids. Twenty amino acids are thought to be sufficient to span the 3-nm thickness of the lipid bilayer (2, 58). The overall structure resembles that of an intrinsic membrane protein, containing multiple membrane-spanning domains. Five potential N-glycosylation sites of the form N-X-T/S were identified at amino acids 29, 58, 115, 170, and 215; three sites are located within predicted transmembrane domains.

The translation product of the 92env reading frame was compared with proteins of other human herpesviruses by using the UWGCG program Gaps. Homologous counterparts were identified in all human herpesviruses whose nucleotide sequences have been completed (4, 15, 44). The greatest homology was observed with the BBRF3 gene product of EBV (23%), followed by UL10 of HSV-1 (21%) and ORF50 of varicella-zoster virus (VZV) (19%). Although the sequences are not highly conserved, they have a number of features in common: (i) a short hydrophilic N-terminal region of approximately 20 amino acids and a larger hydrophilic carboxy-terminal part (25 to 120 amino acids); (ii) eight stretches of hydrophobic sequences, each one long enough to cross the membrane, interrupted by short hydrophilic motifs; (iii) a conserved arrangement with adjacent genes in EBV, VZV, and HSV-1 (see Discussion).

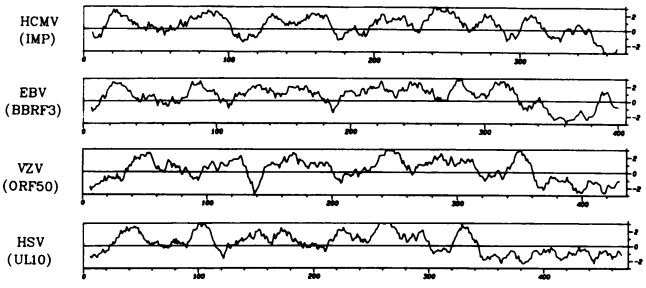

FIG. 8. Hydropathicity plot of the 92env protein of HCMV and the homologous reading frames in EBV (BBRF3), VZV (ORF50), and HSV-1 (UL10). The algorithm of Kyte and Doolittle (34) was used, with a sliding window set to 11 amino acid residues. The scales below the plots give the positions of the respective amino acid sequences.

