JOURNAL OF VIROLOGY, Aug. 1989, p. 3330–3337
0022-538X/89/083330-08$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 63, No. 8

Identification and Expression of a Human Cytomegalovirus Early Glycoprotein

CHIA-PING CHANG,[1] DAVID H. VESOLE,[1] JAY NELSON,[2] MICHAEL B. A. OLDSTONE,[2] AND MARK F. STINSKI[1]*

Department of Microbiology, University of Iowa, Iowa City, Iowa 52242,[1] and Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, California 92537[2]

Received 17 February 1989/Accepted 21 April 1989

A human cytomegalovirus early gene which possesses three temporally regulated promoters is located in the large unique component of the viral genome between 0.054 and 0.064 map units (C.-P. Chang, C. L. Malone, and M. F. Stinski, J. Virol. 63:281–290, 1989). This gene contains a major open reading frame (ORF) located 233 bases downstream of the cap site of an early unspliced RNA. The major ORF predicts a polypeptide of 17 kilodaltons (kDa) which contains a glycoproteinlike signal and anchor domains as well as potential N-glycosylation sites. Antisera were prepared against synthetic peptides derived from amino acid sequences within the major ORF. The antisera detected a viral glycoprotein of 48 kDa in infected cells and recognized the in vitro-translated 17-kDa protein early-gene product. The viral glycoprotein, designated gp48, was modified by N-linked glycans and possibly O-linked glycans. The synthesis of gp48 occurred in the absence of viral DNA replication but accumulated to the highest levels at late times after infection. Since gp48 was found in the virion, it is considered an early structural glycoprotein.

Primary infection by human cytomegalovirus (HCMV) is usually asymptomatic, and the virus remains latent within a healthy individual. However, the virus can cause severe cytomegalic inclusion disease in 1 to 5% of all infants born with an HCMV infection. In addition, immunocompromised patients, such as cancer patients, acquired immunodeficiency syndrome patients, and organ transplant recipients, are at high risk for developing HCMV infection (13, 25).

The genome of HCMV, approximately 240 kilobases of double-stranded DNA, has the capacity to code for approximately 200 proteins. The virion is estimated to consist of 30 to 35 different structural proteins (11, 27, 33), including the matrix-tegument proteins and envelope glycoproteins. The viral glycoproteins play an important role in infectious virus assembly and in the immunobiology of the virus-host interaction. There has been significant progress in the identification of the HCMV glycoprotein genes because of DNA sequencing, but not all of the viral glycoproteins in the infected cell or viral envelope have been identified.

Several distinct viral glycoproteins have been identified in the envelope of HCMV by using monoclonal antibodies (3–5, 9, 10, 19, 24, 26, 27). Some of these glycoproteins are in disulfide-linked complexes that can be separated by rate-zonal centrifugation. Three glycoprotein complexes have previously been identified and designated gcI, gcII, and gcIII (10). The major component of the outer virion envelope is gcI, which contains a homolog of herpes simplex virus gB. gcI consists of a glycoprotein designated gp55 and a heterogeneous glycoprotein designated gp93-130 in a disulfide-linked complex (1, 8). The viral gene maps between 0.344 and 0.360 map units and is generally conserved in herpesviruses (3, 20). A gene family designated HXLF (*Hin*dIII-X left reading frame) consists of five ORFs, with various degrees of homology, that lie in tandem in the short unique component of the viral genome (36). The HXLF1 and HXLF2 genes code for proteins of 21 to 25 and 20 kilodaltons (kDa), respectively. These viral proteins are glycosylated to 47 to 52 kDa (gp47-52) (9). The viral glycoproteins complex with each other and possibly with other members of the HXLF family to make up a family of glycoprotein complexes designated gcII (9). The third envelope glycoprotein complex, gcIII, contains two glycoproteins designated gp86 and gp145 (10). gp86 is a homolog of herpes simplex virus gH (4). This gene was localized between 0.45 and 0.47 map units (4). The gene that encodes gp145 has not been identified. The viral envelope glycoprotein complexes are presumably key determinants in the host immune response to infection, since monoclonal antibodies that react with these glycoproteins neutralize viral infectivity (4, 15, 27). Therefore, it is important to identify the viral envelope glycoproteins and to characterize their genes.

We recently described the transcription of an early gene which maps between 0.054 and 0.064 map units and possesses three temporally regulated promoters (2). The viral mRNAs were analyzed by single-strand-specific nuclease protection experiments and were found to be unspliced mRNAs (2). We demonstrate here that the viral gene product is both an early viral glycoprotein as well as a component of the virion. The early viral glycoprotein was designated gp48. The translation of the viral gene in vitro as well as in virus-infected cells is described.

MATERIALS AND METHODS

Virus and tissue culture. The growth of human foreskin fibroblast cells, the propagation of HCMV (Towne strain), and the use of phosphonoacetic acid (PAA) to inhibit viral DNA replication have been described previously (32, 35). All experiments were conducted at 5 to 10 PFU per cell.

Antisera. The four peptides used for generating antibodies in this study were synthesized by an Applied Biosystems peptide synthesizer and are shown in Fig. 5B. Antipeptide sera were prepared as described previously (12). To reduce nonspecific reactivity, the antisera were passed through a Sepharose 4B (Pharmacia, Inc., Piscataway, N.J.) column conjugated with human foreskin fibroblast cellular proteins.

DNA sequencing. DNA sequence analysis was done by the

* Corresponding author.

chemical cleavage method of Maxam and Gilbert (21) or the dideoxy chain termination method of Sanger et al. (28). A 16-mer oligonucleotide described for primer extension analysis (2) was used as the primer for DNA sequencing.

Plasmid constructions. For constructing a series of deletion plasmids in the 5′ noncoding region (5′ NCR), pUCMS, a pUC19-derived plasmid containing the entire early glycoprotein gene (*Mlu*I to *Sal*I site [Fig. 1]), was linearized at the unique *Nru*I site and digested with exonuclease III, followed by digestion with mung bean nuclease. *Hin*dIII linkers were added to the blunt ends. The shortened inserts were isolated from the pUC19 vector by digestion with *Hin*dIII and *Pst*I and then inserted into the corresponding sites of pSP64. The deletion endpoints of four clones (pSPdl1, pSPdl5, pSPdl13, and pSPdl16) were determined by sequencing the plasmids with an SP6 promoter primer according to the protocol supplied by the manufacturer (Boehringer Mannheim Biochemicals, Indianapolis, Ind.). The TrpE fusion protein vector pATH1 (a gift from Lung-Ji Chang, University of California at San Francisco), which contains a truncated *trpE* gene, was used to construct an ORF3-*trpE* fusion gene by inserting a 640-base-pair DNA fragment into a *Bam*HI-digested pATH1 by methods previously described (29). The resulting plasmid was designated pATH1ORF3.

