JOURNAL OF VIROLOGY, Nov. 1991, p. 6144–6153
0022-538X/91/116144-10$02.00/0

Vol. 65, No. 11

Alternate Promoter Selection within a Human Cytomegalovirus Immediate-Early and Early Transcription Unit (UL119–115) Defines True Late Transcripts Containing Open Reading Frames for Putative Viral Glycoproteins

MARY P. LEATHAM, PAM R. WITTE, AND MARK F. STINSKI*

Department of Microbiology, School of Medicine, The University of Iowa, Iowa City, Iowa 52242

Received 28 May 1991/Accepted 14 August 1991

The human cytomegalovirus open reading frames (ORFs) UL119 through UL115 (UL119–115) are located downstream of the immediate-early 1 and 2 transcription units. The promoter upstream of UL119 is active at all times after infection and drives the synthesis of a spliced 3.1-kb mRNA. The viral mRNA initiates in UL119, contains UL119–117 and UL116, and terminates just downstream of UL115. True late transcripts that are detected only after viral DNA synthesis originate from this transcription unit. True late mRNAs of 2.1 kb, containing ORFs UL116 and UL115, and 1.2 kb, containing ORF UL115 only, are synthesized. The true late viral mRNAs are 3′ coterminal with the 3.1-kb mRNA. This transcription unit is an example of late promoters nested within an immediate-early–early transcription unit. The gene products of UL119–117, UL116, and UL115 are predicted to be glycoproteins. Efficient expression of the downstream ORFs at late times after infection may be related to alternate promoter usage and downstream cap site selection.

Human cytomegalovirus (HCMV) can cause serious morbidity and mortality in individuals with immature or compromised immune systems (1, 12). The virus infects a variety of different human cell types that range from nonpermissive to permissive for virus replication. Among different nonpermissive cell types, the extent of immediate-early (IE) and early viral gene expression varies, but late viral gene expression does not occur (25, 40). Late viral gene expression is restricted to certain cell types in human hosts (2, 8, 26, 31) or in cell culture (25, 40). HCMV is a species-specific virus because late gene expression occurs only in human cells. The virus is classified as a betaherpesvirus.

In permissive cells, the viral genes are transcribed in temporally regulated phases. The first phase, referred to as the IE phase, does not require preceding viral protein synthesis. Cellular transcription factors recognize enhancer-containing elements in the viral genome and promote transcription. HCMV is unique among the known herpesviruses because the IE gene products are both phosphorylated proteins and putative glycoproteins (reviewed in reference 40).

The second phase, referred to as the early phase, requires viral protein synthesis for significant transcription of the early genes but does not require viral DNA synthesis. It has been proposed, on the basis of transient transfection experiments, that the IE gene products of HCMV transactivate the HCMV early promoters (reviewed in references 19, 25, 34, 40, and 41). The early genes specify a complicated array of different types of products which include the enzymes necessary for viral DNA synthesis and nonstructural as well as structural viral proteins and glycoproteins (reviewed in references 33 and 40).

The third phase of viral gene expression, referred to as the late phase, requires viral DNA synthesis. Defining this phase has been difficult for HCMV because most regions of the viral genome, including structural protein genes, are transcribed under early conditions (7, 9, 23, 45). Recently, a few true late genes of HCMV have been identified (17, 29, 35). In the presence of a drug that inhibits viral DNA synthesis, or at the nonpermissive temperature for a virus with a temperature-sensitive DNA synthesis phenotype, no transcription was detected. In general, true late transcription of DNA viruses requires either an undefined *cis*-acting modification of the viral DNA template or late phase-specific *trans*-acting factors.

The initiation of HCMV DNA replication appears to activate at least three different types of late promoters. (i) There are promoters that determine a late transcription initiation site or cap site upstream of a 5′ end of an early transcription unit. For example, the early gene (UL4) that codes for a glycoprotein designated gp48 and the early gene (UL36) that codes for a DNA binding protein designated DBP52 (ICP36) have two early cap sites and one late cap site (4, 5, 18). (ii) There are the conventional types of late promoters located 5′ to a late gene. For example, the promoter for the major capsid protein gene (UL86) is located 5′ to the gene and functions only at late times after infection (29). (iii) There are late promoters nested within an IE or early transcription unit. These late promoters are contained entirely within the transcription unit and promote the synthesis of a late viral mRNA that is 3′ coterminal. The late viral mRNA may permit the expression of an open reading frame (ORF) not efficiently expressed at early times. Alternatively, a truncated version of a early gene product may be made, as recently described by Stenberg et al. (35).

We have investigated the UL119–115 transcription unit. This region of the viral genome was previously designated IE3 because transcription was detected in this region for the Towne strain under IE conditions (42). Since the functions of the viral gene products from this region are not known, we use the term IE to describe a transcription unit that does not require preceding viral protein synthesis. We analyzed IE, early, and late transcription in this region and characterized the structures of the transcripts at various times after infec-

* Corresponding author.

tion. One promoter is active under IE conditions, i.e., in the presence of inhibitors of protein synthesis. The promoter remains active throughout infection. This promoter drives the synthesis of a 3.1-kb mRNA that could code for the expression of UL119, UL118, part of UL117, UL116, and UL115. At late times, two internal promoters are activated and drive the synthesis of a 2.1-kb mRNA and a 1.2-kb mRNA that could code for the expression of UL116–115 and UL115, respectively. The IE-early and late mRNAs are 3′ coterminal. This transcription unit and its associated promoters are an example of late promoters nested within an IE or early transcription unit. This type of HCMV transcription unit allows for amplification of downstream transcription, which could result in differential expression of HCMV putative glycoproteins.

MATERIALS AND METHODS

Cell culture and virus propagation. The growth of human foreskin fibroblast cells and the propagation of the HCMV Towne strain have been described previously (39).

Enzymes. Restriction endonucleases were obtained from either Bethesda Research Laboratories, Inc., Gaithersburg, Md., or New England BioLabs, Inc., Beverly, Mass. T4 polynucleotide kinase and mung bean nuclease were obtained from Pharmacia, Inc., Piscataway, N.J. T7 and SP6 polymerases, T4 ligase, and Klenow DNA polymerase were obtained from Boehringer Mannheim Biochemicals, Indianapolis, Ind. Exonuclease VII was obtained from Bethesda Research Laboratories. RNase A and RNase T_1 were obtained from Sigma Chemical Co., St. Louis, Mo. All enzymes were used according to the manufacturer's specifications.