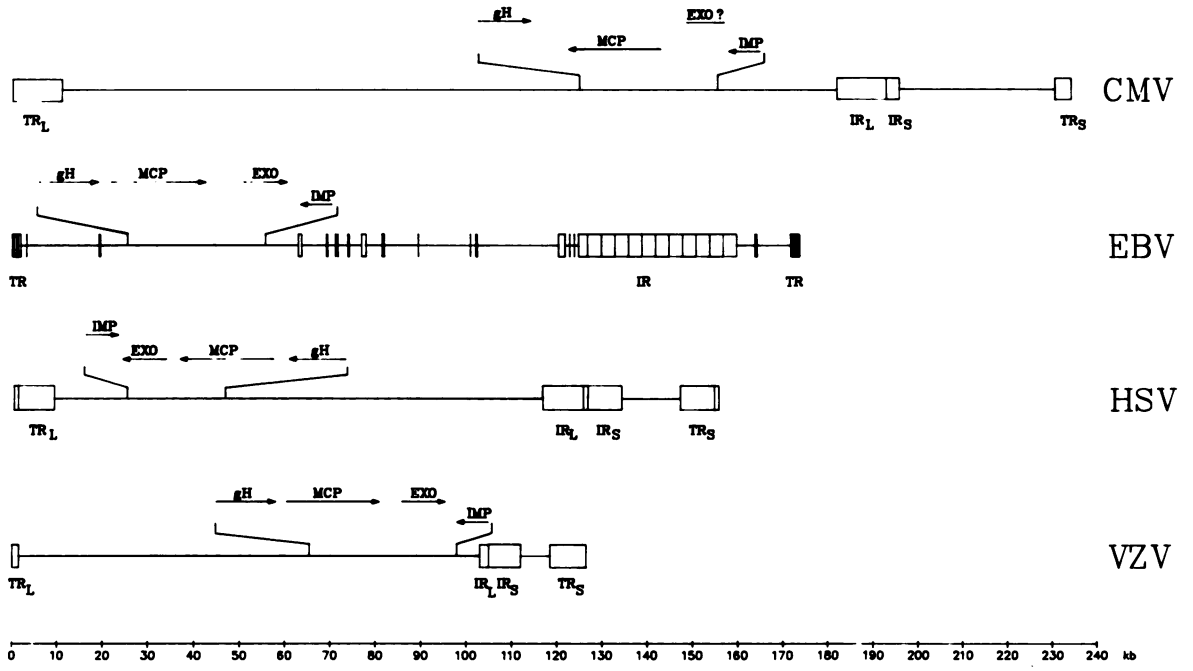

FIG. 9. Conservation of the gene blocks containing the genes coding for the homologous IMPs among the human herpesviruses. The open reading frame and the direction of transcription for the glycoprotein H (gH), the exonuclease (Exo), the major capsid protein (MCP), and the membrane protein (IMP) are indicated. Data are derived from Baer et al. (EBV) (4), Davison and Taylor (VZV) (16), McGeoch et al. (HSV) (44), Chee et al. (HCMV MCP) (11), and Cranage et al. (HCMV gH) (14).

DISCUSSION

In this study, we have described the identification of the gene coding for a 45-kDa envelope protein of HCMV. Until final agreement has been reached on the nomenclature of HCMV-encoded polypeptides, we propose to term this protein IMP (for integral membrane protein).

An immunoreactive lambda gt11 recombinant was initially isolated with the aid of an antiserum raised against solubilized proteins from envelopes of purified virions and NIEPs. The antiserum recognized mainly six polypeptides of 200, 170, 160, 75, 58, and 45 kDa.

The reactivity of the serum includes polypeptides that have been identified in the envelope of HCMV before, i.e., the 200-, 160-, 170-, and 58-kDa proteins (5, 20, 24, 55, 56). The 75-kDa protein could correspond to the 70-kDa glycoprotein, which is able to bind wheat germ agglutinin (5). A polypeptide of 45 kDa has not been described as a constituent of the viral envelope of HCMV. This could be a result of the extraction and detection methods used in this study. In contrast to previous investigations, which used Triton X-100 for the solubilization of envelope material, we used a combination of Nonidet P-40 and sodium deoxycholate. In addition, in previous studies the envelope polypeptides were metabolically labeled with [^{3}H]glucosamine (19, 24, 56) or detected by their affinity for lectins (5). Envelope proteins that are not modified by sugars or that are insoluble in Triton X-100 could not have been detected by these analyses. With the exception of the major glycoprotein gp58, the antiserum against envelope material did not recognize to a significant extent the other identified glycoproteins of HCMV to which monoclonal antibodies have been raised, namely glycoprotein H (14, 51), gp47 to gp52 (23), and gp65 (8). We also did not observe a reaction with the high-molecular-weight protein, which is complexed to gp58 via disulfide bonds (7). Whether this is a result of low antigenicity of these polypeptides in rabbits and/or a small amount of the respective glycoproteins in the envelope material has not been determined.

The coding sequence for the IMP is located within the long segment of the viral genome in the center of *Hin*dIII-R. The gene has a primary translation product of 42.8 kDa, provided that splicing does not occur. The carboxy-terminal part of the reading frame was expressed in *E. coli* as a β-galactosidase fusion protein and used to raise a monospecific antiserum in rabbits. This antiserum identified a 45-kDa viral polypeptide as the gene product, a protein which was also reactive with the anti-envelope serum.