In vitro transcription. The 5′ NCR deletion plasmids were linearized with *Bam*HI or *Eco*RI (pSPdl16 only), extracted with phenol-chloroform, and ethanol precipitated. The optimal conditions for in vitro transcription with SP6 RNA polymerase have previously been determined (22).

In vitro translation. Standard translations, using [^{14}C] leucine (342 mCi/mmol, at 50 μCi/ml; Amersham, Arlington Heights, Ill.) as the labeled amino acid in wheat germ extracts, were performed exactly as described by the supplier (Promega Biotec, Madison, Wis.). Approximately one-fifth of the sample was analyzed directly by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Immunoprecipitation. For immunoprecipitation of in vitro translation products, 540 μl of RIPA buffer (phosphate-buffered saline [pH 7.4], containing 0.5% sodium deoxycholate, 1% Nonidet P-40, 0.1% SDS, 1 mg of ovalbumin per ml, 0.02% sodium azide [NaN_3], and 1 mM phenylmethylsulfonyl fluoride) and 10 μl of antipeptide serum were added to the reaction mixture and incubated at 4°C overnight. For immunoprecipitation of metabolically labeled antigens, infected or mock-infected cells were pulse-labeled from 48 to 72 h postinfection (p.i.) with [^{3}H]glucosamine (33 Ci/mmol, at 40 μCi/ml; Amersham) and then solubilized in RIPA buffer. The clarified lysates were preadsorbed with normal rabbit serum and protein A-Sepharose CL-4B (Pharmacia). An appropriate amount of antiserum was then added to the lysate (5×10^6 to 10×10^6 cpm) and incubated at 4°C overnight. Immune complexes were recovered by incubation with 50 μl of a 50% suspension of protein A-Sepharose CL-4B at room temperature for 1 h. The beads were washed three times with phosphate-buffered saline containing 0.1% Nonidet P-40 and once with sterile water. Antigens were eluted with dissociation buffer and analyzed by SDS-PAGE and autoradiography as described previously (32).

Immunoblotting. Mock- or HCMV-infected cells were washed with phosphate-buffered saline and solubilized in SDS-PAGE sample buffer. The proteins were separated in SDS-polyacrylamide gels (10 or 15% polyacrylamide) and electrophoretically transferred to nitrocellulose sheets by the method of Towbin et al. (34). After transfer, the blots were immersed in milk buffer (Tris-buffered saline containing 10 mM Tris hydrochloride [pH 7.5], 0.9% NaCl, 5% nonfat dry milk, 0.01% [vol/vol] antifoam A emulsion [Sigma Chemical Co., St. Louis, Mo.], and 0.01% NaN_3) at 4°C overnight or at room temperature for 2 h. The blots were incubated with rabbit antipeptide serum (1:100 dilution) for 2 h and then with ^{125}I-labeled protein A (10^6 cpm/ml; Amersham) for 1 h in milk buffer. All incubations were performed at room temperature and terminated by washing the blots four times in milk buffer with 15 min per wash.

Inhibition of glycoprotein biosynthesis and processing. Tunicamycin (Calbiochem-Behring, San Diego, Calif.) or monensin (a gift from Charles Grose, University of Iowa, Iowa City) was added to the medium at 24 h p.i. to a concentration of 0.5 μg/ml or 0.2 μM, respectively. The cells were treated continuously for 48 h and were harvested and analyzed by the Western blot (immunoblot) technique.

Purification of virions and dense bodies. Virions and dense bodies of HCMV were purified from the extracellular fluid of infected human foreskin fibroblast cells as previously described (30).

RESULTS

Two putative ORFs encoded by the viral early gene. ORF analysis of the DNA sequence between 0.054 and 0.064 map units revealed the presence of two overlapping ORFs (ORF1 and ORF3) coding for 84 and 148 amino acids (Fig. 1). The AUG codons of ORF1 and ORF3 are located 285 and 233 base pairs downstream from the major transcription start site, respectively. The unspliced viral mRNAs from this transcription unit have been described previously (2).

Nucleotide sequence and the deduced polypeptide. The nucleotide sequence of a 2.3-kilobase *Mlu*I-*Sal*I segment was determined as described in Materials and Methods. Figure 2 shows the sequence from the major early transcription start site to the downstream polyadenylation signal. The AUG at position 233 is of moderate strength by the rules of Kozak (18). The AUG codons of the upstream short ORFs are at positions 12, 26, and 98, and they lie in an unfavorable context when compared with the AUG of ORF3. The third reading frame has a capacity to encode a 148-residue polypeptide with a molecular mass of 17,132 Da. The 1.5-kilobase major early mRNA would contain a 232-base 5′ untranslated sequence and a 860-base 3′ untranslated sequence. The polyadenylation signal (AATAAA) is located at nucleotides 1507 through 1513. Downstream of this signal is a consensus GT cluster. The hydrophobicity profile of the predicted protein was determined by the method of Hopp and Woods (14). There is a hydrophobic signal sequence at the amino terminus and a hydrophobic carboxy terminus, which resembles a membrane anchorage domain (Fig. 3). In addition, the protein has 10 potential N-linked glycosylation sites (Fig. 2). These properties suggest that the viral gene product could be a membrane-bound glycoprotein.

Translation in vitro. To assign a viral protein(s) to these putative ORFs, translation of hybrid-selected mRNAs in cell extracts was performed. In vitro transcription and translation of the viral gene was also performed. However, both of these experiments failed to detect a protein (data not shown), suggesting that translation of this message was repressed by an unknown mechanism. Since translational regulation via the 5′ untranslated region has been reported in procaryotic and eucaryotic systems, we modified the 5′ leader sequence by progressive deletion analysis. Figure 4A shows the nucleotide sequence of the 5′ leader region of the viral gene, which contains three tandem ORFs upstream of the initiation codon of ORF3. Deletion plasmids containing a