Plasmid constructions. All DNA clones (summarized in Fig. 10) were subclones of the *Xba*I E fragment of HCMV Towne described previously (42). Plasmid DNAs were purified as described previously (43).

Templates for RNA synthesis (summarized in Fig. 10) were obtained by cloning viral DNA restriction endonuclease fragments into vector pGEM3blue (Promega Biotec, Madison, Wis.).

DNA sequence. Base pairs from the HCMV AD169 strain genomic sequence were numbered according to the sequence of the complete genome submitted to EMBL by Chee et al. (6; accession number X17403). The potential ORFs for the entire genome have been designated from left to right for each component of the viral genome (6). The nucleotide sequence of the Towne strain of HCMV was also determined by the method of Maxam and Gilbert (21). The sequences between 0.709 and 0.728 map units for the two strains are very similar. Within this 4.36-kb region of DNA, there are some bases missing and some restriction endonuclease polymorphism. For example, 21 bp are missing from UL119 in strain AD169. For the Towne strain versus the AD169 strain, there are 1 versus 3 *Apa*I sites, 4 versus 3 *Mlu*I sites, 10 versus 11 *Ava*II sites, 11 versus 12 *Taq*I sites, 14 versus 13 *Dde*I sites, 2 versus 1 *Nco*I sites, and 1 versus 0 *Nde*I sites, respectively. An ORF analysis of the Towne strain demonstrated the same major ORFs reported for the AD169 strain (6).

Preparation of RNAs. The isolation of cytoplasmic RNA from mock-infected and infected cells has been described previously (36). The cells were treated with 100 to 200 μg of cycloheximide per ml 30 min before infection, during the virus adsorption period, and until the IE RNA was isolated at 12 h postinfection (p.i.). Early RNA was isolated 48 h p.i. in the presence of 200 μg of phosphonoacetic acid per ml. Under these conditions, we do not detect cytotoxicity due to the phosphonoacetic acid treatment when using primary foreskin fibroblast cells. Late RNA was isolated 72 h p.i. in the absence of any drug treatment.

Northern (RNA) blots. RNAs were fractionated in formaldehyde-agarose gels blotted onto nylon membranes as recommended by the supplier (GeneScreen Plus; DuPont, NEN Research Products, Boston, Mass.). Hybridization was with either ^{32}P-labeled DNA probes labeled by the random primer method as described previously (9) or [^{32}P]UTP-labeled RNA probes prepared from linearized DNA templates as described by Krieg and Melton (15). After hybridization, the blots were washed with 2× SSC (0.3 M NaCl, 0.03 M sodium citrate) at room temperature and then with 0.1× SSC containing 0.1% sodium dodecyl sulfate at 72°C. *Escherichia coli* rRNA standards were purchased from Pharmacia.

Structural analysis of RNA. Determination of RNA structure by the single-strand-specific nuclease mapping technique (3) has been described previously (37, 38). Gel-purified restriction endonuclease DNA fragments were used to prepare end-labeled probes. Labeling at the 5′ end was accomplished with polynucleotide kinase and [γ-^{32}P]ATP as described by Maniatis et al. (20). Labeling at the 3′ end was done with Klenow enzyme and α-^{32}P-labeled deoxynucleoside triphosphates according to the method of Maniatis et al. (20). End-labeled DNA probes were hybridized with 20 μg of denatured cytoplasmic RNA at 50 to 55°C for 5 h or overnight and then treated with either 2 U of mung bean nuclease per μg of RNA or 0.1 U of exonuclease VII per μg of RNA at 37°C for 60 min. The nuclease-protected probes were subjected to electrophoresis in 1.5% agarose denaturing alkaline gels as described previously (37, 38). The gels were dried and submitted to autoradiography.

RNase protection. Twenty micrograms of cytoplasmic RNA was analyzed. [^{32}P]RNA probe synthesis, hybridization, and RNase digestion with RNase A and RNase T_1 were performed as described by Krieg and Melton (15). RNase digestions took place at 30°C for 1 h. The protected fragments were subjected to electrophoresis in 6% polyacrylamide–urea gels. The bands were detected by autoradiography.

RESULTS

RNA size classes from UL119–115 at various times after infection. To investigate the RNA size classes from this region of the viral genome, a ^{32}P-labeled DNA probe originating from a region near the 3′ end of the transcription unit (designated A3) was hybridized to Northern blots as described in the Materials and Methods. A 3.1-kb mRNA was detected at early and late times after infection (Fig. 1, lanes 4 to 8). Two new RNA size classes of 2.1 and 1.2 kb were detected at late times (lanes 7 and 8). In addition, there was a large RNA of 4.2 kb that was in low abundance early and more abundant at late times. Cross-reactivity of the ^{32}P-labeled DNA probe with the smaller size class of rRNA was detected with both mock-infected-cell and infected-cell RNA.

To further characterize the different size classes of viral mRNAs, a series of different ^{32}P-labeled DNA or antisense RNA probes spanning the transcription unit (see Fig. 10) were hybridized to Northern blots of mock-infected- or infected-cell RNA harvested at late times after infection. RNAs of 4.2 and 3.1 kb were detected by all of the ^{32}P-labeled DNA probes designated A1 through A6, B1, and

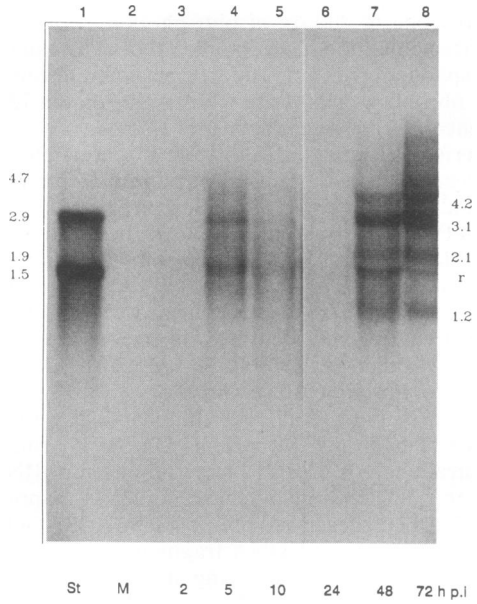

FIG. 1. Viral RNAs from the UL119–115 transcription unit detected at various times after infection. Cytoplasmic RNAs were isolated from early to late times after infection, blotted onto nylon filters, and hybridized with ^{32}P-labeled probe A3 (see Fig. 10) as described in Materials and Methods. Lane 1, ^{32}P-labeled *E. coli* rRNA molecular size standards (St); lane 2, mock-infected (M)-cell RNA. The other lanes contain RNAs isolated at the times p.i. indicated below the gel. Lanes 1 through 5 are a 7-day exposure and lanes 6 through 8 are a 1-day exposure of the autoradiogram. r, rRNA. Sizes in kilobases are given in the margins.