Upon computer analysis, the IMP showed some characteristic features (Fig. 8). The N-terminal region consists of a short hydrophilic sequence and resembles the hydrophobic signal sequences usually found in membrane-bound glycoproteins and considered to be necessary for the translocation of the growing polypeptide chain across membranes (43). The hydropathicity plot of the entire protein shows eight sections of hydrophobic residues, each representing a potential membrane-spanning domain. The hydrophobic stretches are interrupted by hydrophilic amino acids, a feature characteristic of membrane-spanning proteins. The actual topology of the protein within the membrane, however, cannot be deduced from computer predictions; it has to be determined experimentally. Such experiments are currently in progress. The carboxy-terminal part is remarkably hydrophilic and acidic. Of the last 50 amino acids, 26 have charged side chains, and of the last 14 amino acids, 12 are negatively charged. The high percentage of hydrophilic residues at the C terminus might explain the antigenicity of this part of the molecule, which was deduced from the fact that the expression of nine amino acids in the lambda gt11 recombinant was sufficient for it to be recognized by the anti-envelope serum. The isolation of lambda gt11 recombinants expressing only a

few amino acids for foreign sequences does not seem to be an unlikely phenomenon since we reported previously the isolation of a cDNA clone expressing nine amino acids of gp58 of HCMV (39).

The IMP shows homology to predicted translation products from EBV, HSV-1, and VZV. These proteins have also been suggested to be membrane associated but to our knowledge have not been identified as constituents of the respective viruses (4, 15, 44). In general, they have a similar distribution of hydrophilic and hydrophobic residues across the molecule. Differences are mainly restricted to the variable length of the hydrophilic carboxy-terminal part of the proteins. For HCMV, this part contains 50 residues, whereas in HSV-1 the carboxy-terminal 120 amino acids are strongly hydrophilic (Fig. 8).

The IMP gene is transcribed into a 1.5-kb RNA which was detectable only late in the course of infection. The transcripts for the homologous reading frames in EBV and HSV-1 have also been determined to be late (4, 45).

With increasing sequence data available, it becomes apparent that the various human herpesviruses contain not only homologous genes but also, in some areas of the genome, a conserved gene arrangement (16). This has been shown for the genomic area encoding the DNA polymerase, the major glycoprotein, and a DNA-binding protein (32). Other areas of the genome, such as the entire U_S region of HCMV, do not show any homology to the other herpesviruses (60). The IMP of HCMV and its homologs in EBV, HSV, and VZV are conserved in a block that includes the genes coding for glycoprotein H, the major capsid protein, and the exonuclease. However, the gene for the major capsid protein is inverted in orientation in HCMV compared with the orientation in the other viruses (Fig. 9). Although it has not been proven experimentally, we conclude that the HCMV exonuclease gene is located between the major capsid protein and the IMP. In fact, there is a reading frame upstream of pp28 which shows homology to the BGLF5 of EBV (46), a suggested exonuclease (16). The high degree of conservation during evolution of the alpha-, beta-, and gammaherpesviruses suggests that these gene blocks reflect important functional domains.

In summary, we have identified a structural protein of HCMV which shows several membrane-spanning domains representing a type of protein not detected before as a constituent of HCMV envelopes. The conservation of the gene among four human herpesviruses indicates an important function in the replication cycle of the virus. As a part of the viral envelope, this protein most likely represents a target for the immune system. Experiments are currently in progress to establish the role of IMP in the humoral and cellular immune responses.

ACKNOWLEDGMENTS

This work was supported by the Johannes und Frieda Marohn Stiftung and the Wilhelm Sander Stiftung.

We appreciate the helpful discussions and support by B. Fleckenstein. We also thank T. Stamminger for preparing the oligonucleotides.