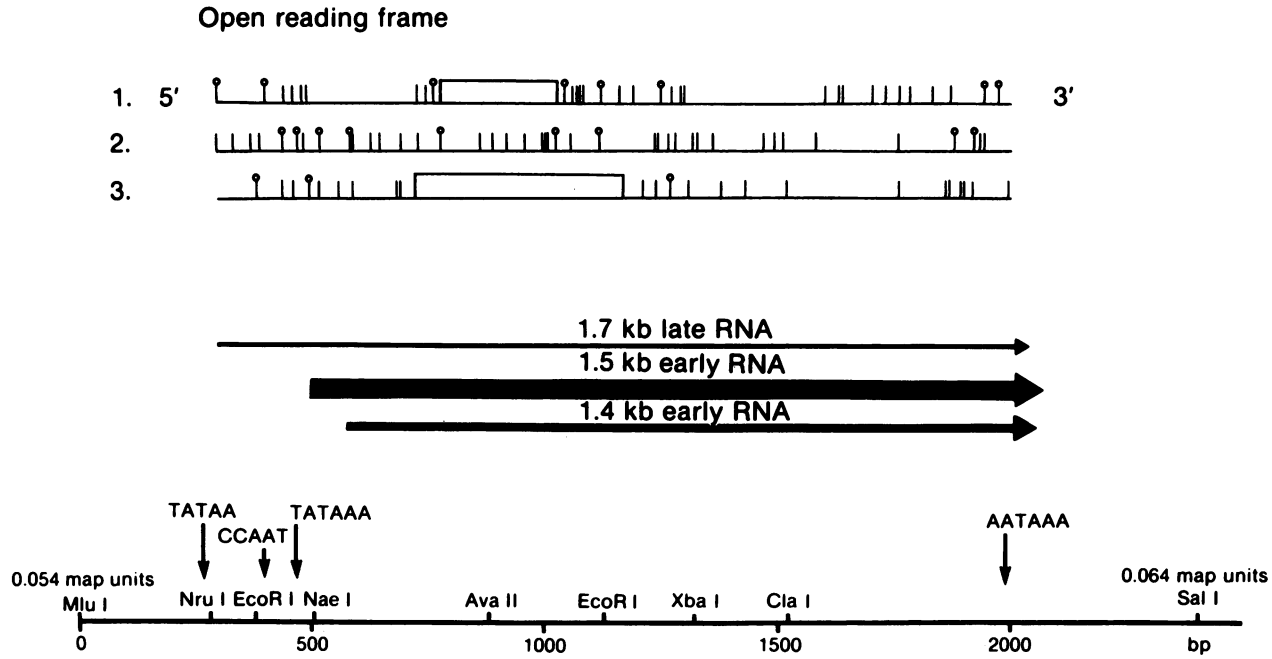

FIG. 1. Organization of an HCMV early transcription unit between 0.054 and 0.064 map units. The gene is drawn in a 5'-to-3' orientation, with a bar diagram showing the translation initiation codons (♀) and stop codons ( ) in all three reading frames. Potential protein-coding regions are indicated (□). The three unspliced transcripts, represented by three arrows above the restriction map, have been described previously (2). The thickness of each arrow represents the approximate abundance of each RNA in infected cells. The promoter consensus sequences and the polyadenylation signal are indicated. kb, Kilobase; bp, base pair.

portion of the 5' leader were generated by exonuclease III digestion from the parental vector pUCMS, as described in Materials and Methods. The effect of each deletion on the translational efficiency was determined by a combination of in vitro transcription and translation. A deletion to +20 (pSPdl1), which removed the first short ORF, did not release the block (Fig. 4B, lane 2). A deletion to +81 (pSPdl5) resulted in the appearance of a protein of approximately 17 kDa (Fig. 4B, lane 3). A further deletion to +138 (pSPdl13), removing all three upstream ORFs, slightly increased the translation of the 17-kDa protein (Fig. 4B, lane 4). Removal of additional nucleotides from +138 to +194 (pSPdl16) greatly enhanced translation (Fig. 4B, lane 5).

The protein translated from pSPdl16 was initiated from the AUG of ORF3. The evidence for this was provided by in vitro transcription and translation of *Eco*RI-linearized pSPdl16. Since *Eco*RI cleaved outside of ORF1 but within the carboxy terminus of ORF3, a truncated protein would be predicted. Indeed, a protein of ~15 kDa was detected (Fig. 4B, lane 6).

Immunoprecipitation of in vitro translation products. To eventually determined whether the early gene specified a bona fide viral protein, antibodies directed against synthetic peptides corresponding to different regions of ORF1 and ORF3 were prepared (Fig. 4B). The deletion clone pSPdl16 was linearized at the *Bam*HI site and used as a template for in vitro transcription. Translation was carried out either in the presence or absence of in vitro-synthesized RNA. The translation products were subsequently immunoprecipitated with serum 2-1 against peptides (antipeptide 2-1 serum) or a human anti-HCMV convalescent serum. Both antisera reacted with a protein of ~17 kDa (Fig. 5A, lanes 3 and 5). In addition, antipeptide 2-1 serum recognized a smaller protein which may represent premature termination of translation or a degraded translation product. Antibodies directed against synthetic peptides corresponding to different regions of ORF1 failed to detect viral antigens. These results demonstrate that the early gene can direct the synthesis of a protein encoded by ORF3 and that antipeptide 2-1 serum or human convalescent serum had antibodies to this viral antigen.

Identification of a viral protein in infected cells. Antipeptide 2-1 serum was also tested for its ability to recognize a specific viral antigen in infected cells. Cell lysates obtained from mock-infected or infected cells at various times after infection were analyzed by Western blot analysis using antipeptide 2-1 serum as described in Materials and Methods. A protein of approximately 48 kDa was weakly detected at 24 h p.i. (Fig. 6A, lane 2), and accumulated to higher levels at late times after infection (Fig. 6A, lanes 3 and 4). This protein was not present in mock-infected cells (Fig. 6A, lane 1) and was not detected with preimmune serum (data not shown). Synthesis of this viral protein was independent of viral DNA replication, as demonstrated by its insensitivity to PAA (Fig. 6B). The sensitivity of HCMV late protein synthesis to PAA has been previously demonstrated (31). For example, the HCMV glycoprotein gp47-52 was not detected in the presence of PAA (D. R. Gretch and M. F. Stinski, unpublished data). Antipeptide 2-1 serum also recognized the TrpE-ORF3 fusion protein made in *Escherichia coli* (Fig. 6A, lane 5), but not the TrpE truncated protein (Fig. 6A, lane 6). Since the TrpE-ORF3 fusion protein contains the amino acid sequence of ORF3 downstream from the *Ava*II site (Fig. 1), this result confirmed that the 48-kDa viral protein is encoded by ORF3. None of the antisera directed against ORF1 peptides recognized a specific viral antigen by Western blot analysis or by immunoprecipitation of [^{35}S]cysteine- or [^{35}S]methionine-labeled cells (data not shown).