B2 (Fig. 2) and A7 (data not shown). In contrast, the 2.1-kb mRNA was detected by only the A1 through A4, B1, and B2 probes. The 1.2-kb mRNA was detected by only A1 through A3 (Fig. 2). These data indicated that the 4.2- and 3.1-kb RNAs originated from the entire transcription unit but the 2.1- and 1.2-kb mRNAs originated from only the 3′ half of the transcription unit (summarized in Fig. 10).

Antisense RNA probes were used to determine the direction of transcription and also to better map the origins of the different viral transcripts. ^{32}P-labeled RNA probes were prepared and hybridized to Northern blots of mock-infected-cell or infected-cell RNA harvested at IE, early, or late times after infection as described in Materials and Methods. Northern blot analysis indicated that all viral RNAs originated from the same strand of the viral DNA. The 4.2- and 3.1-kb RNAs were detected by probes R1 through R4 (Fig. 3), R5 and R6 (data not shown), and R8 and R9 (Fig. 3). The 4.2-kb mRNA was detected by R7, but the 3.1-kb mRNA was not (data not shown). The R3 probe detected the 3.1-kb mRNA in IE, early, and late infected-cell RNA (Fig. 3). In contrast, the 2.1- and 1.2-kb mRNAs were detected only at late times after infection. Occasionally, the 2.1-kb mRNA did not separate well from rRNA in 1.5% agarose denaturing gels. Again, the 2.1- and 1.2-kb mRNAs originated from the 3′ half of the transcription unit (summarized in Fig. 10). The 4.2-kb RNA was detected primarily at late times after infection.

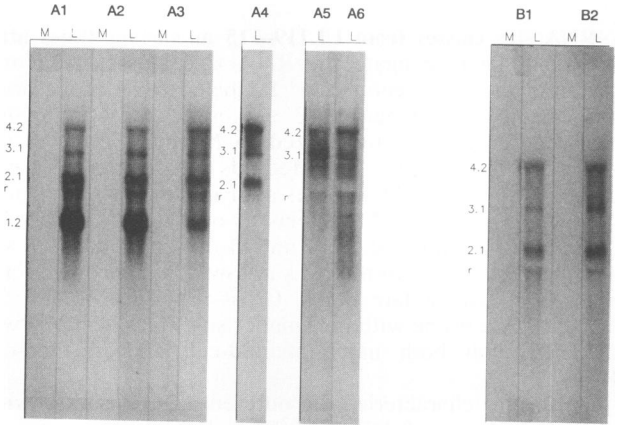

FIG. 2. Viral RNAs from the UL119–115 transcription unit detected at late times after infection. ^{32}P-labeled DNA probes A1 through A6, B1, and B2 (see Fig. 10) were hybridized to Northern blots of mock-infected-cell RNA (M) or infected-cell RNA harvested at late times after infection (L) as described in Materials and Methods. RNA size classes are indicated in kilobases. r, rRNA.

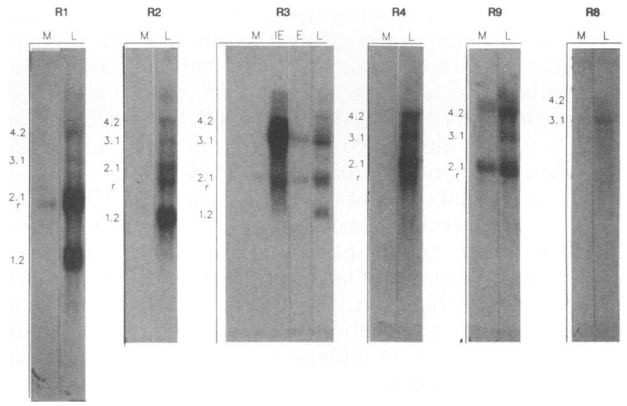

FIG. 3. Direction of transcription and fine mapping of viral RNAs from the UL119–115 transcription unit detected at IE, early (E), or late (L) times after infection. ^{32}P-labeled antisense RNA probes (see Fig. 10) were hybridized to Northern blots of mock-infected-cell RNA or infected-cell RNA harvested under IE, early, or late conditions as described in Materials and Methods. RNA sizes are indicated in kilobases. r, rRNA.

Mapping the 3′ ends of the transcription units at various times after infection. The Northern blot analysis suggested that at late times after infection, there were at least three different size classes of mRNAs with coterminal 3′ ends. To map the 3′ ends of these viral mRNAs, a DNA probe which extended downstream was uniquely end labeled at the *Sal*I site (see Fig. 10) and hybridized to mock-infected cellular RNA, IE RNA, or RNA harvested at various times after infection. The hybrids were treated with either mung bean nuclease or exonuclease VII, and the protected ^{32}P-labeled DNA probe was subjected to electrophoresis on denaturing alkaline gels as described in Materials and Methods. After mung bean nuclease treatment, protected probes of 1,400 and 400 nucleotides (nt) were detected (Fig. 4A). After exonuclease VII treatment, only the 1,400-nt protected band was detected (Fig. 4B). The amount of the 400-nt protected band at various times after infection was always smaller than the amount of IE RNA. These data indicated the presence of a coterminal 3′ end approximately 1,400 nt downstream from the *Sal*I site and a donor splice site approximately 400 nt downstream from the *Sal*I site. The donor splice site was highly used under IE conditions but not at later times after infection. Examination of the Towne and AD169 strain DNA