LITERATURE CITED

1. **Adams, G. A., and J. K. Rose.** 1985. Incorporation of charged amino acids into the membrane-spanning domain blocks cell surface transport but not membrane anchoring of a viral glycoprotein. Mol. Cell. Biol. **5**:1442–1448.
2. **Adams, G. A., and J. K. Rose.** 1985. Structural requirements of a membrane spanning domain for protein anchoring and cell surface transport. Cell **41**:1007–1015.
3. **Alford, C. A., S. Stagno, and R. F. Pass.** 1980. Natural history of perinatal cytomegaloviral infection, p. 125–147. *In* Perinatal infections, vol. 77. Elsevier/North-Holland Publishing Co., New York.
4. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrel, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Sèguin, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310**:207–211.
5. **Benko, D. M., and W. Gibson.** 1986. Primate cytomegalovirus glycoproteins: lectin-binding properties and sensitivities to glycosidases. J. Virol. **59**:703–713.
6. **Birnstiel, L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3'-processing: the end is in site! Cell **41**: 349–359.
7. **Britt, W. J.** 1984. Neutralizing antibodies detect a disulfide linked glycoprotein complex within the envelope of human cytomegalovirus. Virology **135**:369–378.
8. **Britt, W. J., and D. Auger.** 1985. Identification of a 65,000 dalton virion envelope protein of human cytomegalovirus. Virus Res. **4**:31–36.
9. **Bröker, M.** 1986. Vectors for the high level expression of proteins fused to truncated forms of *E. coli* β-galactosidase. Gene Anal. Technol. **3**:53–57.
10. **Burnette, W. N.** 1981. "Western blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal. Biochem. **112**:195–203.
11. **Chee, M., S.-A. Rudolph, B. Plachter, B. Barrell, and G. Jahn.** 1989. Identification of the major capsid protein gene of human cytomegalovirus. J. Virol. **63**:1345–1353.
12. **Claesson-Welsh, L., and P. Spear.** 1987. Amino-terminal sequence synthesis and membrane insertion of glycoprotein B of herpes simplex virus type 1. J. Virol. **61**:1–7.
13. **Craighead, J. E., R. E. Kanich, and J. D. Almeida.** 1972. Non-viral microbodies with viral antigenicity produced in cytomegalovirus-infected cells. J. Virol. **10**:766–775.
14. **Cranage, M. P., G. L. Smith, S. E. Bell, H. Hart, C. Brown, A. T. Bankier, P. Tomlinson, B. G. Barrell, and T. C. Minson.** 1988. Identification and expression of a human cytomegalovirus glycoprotein with homology to the Epstein-Barr virus BXLF2 product, varicella-zoster virus gpIII, and herpes simplex virus type 1 glycoprotein H. J. Virol. **62**:1416–1422.
15. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella zoster virus. J. Gen. Virol. **67**:1759–1816.
16. **Davison, A. J., and P. Taylor.** 1987. Genetic relation between varicella zoster virus and Epstein Barr virus. J. Gen. Virol. **68**:1067–1079.
17. **Debroy, C., N. Pederson, and S. Person.** 1985. Nucleotide sequence of a herpes simplex virus type 1 gene that causes cell fusion. Virology **145**:36–48.
18. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the Vax. Nucleic Acids Res. **12**:387–395.
19. **Farrar, G. H., and P. J. Greenaway.** 1986. Characterization of glycoprotein complexes present in human cytomegalovirus envelopes. J. Gen. Virol. **67**:1469–1473.
20. **Farrar, G. H., and J. D. Oram.** 1984. Characterization of human cytomegalovirus envelope glycoproteins. J. Gen. Virol. **65**: 1991–2002.
21. **Fleckenstein, B., I. Müller, and J. Collins.** 1982. Cloning of the complete human cytomegalovirus genome in cosmids. Gene **18**:39–46.
22. **Gibson, W.** 1983. Protein counterparts of human and simian cytomegalovirus. Virology **128**:391–406.
23. **Gretch, D. R., B. Kari, R. C. Gehrz, and M. F. Stinski.** 1988. A multigene family encodes the human cytomegalovirus glycoprotein complex gcII (gp47-52 complex). J. Virol. **62**:1956–1962.
24. **Gretch, D. R., B. Kari, L. Rasmussen, R. C. Gehrz, and M. F. Stinski.** 1988. Identification and characterization of three distinct families of glycoprotein complexes in the envelope of