The viral early gene encodes a glycoprotein. The size difference between the sequence-predicted polypeptide and

```
                                                    +1
GGA TAT GTA TAA ATG AGG AGT GAA ATA GGC ACA GGG AAT CAG +10

ATG CCG GCC TTG TGA TGC AGC CGC TGG TTC TCT CGG CGA AAA +52

AAC TGT CGT CTT TGC TGA CTT GCA AAT ACA TCC CGC CTT AAG +94

TGA TGA GTC TAT AAA GCA CCG TTG TCT GGG TAC GGT AAA AGT +136

GAC TCG GAT TGT AGC ACG TCA TTT TTT TTT GTT TTT GCA TCG +178

TTT ATC GTC ACC ACT AGT GCA ATA TTT TGA TCG TAA GGC TGA +220

AAG AGT ATC GTT ATG ATG CTT AGA GCG TGG AGA TTG ATG GTA +262
                Met Met Leu Arg Ala Trp Arg Leu Met Val (10)

CTA CTT GCC GCG TAC TGT TAT TAT GTT TTT GCG AAT TGT TCA +304
Leu Leu Ala Ala Tyr Cys Tyr Tyr Val Phe Ala Asn Cys Ser (25)

ATC AGC ACG ACG ACT GCT CCT GTG GAA TGG AAG TCT CCC AAC +346
Ile Ser Thr Thr Thr Ala Pro Val Glu Trp Lys Ser Pro Asn (39)

CGT CAG ATT CCC AAG AAT ATT ACT TGC GCT AAT TAC TCA GGG +388
Arg Gln Ile Pro Lys Asn Ile Thr Cys Ala Asn Tyr Ser Gly (53)

ACC GTC GGC GGT AAC GTT ACT TTT CAG GGT CTC AAG AAT AAA +430
Thr Val Gly Gly Asn Val Thr Phe Gln Gly Leu Lys Asn Lys (67)

ACG GAA GAT TTT TTA TCC TGG CTA CTC GGG TCT GGT TAT AAG +472
Thr Glu Asp Phe Leu Ser Trp Leu Leu Gly Ser Gly Tyr Lys (81)

TCC ATT TGC TCG TTC TTC CCG CAA CTC CCT GGT GAT TCT AAT +514
Ser Ile Cys Ser Phe Phe Pro Gln Leu Pro Gly Asp Ser Asn (95)

GAG CAG CAT TAC AGA TAT GAA GTA ACC AAC CTC ACG TAC AAT +556
Glu Gln His Tyr Arg Tyr Glu Val Thr Asn Leu Thr Tyr Asn (109)

TGC ACC TAT GAC CGC CTG ACG TTA CTG AAT CTG ACA ACG GAA +598
Cys Thr Tyr Asp Arg Leu Thr Leu Leu Asn Leu Thr Thr Glu (123)

AAC AGC AGG AAT TAC TAT TTC AGA AGA GAA GAT GCG AAT TCC +640
Asn Ser Arg Asn Tyr Tyr Phe Arg Arg Glu Asp Ala Asn Ser (137)

ACC TTC TAT TAC TCT TGT TAC AAT CTG ACC GTG TCC TAA AGA +682
Thr Phe Tyr Tyr Ser Cys Tyr Asn Leu Thr Val Ser Stop    (148)

TCG CAC GTG AAG TTC CAC AGA AAC CGC GCA GCT GTA GCT ATT +724

GTG TTT ACG TTG CTT TTG AAA TGT TAA GCG TCC CTA CGG CGC +766

TAA CAT GTT TCT AGG CTA CTC TGA CTG TGT AGA TCC CGG CCT +808

TGC TGC GTA TCG TGT ATC TAG ATC ACG CTT AAA GCT CGT GTT +850

GTC TTT TGT GTG GTT GAT CGG TTT GCG TCT CCA TGA TTG TGC +892

CAC GTT CGA GTC CTG CTG TTA CGA CAT CAC CGA GGC GGA GAG +934

TAA CAA GGC TAT ATC AAG GGA CGA AGC AGC ATT CAC CTC CAG +976

CGT GAG CAC TCG TAC ACC ATC CCT AGC GAT CGC GCC CCT CCT +1018

CCT GAC CGA TCG ATG CTG TTG TCG CGG GAG GAA GAA CTC GTT +1060

CCG TGG AGT CGT CTC ATC ATC ACT AAG CAG TTC TAC GGA GGC +1102

CTG ATT TTC CAC ACC ACC T-G GTC ACC CGG CTT CGT CTT GCT +1144

AGG ACT TTT GAC GCT TTT CGC CAG CCT GTT TCG CGT ACC CAC +1186

AAT CCA TCT GTC GTT TCT GCA TAG ACC GTC TCC GGG ACA TCG +1228

CCC GTC CTC TGA AAT ACC GCT ATC AAC GTC TCG TCG CTA CCG +1270

TGT AGC TAG TTA GCC AGC TGT GTA TAG TTT GTT GTG TTT TGC +1312

TTT TGC ATA TTT GTT TTC AGT CAG AGA GTC TGA AAC GGG GTG +1354

GGA GGG ACT TTT GCG GGT AGT GCA CGC TAA GAT GAA CGG GTG +1396

GGC TGG GGT GTG CTT GAT AAC TCA CTG TTT GAA TAC CGC TCA +1438

CGC ACA TAT GTA GCA CTC AAC ATG TTA GCT TTT GCC CGC ACG +1480

CCC CGG GGC ATG CCG AGC TGC CTT TTT AAT AAA GTC TGG GTT +1522

TCC AGA TAC GCG CTG GTT CTG ATT TTG ATG GTT TGT GCC TCT +1564
```

FIG. 2. DNA sequence and the predicted amino acid sequence of the early gene. The nucleotide sequence (top line) and the amino acid sequence of ORF3 (bottom line) are numbered. The transcription start site of the major early RNA is numbered +1. The peptide of 18 amino acids used to generate antipeptide 2-1 serum is boxed. Potential N-glycosylation sites (Asn-X-Ser and -Thr) and the polyadenylation signal are underlined. The nucleotide at position +1122 was not determined.