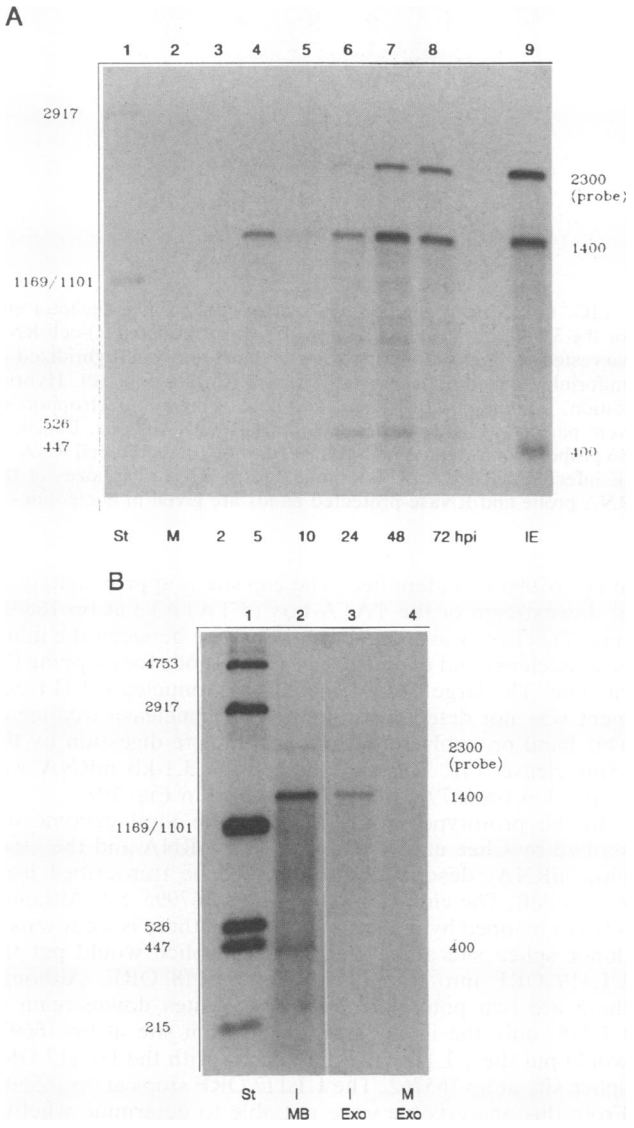

FIG. 4. Mapping the 3′ ends of viral RNAs downstream of UL115 at various times after infection. Cytoplasmic RNAs were harvested from mock-infected (M) or infected (I) cells under IE conditions or at various times after infection. (A) A DNA probe uniquely 3′ end labeled at the *Sal*I site (see Fig. 10) was hybridized to RNA, digested with mung bean nuclease, and subjected to alkaline gel electrophoresis as described in the Materials and Methods. Lane 1, molecular size standards (St); lane 2, mock-infected-cell RNA. Lanes 3 to 8 contain RNAs harvested at the following times p.i.: 3, 2 h; 4, 5 h; 5, 10 h; 6, 24 h; 7, 48 h; 8, 72 h. Lane 9, IE RNA. (B) Mock-infected- or infected-cell RNAs harvested under IE conditions were hybridized to the DNA probe uniquely 3′ end labeled at the *Sal*I site (see Fig. 10) and treated with either mung bean nuclease (MB) or exonuclease VII (Exo). Lanes: 1, molecular size standards (St); 2, infected-cell RNA plus MB; 3, infected-cell RNA plus Exo; 4, mock-infected-cell RNA plus Exo. The sizes of DNA standards and protected bands are given in nucleotides in both panels.

sequences indicated a consensus cleavage and polyadenylation signal at bp 163683 and a donor splice site at bp 164661 (summarized in Fig. 10).

Mapping the structure of the 3.1-kb mRNA. Since the 3.1-kb mRNA was the major viral RNA size class detected under IE conditions, IE RNA was hybridized to different ^{32}P-labeled DNA probes (summarized in Fig. 10) as described in Materials and Methods. The probes will be discussed starting from the 3′ end of the transcription unit and progressing to the 5′ end. After mung bean nuclease treatment, a 770-nt and a 750-nt protected band for the probes uniquely end labeled at the 5′ end of the *Ava*II and *Ava*I sites were detected, respectively (Fig. 5A and B). After exonuclease VII treatment, the entire DNA probes were protected. The minor protected bands may be artifactual or alternate splices that were not investigated further. These data indicated the presence of a splice acceptor site approximately 770 nt upstream of the *Ava*II site and 750 nt upstream of the *Ava*I site. DNA sequence analysis identified a splice acceptor site at bp 164607 (summarized in Fig. 10).

After mung bean nuclease treatment, an 850-nt protected band for the probe uniquely end labeled at the 5′ end of the *Dde*I site was detected (Fig. 5C). After exonuclease VII treatment, a 770-nt protected band for a probe uniquely end labeled at the 5′ end of the *Mlu*I site was also detected (Fig. 5D). In each case, the amount of probe protected by infected-cell RNA was significantly greater than mock-infected-cell RNA. These data suggested that either a 5′ end or a splice acceptor site is located approximately 850 nt upstream of the *Dde*I site and 770 nt upstream of the *Mlu*I site. DNA sequence analysis of this region of both the AD169 and Towne strain DNAs detected both splice acceptor sites and potential TATA boxes.

To more accurately map the 5′ end upstream of the *Dde*I and *Mlu*I sites, an RNase protection experiment with mock-infected, IE, and late RNA and the ^{32}P-labeled antisense RNA probe designated R6 (see Fig. 10) was done as described in Materials and Methods. A unique 285-nt RNase-protected band was detected (Fig. 6). There is a splice acceptor site 285 bp away from the *Eco*RI site (see Fig. 10) at bp 165762. The 397-nt band represents protection from the 4.2-kb RNA, which is found in small amounts under IE conditions and in larger amounts at late times after infection (Fig. 6). Therefore, the DNA probes uniquely end labeled at the *Dde*I or *Mlu*I site and the antisense RNA probe R6 suggested the presence of a large intron upstream of the splice acceptor site at bp 165762 and the presence of an exon in the UL118 region. An exonuclease VII protection experiment with a probe uniquely end labeled at the *Sst*I site in UL115 also suggested the presence of a 5′ end upstream of UL118 (data not shown).

In order to define the exons upstream, a 5′-end-labeled probe within the UL118 region was selected. After mung bean nuclease treatment, a 450-nt protected band for the probe uniquely end labeled at the 5′ end of the *Taq*I site was detected (Fig. 5E). After exonuclease VII treatment, only a 1,000-nt band was detected (Fig. 5E). These data indicated the presence of a splice acceptor site approximately 450 nt upstream of the *Taq*I site and a 5′ end approximately 1,000 nt upstream of the *Taq*I site. DNA sequence analysis identified a splice acceptor site at bp 167474 (summarized in Fig. 10).

After mung bean nuclease or exonuclease VII treatment, a 400-nt protected band was detected for a DNA probe in the UL119 region uniquely end labeled at the *Ava*II site (Fig. 5F). The 5′ end of the 3.1-kb mRNA was approximately 400 nt upstream of the *Ava*II site.