human cytomegalovirus. J. Virol. **62**:875–881.
25. **Hu, N. T., and J. Messing.** 1982. The making of strand-specific M13 probes. Gene **17**:271–277.
26. **Irmiere, A., and W. Gibson.** 1985. Isolation of human cytomegalovirus intranuclear capsids, characterization of their protein constituents, and demonstration that the B-capsid assembly protein is also abundant in noninfectious enveloped particles. J. Virol. **56**:277–283.
27. **Jahn, G., T. Kouzarides, M. Mach, B.-C. Scholl, B. Plachter, B. Traupe, E. Preddie, S. C. Satchwell, B. Fleckenstein, and B. G. Barrell.** 1987. Map position and nucleotide sequence of the gene for the large structural phosphoprotein of human cytomegalovirus. J. Virol. **61**:1358–1367.
28. **Jahn, G., B.-C. Scholl, B. Traupe, and B. Fleckenstein.** 1987. The two major structural phosphoproteins (pp65 and pp150) of human cytomegalovirus and their antigenic properties. J. Gen. Virol. **68**:1327–1337.
29. **Kanich, R. E., and J. E. Craighead.** 1972. Human cytomegalovirus infection of cultured fibroblasts. II. Viral replicative sequence of a wild and an adapted strain. Lab. Invest. **27**: 273–282.
30. **Kari, B., N. Lussenhop, R. Goertz, M. Wabuke-Bunoti, R. Radeke, and R. Gehrz.** 1986. Characterization of monoclonal antibodies reactive to several biochemically distinct human cytomegalovirus glycoprotein complexes. J. Virol. **60**:345–352.
31. **Kouzarides, T., A. T. Bankier, S. C. Satchwell, E. Preddy, and B. G. Barrell.** 1988. An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. Virology **165**:151–164.
32. **Kouzarides, T., A. T. Bankier, S. C. Satchwell, K. Weston, P. Tomlinson, and B. G. Barrell.** 1987. Large-scale rearrangement of homologous regions in the genomes of HCMV and EBV. Virology **157**:397–413.
33. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eucaryotic mRNAs. Nucleic Acids Res. **12**:857–873.
34. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
35. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
36. **Landini, M. P., B. Severi, L. Badiali, E. Gonczol, and G. Mirolo.** 1987. Structural components of human cytomegalovirus: in situ localization of the major glycoprotein. Intervirology **27**:154–160.
37. **Landini, M. P., B. Severi, G. Furlini, and L. Badiali DeGiorgi.** 1987. Human cytomegalovirus structural components: intracellular and intraviral localization of pp28 and pp65-69 by immunoelectron microscopy. Virus Res. **8**:15–23.
38. **Liebowitz, D., D. Wang, and E. Kieff.** 1986. Orientation and patching of the latent infection membrane protein encoded by Epstein-Barr virus. J. Virol. **58**:233–237.
39. **Mach, M., U. Utz, and B. Fleckenstein.** 1986. Mapping of the major glycoprotein gene of human cytomegalovirus. J. Gen. Virol. **67**:1461–1467.
40. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Matteucci, M. D., and M. H. Caruthers.** 1981. Synthesis of deoxynucleotides on a polymer support. J. Am. Chem. Soc. **103**:3185–3191.
42. **McGavran, M. H., and M. G. Smith.** 1965. Ultrastructural, cytochemical and microchemical observations on cytomegalovirus (salivary gland virus) infection of human cells in tissue culture. Exp. Mol. Pathol. **4**:1–10.
43. **McGeoch, D. J.** 1985. On the predictive recognition of signal peptide sequences. Virus Res. **3**:271–286.
44. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**:1531–1574.