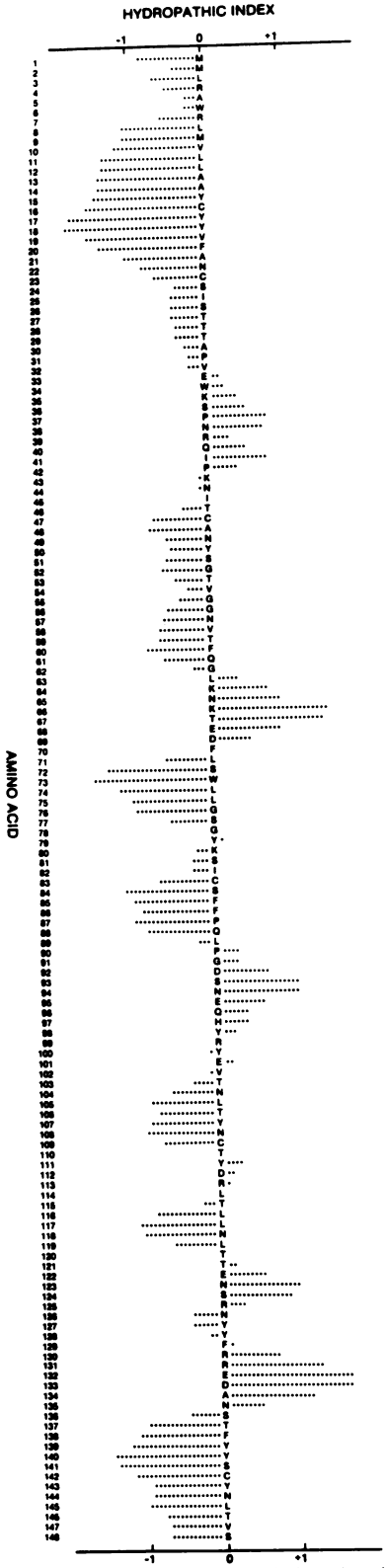

FIG. 3. Hydrophobicity profile of the deduced early protein. The hydropathy value was defined by the method of Hopp and Woods (14), with a moving window of six amino acids. The average hydrophilicity (+) and hydrophobicity (−) are indicated by the horizontal bars. The predicted amino acid sequence of ORF3 is given by using the single-letter code.

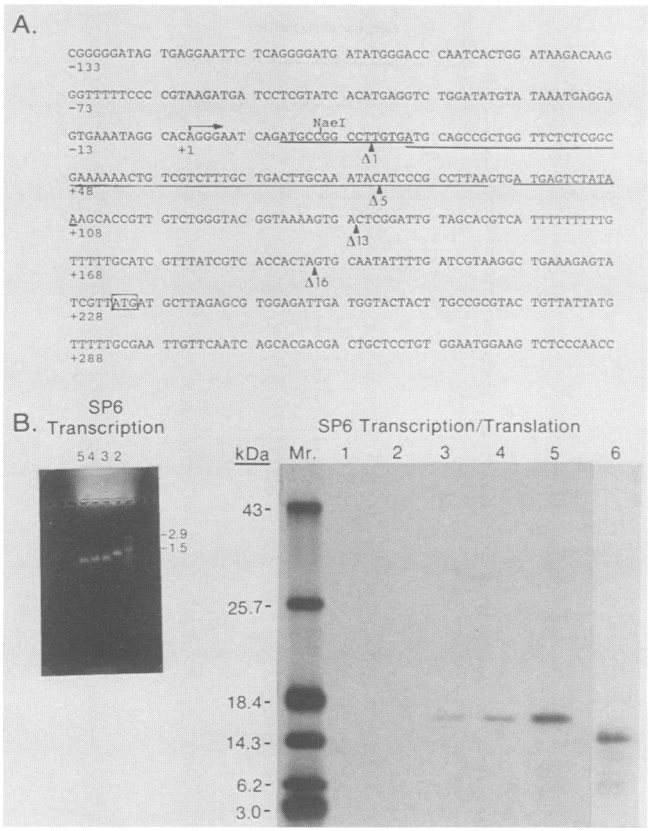

FIG. 4. 5′ NCR of the early gene and in vitro translation. (A) Nucleotide sequence of the 5′ NCR of the early gene and its deletion clones. The major early transcription start site is numbered +1 and marked with an arrow. The first ATG of ORF3 is boxed, and the three short upstream ORFs are underlined. The 3′ boundaries of deletion clones are indicated by arrowheads. Δ1, pSPdl1; Δ5, pSPdl5; Δ13, pSPdl13; Δ16, pSPdl16. (B) The 5′ NCR deletion clones were transcribed by using SP6 RNA polymerase as described in the text. Samples of the synthesized RNAs were examined by electrophoresis in a formaldehyde gel (left panel). The positions of *E. coli* rRNA size markers are indicated in kilobases. The same quantity (2 μg) of RNAs derived from each deletion clone was added to wheat germ extracts and translated by using [^{14}C]leucine as a radiolabeled amino acid. The translation products were analyzed by SDS-PAGE and autoradiography (right panel). Lanes: 1, no RNA control; 2, *Bam*HI-linearized pSPdl1; 3, *Bam*HI-linearized pSPdl5; 4, *Bam*HI-linearized pSPdl13; 5, *Bam*HI-linearized pSPdl16; 6, *Eco*RI-linearized pSPdl16; Mr., Protein molecular size standards.