To map the 5′ end and to identify the location of a potential TATA box, a DNA probe uniquely 5′ end labeled at the *Ava*I site in the UL119 region was hybridized to mock-infected-cell or IE RNA. After mung bean nuclease or exonuclease VII treatment, the cap site of the 3.1-kb mRNA

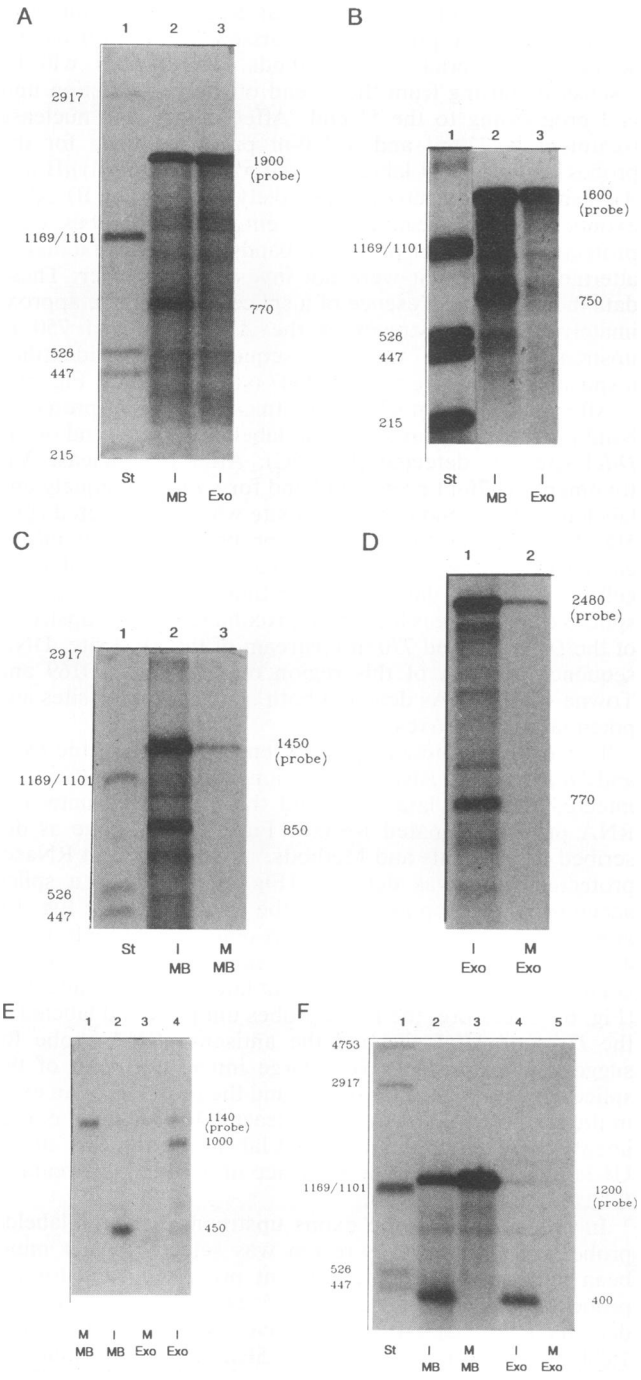

FIG. 5. Mapping of the 5′ ends of the viral 3.1-kb RNA from the UL119–115 transcription unit. Mock-infected-cell RNA (M) or infected-cell RNA (I) harvested under IE conditions was hybridized to several different end-labeled DNA probes (see Fig. 10). The restriction endonuclease sites and abbreviations of restriction endonuclease sites used to prepare uniquely end-labeled DNA are indicated at the top and in the middle of Fig. 10, respectively. Hybrids were treated with mung bean nuclease (MB) or exonuclease VII (Exo), and the protected DNA fragments were subjected to alkaline gel electrophoresis as described in Materials and Methods. (A) DNA probe uniquely 5′ end-labeled at the *Ava*II site. Lanes: 1, molecular size standards (St); 2, infected-cell RNA plus MB; 3, infected cell RNA plus Exo. (B) DNA probe uniquely 5′ end labeled at the *Ava*I site. Lanes: 1, molecular size standards; 2, infected-cell RNA plus MB; 3, infected-cell RNA plus Exo. (C) DNA probe uniquely 5′ end labeled at the *Dde*I site. Lanes: 1, molecular size standards; 2, infected-cell RNA plus MB; 3, mock-infected-cell RNA plus MB. (D) DNA probe uniquely 5′ end labeled at the *Mlu*I site. Lanes: 1, infected-cell RNA plus Exo; 2, mock-infected-cell RNA plus Exo. (E) DNA probe uniquely 5′ end labeled at the *Taq*I site. Lanes: 1, mock-infected-cell RNA plus MB; 2, infected-cell RNA plus MB; 3, mock-infected-cell RNA plus Exo; 4, infected-cell RNA plus Exo. (F) DNA probe uniquely 5′ end labeled at the *Ava*II site. Lanes: 1, molecular weight standards; 2, infected-cell RNA plus MB; 3, mock-infected-cell RNA plus MB; 4, infected-cell RNA plus Exo; 5, mock-infected-cell RNA plus Exo. The sizes of DNA standards and protected bands are given in nucleotides in all panels.

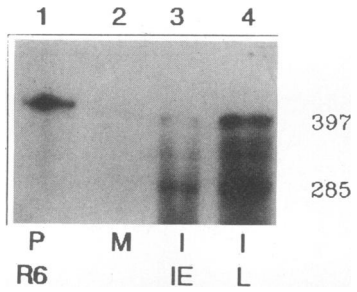

FIG. 6. RNase protection and mapping of a splice acceptor site for the 3.1-kb mRNA. Mock-infected (M)- or infected (I)-cell RNA harvested at IE or late (L) times after infection was hybridized to uniformly labeled antisense RNA probe R6 (see Fig. 10). Hybridization, treatment with RNase, and denaturing gel electrophoresis were performed as described in Materials and Methods. Lanes: 1, R6 probe (P) not treated with RNase; 2, mock-infected-cell RNA; 3, IE infected-cell RNA; 4, late infected-cell RNA. The sizes of the RNA probe and RNase-protected bands are given in nucleotides.

at bp 167995 was identified. The cap site is approximately 22 nt downstream of the TATA box (5′TATA3′) at bp 168020 (Fig. 7). There was a one-base difference between the mung bean nuclease and exonuclease VII methods for mapping the cap site. The large band detected by exonuclease VII treatment was not detected by mung bean nuclease treatment. This band probably represents incomplete digestion by the exonuclease VII. The cap site for the 3.1-kb mRNA was mapped at bp 167995 ± 1 (summarized in Fig. 10).