45. **McGeoch, D. J., A. Dolan, and M. C. Frame.** 1986. DNA sequence of the region in the genome of herpes simplex virus type I containing the exonuclease gene and neighbouring genes. Nucleic Acids Res. **14**:3435–3448.
46. **Meyer, H., A. T. Bankier, M. P. Landini, C. M. Brown, B. G. Barrell, B. Rüger, and M. Mach.** 1988. Identification and procaryotic expression of the gene coding for the highly immunogenic 28-kilodalton structural phosphoprotein (pp28) of human cytomegalovirus. J. Virol. **62**:2243–2250.
47. **Meyers, J. D.** 1982. Nonbacterial pneumonia after allogenic marrow transplantation: a review of ten years experience. Rev. Infect. Dis. **4**:1119–1132.
48. **Murray, R. J., D. Wang, L. S. Young, F. Wang, M. Rowe, E. Kieff, and A. B. Rickinson.** 1988. Epstein-Barr virus-specific cytotoxic T-cell recognition of transfectants expressing the virus-coded latent membrane protein. J. Virol. **62**:3747–3755.
49. **Nowak, B., C. Sullivan, P. Sarnow, R. Thomas, F. Bricant, I. C. Nicolas, B. Fleckenstein, and A. J. Levine.** 1984. Characterization of monoclonal antibodies and polyclonal immune sera directed against human cytomegalovirus virion proteins. Virology **132**:325–338.
50. **Pereira, L., M. Hoffman, M. Tatsuno, and D. Dondero.** 1984. Polymorphism of human cytomegalovirus glycoproteins characterized by monoclonal antibodies. Virology **139**:73–86.
51. **Rasmussen, L. D., R. Nelson, U. Kelsa, and T. Merigan.** 1984. Murine monoclonal antibody to a single protein neutralizes the infectivity of human cytomegalovirus. Proc. Natl. Acad. Sci. USA **81**:876–880.
52. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
53. **Schreier, P. H., and R. Cortese.** 1979. A fast and simple method for sequencing DNA in the single stranded bacteriophage M13. J. Mol. Biol. **129**:169–172.
54. **Spector, S. A., K. K. Hirata, and T. R. Neuman.** 1984. Identification of multiple cytomegalovirus strains in homosexual men with acquired immune deficiency syndrome. J. Infect. Dis. **150**:953–955.
55. **Stinski, M. F.** 1976. Human cytomegalovirus: glycoproteins associated with virions and dense bodies. J. Virol. **19**:594–609.
56. **Stinski, M. F.** 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. J. Virol. **23**:751–767.
57. **Talbot, P., and J. D. Almeida.** 1977. Human cytomegalovirus: purification of enveloped virions and dense bodies. J. Gen. Virol. **36**:345–349.
58. **Tanford, C.** 1978. The hydrophobic effect and the organization of living matter. Science **200**:1012–1018.
59. **Thorley-Lawson, D. A., and E. S. Israelson.** 1987. Generation of specific cytotoxic T-cells with a fragment of the Epstein Barr virus encoded p63/latent membrane protein. Proc. Natl. Acad. Sci. USA **84**:5384–5388.
60. **Weston, K., and B. G. Barrell.** 1986. Sequence of the short unique region, short repeats, and part of the long repeats of human cytomegalovirus. J. Mol. Biol. **192**:177–208.
61. **Wilson, I. A., J. J. Skehel, and D. C. Wiley.** 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. Nature (London) **289**:366–373.
62. **Xu-Bin, T. Murayama, K. Ishida, and T. Furukawa.** 1989. Characterization of IgG Fc receptors induced by human cytomegalovirus. J. Gen. Virol. **70**:893–900.
63. **Yanish-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC vectors. Gene **33**:103–119.
64. **Young, R. A., and R. W. Davis.** 1983. Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA **80**: 1194–1198.