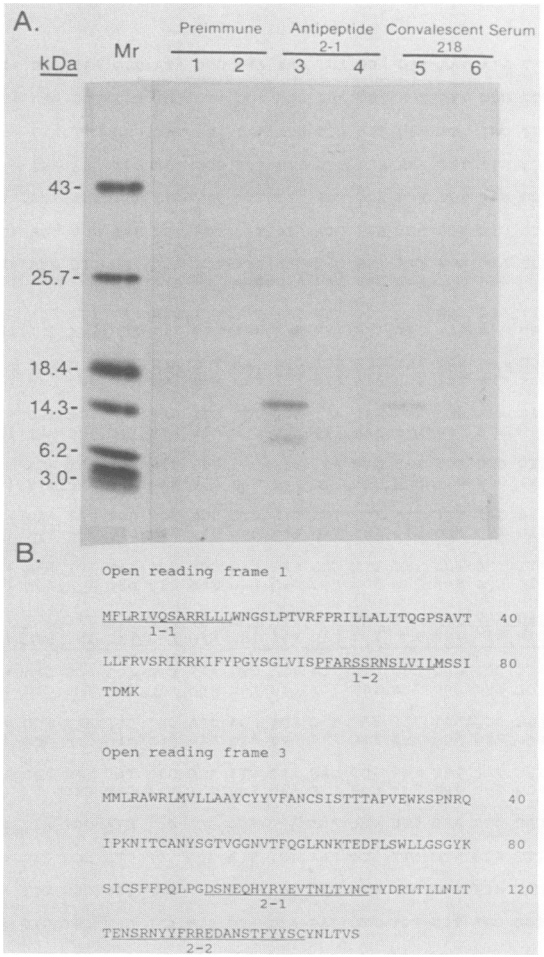

FIG. 5. Immunoprecipitation of the in vitro translation products by antipeptide 2-1 serum and a human anti-HCMV convalescent serum. (A) pSPdl16 linearized with *Bam*HI was transcribed and translated as described in Fig. 4. Proteins were translated either in the presence (lanes 1, 3, and 5) or the absence (lanes 2, 4, and 6) of synthetic RNA and immunoprecipitated with preimmune serum (lanes 1 and 2), antipeptide 2-1 serum (lanes 3 and 4), or human anti-HCMV convalescent serum 218 (lanes 5 and 6). A fluorograph of a SDS-polyacrylamide gel (15% polyacrylamide) is shown. Mr, Protein molecular size standards. (B) The predicted amino acid sequences of ORF1 and ORF3 are given by using the single-letter code. The sequences of synthetic peptides used for generation of antibodies are underlined and numbered. The amino acid numbers are indicated on the right margin.

the protein detected from infected cell lysates suggested that posttranslational modifications had occurred in vivo. The presence of a number of potential N-linked glycosylation sites further suggested that the early gene might encode an as yet unidentified viral glycoprotein. To test this possibility, infected cells were radiolabeled with [^{3}H]glucosamine from 48 to 72 h p.i. The antigens were solubilized in RIPA buffer and immunoprecipitated with antipeptide 2-1 or 2-2 serum (Fig. 5). A diffuse glycoprotein band was detected by both antisera (Fig. 7, lanes 2 to 5). However, both antisera functioned poorly in the presence of SDS, and consequently, only a relatively low amount of [^{3}H]glucosamine-labeled viral antigen was immunoprecipitated. The sizes of the ^{3}H-labeled glycoproteins detected by immunoprecipitation agree with the protein size detected by Western blot analysis. We designated the product of this early gene as gp48.

Effects of tunicamycin and monensin. To further substantiate that gp48 has glycan modifications, we analyzed the effects of tunicamycin and monensin as described in Materials and Methods. Since gp48 accumulates slowly from early to late times, the drugs were added at 24 h p.i. and the cells were harvested 48 h later. Treatment of infected cells with tunicamycin, an antibiotic which block N-linked glycoprotein synthesis, resulted in accumulation of a precursor protein of 27 kDa detected by Western blot analysis with antipeptide 2-1 serum (Fig. 8, lane 1). Since the protein is 10 kDa larger than the polypeptide backbone and the addition of O-linked oligosaccharides to nascent polypeptides is not blocked by tunicamycin, the 27-kDa protein most likely contains O-linked oligosaccharides. The above precursor antigen was not detected by preimmune serum (data not shown). No precursor viral antigens were detected in the

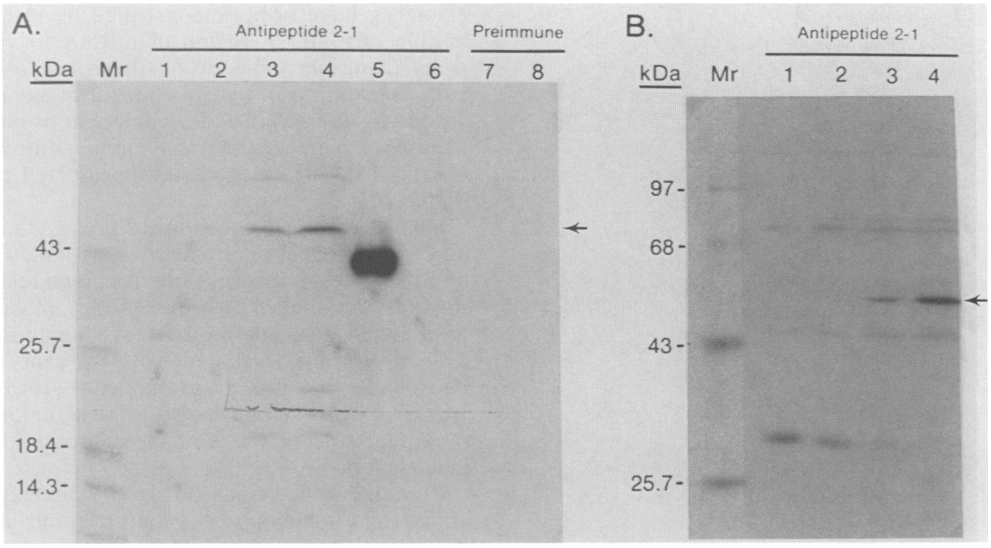

FIG. 6. Immunoblot analysis of infected cell lysates and the TrpE-ORF3 fusion protein. Human foreskin fibroblast cells were mock infected or infected with HCMV either in the absence (A) or in the presence (B) of PAA. Cells were harvested at various times p.i., lysed, and fractionated in SDS-polyacrylamide gels (15% polyacrylamide) along with the TrpE truncated protein and TrpE-ORF3 fusion protein made in *E. coli*. The proteins were transferred to nitrocellulose and probed with preimmune (lanes 7 and 8) or antipeptide 2-1 serum (lanes 1 to 6). Lanes: 1, mock-infected cell lysate; 2, lysate at 24 h p.i.; 3, lysate at 48 h p.i.; 4, lysate at 72 h p.i.; 5 and 7, TrpE-ORF3 fusion protein; 6 and 8, TrpE truncated protein; Mr, Protein molecular size standards. The location of the viral antigen is designated by an arrow.

monensin-treated cells (Fig. 8, lane 2). This negative result may reflect instability of the precursor glycoprotein or a general inhibitory effect on protein synthesis by monensin. Nevertheless, both treatments decreased the synthesis of gp48. These results suggest that gp48 contains glycans that are either N-linked or possibly O-linked. The glycan moieties of gp48 require further investigation.