In the prototype arrangement of the viral genome described by Chee et al. (6), the 3.1-kb mRNA and the other viral mRNAs described below would be transcribed from right to left. The viral RNA starts at bp 167995 ± 1. Although it is not mapped by nuclease protection, there is a consensus donor splice site at bp 167563. This splice would put the UL119 ORF into frame with the UL118 ORF. Although there are two potential donor splice sites downstream of UL118, only the immediate downstream site at bp 166978 would put the UL118 ORF into frame with the UL117 ORF splice site at bp 165762. The UL117 ORF stops at bp 165590. From this analysis we were not able to determine whether the second donor splice site would place UL119 and UL118 into frame with downstream ORFs such as UL116 and UL115. The viral RNA terminates at the common 3′ end at bp 163683 (summarized in Fig. 10).

Mapping the 5′ ends of the 2.1- and 1.2-kb true late mRNAs by RNase protection. The Northern blot analysis of true late RNAs suggested that the 2.1- and 1.2-kb mRNAs had 5′ ends within the 3.1-kb transcription unit. Two antisense RNA probes were selected for finer mapping of these 5′ ends. Probe R6 spans two potential TATA boxes and an initiator-

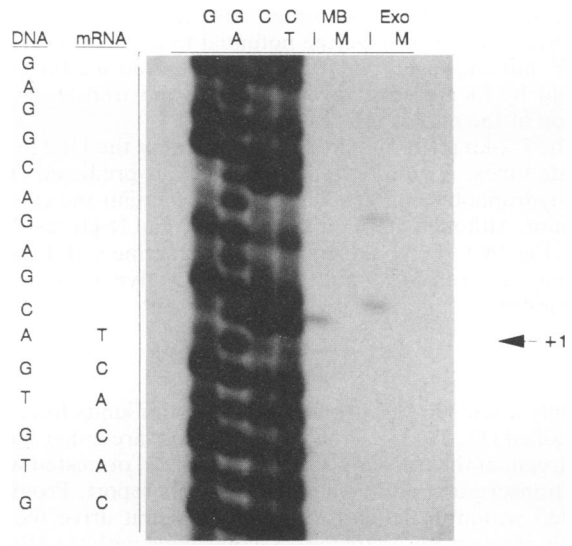

FIG. 7. Mapping the cap site of the 3.1-kb mRNA. Mock-infected (M)- or infected (I)-cell RNA harvested under IE conditions was hybridized to a DNA probe uniquely 5′ end labeled at the *Ava*I site and treated with mung bean nuclease (MB) or exonuclease VII (Exo). The protected DNAs were subjected to electrophoresis adjacent to a Maxam and Gilbert (21) chemical cleavage analysis of the 5′-end-labeled DNA probe as described in Materials and Methods. The approximate location of the 5′ end of the viral IE RNA is designated +1.

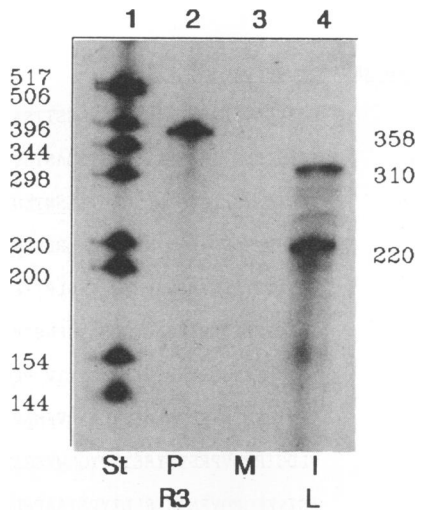

FIG. 8. Mapping the 5′ end of the true late 1.2-kb mRNA by RNase protection. Mock-infected (M)- or infected (I)-cell RNA harvested at late times after infection was hybridized to uniformly labeled antisense RNA probe R3 (see Fig. 10). Hybridization, treatment with RNase, and denaturing gel electrophoresis were performed as described in Materials and Methods. Lanes: 1, molecular size standards (St); 2, probe (P) R3 not treated with RNase; 3, mock-infected-cell RNA; 4, late infected-cell RNA. The size of the ^{32}P-labeled RNA probe reflects transcription from a chimera of viral and vector DNA template. The sizes of the DNA standards, the RNA probe, and the RNase-protected bands are given in nucleotides.

like sequence (5′CCGCGTTCTGGACT3′) (32). Probe R3 spans one potential TATA box, according to the DNA sequence (6). After RNase treatment, a 285-nt protected band was detected with both IE and late infected-cell RNA and the ^{32}P-labeled R6 probe (Fig. 6). This defined an upstream splice acceptor site as described above. Lower protected bands were difficult to identify because the ^{32}P-labeled RNA probe is AU rich and, consequently, RNase treatment caused cleavage of the RNA-RNA hybrids and smearing in the lanes (Fig. 6, lanes 3 and 4). There was no band to suggest that the initiatorlike sequence was used. There was a band indicating that the TATA box at bp 165590 was used at late times to determine the 5′ end of the 2.1-kb mRNA (data not shown). It was not possible to clearly demonstrate the 5′ end or cap site of the 2.1-kb mRNA by RNase protection assay. Primer extension to map the 5′ end of the 2.1-kb mRNA was not done. Therefore, the cap site of the RNA was tentatively positioned downstream of the TATA box at bp 165590 because an unspliced transcript from this promoter would generate an RNA of 2.1 kb after cleavage and polyadenylation at the common 3′ end at bp 163683. With a 5′ end at bp 165590, the 2.1-kb mRNA would have a cap site just upstream of the UL116 ORF (proposed in Fig. 10).

After RNase treatment, a 220-nt protected band was detected with the ^{32}P-labeled R3 probe (Fig. 8, lane 4). There is no splice acceptor site 220 bp away from the *Mlu*I site (see Fig. 10). Therefore, the 5′ end is a potential cap site at bp 164593 which is 24 bases downstream of a TATA box (5′TATATAA3′) at bp 164621. A cap site downstream of this TATA box would generate an RNA of approximately 1.2 kb after cleavage and polyadenylation at the common 3′ end. The RNase protection experiment detected little splicing of the downstream intron at late times. The splice acceptor site is 234 bp upstream of the *Mlu*I site. If significant splicing occurred at late times, the splice acceptor site would have defined an intense band at 234 nt (Fig. 8, lane 4). The 310-nt band represents protection of the antisense RNA probe by the 3.1-kb mRNA containing the downstream intron. The lack of splicing of the downstream intron at late times after infection is in agreement with the data in Fig. 4A. The 1.2-kb mRNA would have a cap site just upstream of the translation initiation site in UL115 at bp 164530, as summarized in Fig. 10.