Association of gp48 with the virion or dense bodies. To determine if gp48 is associated with the virion, extracellular viral particles were purified as described in Materials and Methods and subjected to Western blot analysis. Antipeptide 2-1 serum detected a protein of 48 kDa in the infected cell lysate (Fig. 9, lane 2) and in purified virions and dense bodies (Fig. 9, lane 3), but this antigen was not detected with preimmune serum (data not shown). These results indicated that gp48 is associated with virions or dense bodies. How-

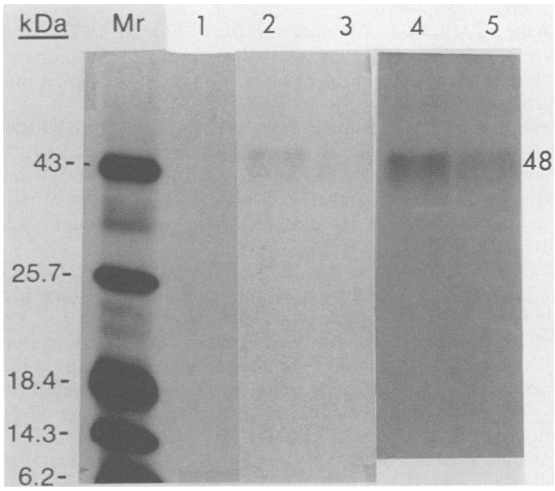

FIG. 7. Immunoprecipitation of a viral glycoprotein. Infected cells were labeled with [^{3}H]glucosamine from 48 to 72 h p.i. Cell lysates (~10^6 cpm) were immunoprecipitated with preimmune serum or antisera against synthetic peptides. Bound proteins were eluted and subjected to SDS-PAGE. Lanes: 1, preimmune serum; 2, antipeptide 2-1 serum; 3, antipeptide 2-2 serum; 4 and 5, longer exposure of lanes 2 and 3; Mr, Protein molecular size standards.

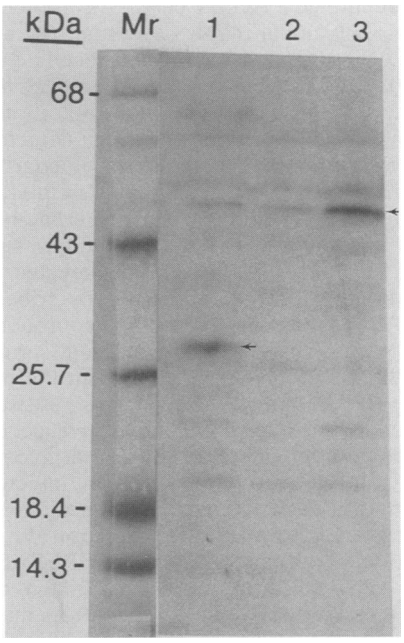

FIG. 8. Effects of tunicamycin and monensin on the processing of the early glycoprotein, gp48. Infected cells were treated with tunicamycin or monensin at 24 h p.i. for 48 h and analyzed by Western blot with antiserum against peptide 2-1 as described in Materials and Methods. Lanes: 1, tunicamycin treated-lysate; 2, monensin-treated lysate; 3, untreated lysate; Mr, Protein molecular size standards.

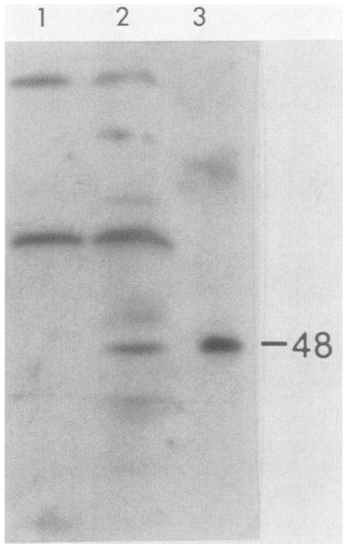

FIG. 9. Immunoblot analysis of a virion-associated glycoprotein. Virions and dense bodies of HCMV were purified as described in Materials and Methods. The virion-associated antigens and lysates of mock-infected or infected cells were analyzed by the Western blot technique with antipeptide 2-1 serum. Lanes: 1, mock-infected cell lysate; 2, cell lysate at 72 h p.i.; and 3, virion- and dense body-associated antigens.

ever, the relative amount of gp48 in virions or dense bodies appears to be low compared with the other viral glycoproteins, such as the gB homolog.

DISCUSSION

We have identified a HCMV early glycoprotein designated gp48 using a polyclonal rabbit antiserum directed against a region of ORF3 encoded by the unspliced RNAs mapped between 0.054 and 0.064 map units. Although ORF1 encoded by the same RNAs has the potential to code for a 10-kDa polypeptide, a viral protein has not been detected by any of the immune sera tested. The gene for glycoprotein gp48 is the first early glycoprotein gene identified in HCMV. However, a potential immediate-early membrane-bound glycoprotein identified only by DNA sequencing was proposed by Kouzarides et al. (16). gp48 was not detected in the three major HCMV glycoprotein complexes described by Gretch et al. (10). Although gp48 is a minor glycoprotein in infected cells and the virions, human convalescent serum does recognize gp48. The involvement of gp48 in the host immune response to HCMV infection remains to be determined.

Even though the RNAs encoding gp48 accumulated to high abundance after infection, this viral glycoprotein was synthesized at relatively low levels in infected cells. In contrast, other HCMV structural viral proteins accumulate to high levels at late times after infection (7, 10, 23, 31). Since deletion of all three upstream ORFS did not fully relieve the in vitro translational repression of the gp48 mRNAs, the presence of the upstream ORFs may not play a major role in gp48 expression. It is more likely that translational repression of gp48 mRNA is due to a secondary structure or some *cis*- or *trans*-acting phenomenon. Computer-aided analysis of the 5′ NCR of gp48 mRNA did not reveal a significant secondary structure. The predicted Gibbs free energy for the 232-nucleotide leader is −72 kcal/mol (1 cal = 4.184 J). Effects of secondary structure on translational efficiency have been demonstrated by insertion of a stable hairpin near the 5′ region of mRNA in vitro (17). *cis*- or *trans*-acting phenomenon has been proposed for a HCMV early transcript (6), but this transcript has not been demonstrated to be mRNA and a viral gene product has not been identified. Whether the translational control of gp48 mRNA observed in vitro also occurs in infected cells is presently unknown.

gp48 is modified by N-linked glycosylation and possibly other posttranslational reactions. O-linked glycosylation is also likely because serine and threonine residues, the acceptor sites for O-linked oligosaccharides, make up 18.9% of the total amino acid composition. A similar structure and hydropathy profile is shared by membrane proteins bearing phosphatidylinositol glycan-linked carboxy termini (19). However, the significance of these modifications to the biological role of the viral glycoproteins requires further investigation.

Computer-aided analysis did not detect any significant amino acid homology between gp48 and other known herpesvirus glycoproteins. Although gp48 is an envelope glycoprotein, it may have early functions in infected cells. For example, the viral glycoprotein could be localized in the nuclear membrane or the intracytoplasmic membranes where the viral glycoprotein could be involved either in translocation of viral proteins between the cytoplasm and nucleus or in stimulation of host cell macromolecular synthesis, respectively.