Predicted viral gene products. The ORFs in this region of the viral genome were predicted to be glycoproteins by Chee et al. (6). The 3.1-kb IE, early, and late mRNA is predicted to express UL119, UL118, and part of UL117. The first ATG in UL119 is at bp 167983 (see Fig. 10). The splices summarized in Fig. 10 put UL119 into frame with UL118 and UL117. A stop codon is encountered in UL117 at bp 165590. The 3.1-kb mRNA is predicted to code for a protein of 54.3 kDa. The predicted viral protein has hydrophobic regions toward the amino and carboxyl termini and 13 potential N-glycosylation sites (Fig. 9A). Therefore, it is predicted that the 3.1-kb mRNA codes for a viral glycoprotein at early and late times after infection. Although the 3.1-kb mRNA would have the potential for expression of the UL116 or UL115 ORF, these ORFs are positioned downstream of the translational stop codon at the end of UL117. Whether alternate splicing places UL119–118 into frame with UL116 or UL115 was not determined in this investigation.

The 2.1-kb mRNA is predicted to express the UL116 ORF at late times after infection. A viral protein of 39.1 kDa is predicted. The amino acid sequence shows that there are hydrophobic amino acids spanning the amino and carboxyl termini and 13 potential N-glycosylation sites (Fig. 9B). Therefore, it is predicted that the 2.1-kb mRNA codes for a

A. 3.1 kb mRNA product of UL119-117:

MCSVLAIALVVALLGDMHPGVKSSTTSAVTSPSNTTVTST
TSISTSNNVTSAVTTTVQTSTSSASTSVIATTQKEGHLYT
VNCEASYSHDQVSLNATCKVILLNNTKNPDILSVTCYART
↓
DCKGPFTQVGYLSAFPPDNEGKLHLSYNATAQELLISGLR
PQETTEYTCSFFSWGRHHNATWDLFTYPIYAVYGTRLNAT
TMRVRVLLQEHEHCLLNGSSLYHPNSTVJLJQGDQLIPPW
NISNVTYNGQRLREFVFYLNGTYTVVRLHVQIAGRSFTTT
↓
YVFIKSDPLFEDRLLAYGVLAFLVFMAVATFKFFHQDPNR
VLDCIRPVVPRSTSYHETGVYQMWVSGATKKDLFDAVTLC
ASIVEKQPDVFNINVSLLTYPSIAAPHLPLYNEFTSFRLP
TS

B. 2.1 kb mRNA product of UL116:

MKRRRRWRGWLLFPALCFCLLCEAVETNATTVTSTTAAAA
TTNTTVATTGTTTTSPNVTSTTSNTVTTPTTVSSVSNLTS
STTSIPISTSTVSGTRNTGNNNTTTIGTNATSPSPSVSIL
TTVTPAATSTISVDGVVTASDYTPTFDDLENITTTRAPTR
PPAQDLCSHNLSIILYEEESQSSVDIAVDEEEPELEDDDE
YDELWFPLYFEAECNRNYTLHVNHSCDYSVRQSSVSFPPW
RDIDSVTFVPRNLSNCSAHGLAVIVAGNQTWYVNPFSLAH
LLDAIYNVLGIEDLSANFRRQLAPYRHTLIVPQT

C. 1.2 kb mRNA product of UL115:

MCRRPDCGFSFSPGPVVLLWCCLLLPIVSSVAVSVAPTAA
EKVPAECPELTRRCLLGEVFQGDKYESWLRPLVNVTRRDG
PLSQLIRYRPVTPEAANSVLLDDAFLDTLALLYNNPDQLR
ALLTLLSSDTAPRWMTVMRGYSECGDGSPAVYTCVDDLCR
GYDLTRLSYGRSIFTEHVLGFELVPPSLFNVVVAIRNEAT
RTNRAVRLPVSTAAAPEGITLFYGLYNAVKEFCLRHQLDP
PLLRHLDKYYAGLPPELKQTRVNLPAHSRYGPQAVDAR

FIG. 9. Amino acid sequences of the predicted viral glycoprotein products encoded by the various mRNAs from the UL119–115 transcription unit. (A) 3.1-kb mRNA product encoded by UL119, UL118, and part of UL117. (B) 2.1-kb mRNA UL116 product. (C) 1.2-kb mRNA UL115 product. The hydrophobic region at the N terminus (potential signal sequence) and the hydrophobic stretch near the C terminus (potential transmembrane region) for each protein are underlined by a dashed line. Possible N-glycosylation sites (NXS or NXT, where X is any residue) are underlined. The vertical arrows indicate the junctions between different exons. The hydrophilicity was determined according to the method of Kyte and Doolittle (16) by averaging over a window of 7.

viral glycoprotein at late times after infection. The 2.1-kb mRNA would also have the potential to express the UL115 ORF, but once again the translation initiation site for UL115 would be located just downstream of the translation stop codon at the end of UL116.

The 1.2-kb mRNA is predicted to express the UL115 ORF at late times. A viral protein of 34.8 kDa is predicted. There are hydrophobic amino acids toward the amino and carboxyl termini. Although there is only one potential N-glycosylation site (Fig. 9C), the viral protein is rich in serine and threonine amino acid residues, and therefore O glycosylation is a possibility.

DISCUSSION

Only a few HCMV true late transcription units have been identified (17, 29, 35). True late promoters are either located upstream of the transcription unit (4, 18, 29) or nested within the transcription unit as described in this report. Promoters nested within an IE-early transcription unit drive the synthesis of late mRNAs that are 3′ coterminal with the IE-early mRNAs. These true late mRNAs either code for a truncated version of the early viral protein, as described by Stenberg et al. (35), or allow for efficient expression of an ORF, as proposed in this report.

HCMV early promoters are activated in *trans* by the IE2 protein (4, 11). The viral IE1 protein acts synergistically as previously described (reviewed in references 19, 25, 34, 40, and 41). Why true late promoters are not activated until after the initiation of viral DNA replication is not understood. Viral DNA replication may modify the template, thus allowing recognition of the true late promoter. Alternatively, inhibitory proteins of viral or cellular origin may be present on the viral DNA template and either removed by DNA replication or titrated out by the amplification of the template. It is also possible that there are HCMV-specified or -induced proteins that activate true late promoters in *trans*.