ACKNOWLEDGMENTS

We thank D. Gretch, T. Hermiston, and M. Leatham for a critical review of the manuscript.

This work was supported by Public Health Service grants HD-19937 (M.F.S.), AI21640 (J.N.), AI07007 (M.B.A.O.), and NS-12428 (M.B.A.O.) from the National Institutes of Health and by grant 1-697 from the National Foundation of the March of Dimes (M.F.S.). J.N. has a Faculty Award from the American Cancer Society.

LITERATURE CITED

1. **Britt, W. J., and L. G. Vugler.** 1989. Processing of the gp55-116 envelope glycoprotein complex (gB) of human cytomegalovirus. J. Virol. **63**:403–410.
2. **Chang, C.-P., C. L. Malone, and M. F. Stinski.** 1989. A human cytomegalovirus early gene has three inducible promoters that are regulated differently at various times after infection. J. Virol. **63**:281–290.
3. **Cranage, M. P., T. Kouzarides, A. T. Bankier, S. Satchwell, K. Weston, P. Tomlinson, B. Barrell, H. Hart, S. E. Bell, A. C. Minson, and G. L. Smith.** 1986. Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. EMBO J. **5**:3057–3063.
4. **Cranage, M. P., G. L. Smith, S. E. Bell, H. Hart, C. Brown, A. T. Bonkier, P. Tomlinson, B. Barrell, and T. C. Minson.** 1988. Identification and expression of human cytomegalovirus (HCMV) glycoprotein H: neutralization of HCMV infectivity with anti-gH monoclonal antibody. J. Virol. **62**:1416–1422.
5. **Farrar, G. H., and P. J. Greenaway.** 1986. Characterization of glycoprotein complexes present in human cytomegalovirus envelopes. J. Gen. Virol. **67**:1469–1473.
6. **Geballe, A. P., R. R. Spaete, and E. S. Mocarski.** 1986. A cis-acting element within the 5′ leader of a cytomegalovirus transcript determines kinetic class. Cell **46**:865–872.
7. **Geballe, A. P., R. R. Spaete, and E. S. Mocarski.** 1986. Regulation of cytomegalovirus late gene expression: γ genes are controlled by posttranscriptional events. J. Virol. **57**:864–874.
8. **Gretch, D. G., R. C. Gehrz, and M. F. Stinski.** 1988. Characterization of a human cytomegalovirus glycoprotein complex (gcl). J. Gen. Virol. **69**:1205–1215.

9. **Gretch, D. G., 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.
10. **Gretch, D. G., B. Kari, L. Rasmussen, R. C. Gehrz, and M. F. Stinski.** 1988. Identification and characterization of three distinct families of glycoprotein complexes present in the envelopes of human cytomegalovirus. J. Virol. **62**:875–881.
11. **Griffiths, P. D., and J. E. Grundy.** 1987. Molecular biology and immunology of cytomegalovirus. Biochem. J. **241**:313–324.
12. **Hermiston, T. W., C. L. Malone, P. M. Witte, and M. F. Stinski.** 1987. Identification and characterization of the human cytomegalovirus immediate-early region 2 gene that stimulates gene expression from an inducible promoter. J. Virol. **61**:3214–3221.
13. **Ho, M.** 1982. Cytomegalovirus biology and infection, p. 1–275. Plenum Publishing Corp., New York.
14. **Hopp, T. P., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**:3823–3828.
15. **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.
16. **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.
17. **Kozak, M.** 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosome. Proc. Natl. Acad. Sci. USA **83**:2850–2854.
18. **Kozak, M.** 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J. Mol. Biol. **196**:947–950.
19. **Low, M. G., and A. R. Saltiel.** 1988. Structural and functional roles of glycosyl-phosphatidylinositol in membranes. Science **239**:268–275.
20. **Mach, M., U. Utz, and B. Fleckenstein.** 1986. Mapping of the major glycoprotein genes of human cytomegalovirus. J. Gen. Virol. **67**:1461–1467.
21. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
22. **Melton, D. A., P. A. Grieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
23. **Meyer, H., A. T. Bankier, M. P. Landini, C. M. Brown, B. G. Barrell, B. Ruger, and M. Mach.** 1988. Identification and prokaryotic expression of the gene coding for the highly immunogenic 28-kilodalton structural phosphoprotein (pp28) of human cytomegalovirus. J. Virol. **62**:2243–2250.
24. **Pereira, L., M. Hoffmann, D. Gallo, and D. Dondero.** 1984. Polymorphism of human cytomegalovirus glycoproteins characterized by monoclonal antibodies. Virology **139**:73–86.
25. **Rapp, F.** 1980. Persistence and transmission of cytomegalovirus. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 16, p. 193–232. Plenum Publishing Corp., New York.
26. **Rasmussen, L., J. Mullenaz, R. Nelson, and T. C. Merigan.** 1985. Viral polypeptides detected by a complement-dependent neutralizing murine monoclonal antibody to human cytomegalovirus. J. Virol. **55**:274–280.
27. **Rasmussen, L., R. Nelson, D. Kelsall, and T. C. Merigan.** 1984. Murine monoclonal antibody to a single protein neutralizes the infectivity of human cytomegalovirus. Proc. Natl. Acad. Sci. USA **91**:876–880.
28. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
29. **Spindler, K. R., D. S. E. Rosser, and A. J. Berk.** 1984. Analysis of adenovirus transforming proteins from early regions 1A and 1B with antisera to inducible fusion antigens produced in *Escherichia coli*. J. Virol. **49**:132–141.
30. **Stinski, M. F.** 1976. Human cytomegalovirus: glycoproteins associated with virions and dense bodies. J. Virol. **19**:594–609.
31. **Stinski, M. F.** 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. J. Virol. **23**:751–767.
32. **Stinski, M. F.** 1978. Sequence of protein synthesis in cells infected by human cytomegalovirus: early and late virus-induced polypeptides. J. Virol. **26**:686–701.
33. **Stinski, M. F.** 1983. Molecular biology of cytomegalovirus. *In* B. Roizman (ed.), The herpesviruses, p. 67–113. Plenum Publishing Corp., New York.
34. **Towbin, H., T. Staehlin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
35. **Wathen, M. W., and M. F. Stinski.** 1982. Temporal patterns of human cytomegalovirus transcription. Mapping the viral RNAs synthesized at immediate early, early, and late times after infection. J. Virol. **41**:462–477.
36. **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.