Maximum expression of the adenovirus true late promoter for the IX gene requires viral DNA replication plus the E1b gene product (44). Maximum expression of vaccinia virus true late genes requires three viral proteins with properties similar to transcriptional regulatory proteins (13). The vaccinia virus regulatory proteins are expressed immediately after viral DNA replication, and therefore they are referred to as intermediate viral proteins. With herpes simplex virus, an alphaherpesvirus, true late or $\gamma 2$ gene expression requires DNA replication and is influenced by an IE or α gene product designated ICP27 (10, 22, 24, 27, 28, 30). Whether this viral transactivator functions to stimulate transcription or to stabilize the viral RNA or both requires further investigation.

For HCMV, true late promoters nested within an IE-early transcription unit were not efficiently utilized until after viral DNA replication, even though the late promoters are included within active transcription units. The promoter containing a typical TATA box upstream of UL119 is active at all times after infection. The far upstream major IE promoter drives transcription through the viral genes designated IE1 (UL122) and IE2 (UL121) (11, 38). Some of the IE transcription may continue through UL120 and UL119 (IE2b). A polyadenylation signal is located within the intron downstream of UL119 (Fig. 10). Therefore, the TATA box upstream of UL119 is also nested within a transcription unit. IE transcription occurs across the UL119 promoter and is also initiated downstream of the UL119 promoter. It is possible that some of the RNA polymerase II complex originally

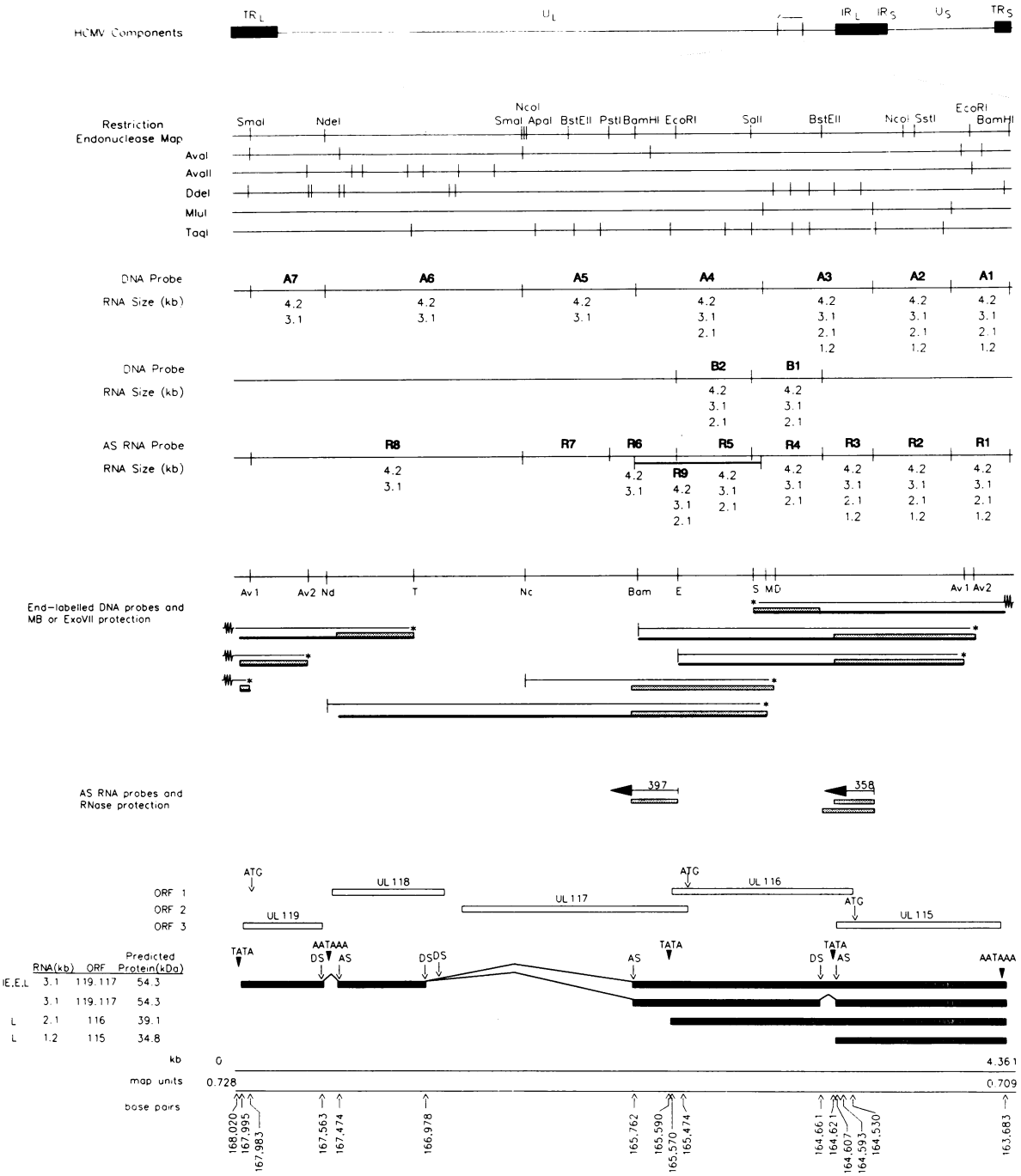

FIG. 10. The UL119–115 transcription unit. The prototype arrangement of the viral genome and the direction of transcription in UL119–115 are indicated. The genome consists of a long (L) and a short (S) component. Each unique sequence is bounded by terminal repeats (TR) and inverted internal repeats (IR). The restriction endonuclease sites and abbreviations thereof are indicated at the top and in the middle, respectively (Av1, *Ava*I; Av2, *Ava*II; Nd, *Nde*I; T, *Taq*I; Nc, *Nco*I; Bam, *Bam*HI; E, *Eco*RI; S, *Sal*I; M, *Mlu*I; D, *Dde*I). DNA probes, antisense (AS) RNA probes, and end-labeled DNA probes are indicated. The size classes of mRNAs detected by Northern blot analysis with the various DNA or antisense RNA probes are summarized. End-labeled DNA probes are shown as thin lines with an asterisk at the labeled end. Protection of end-labeled DNA probes from mung bean nuclease (MB) digestion is indicated by lightly stippled bars. Protection from exonuclease VII (ExoVII) digestion is indicated by thick lines. Protection of antisense RNA probes from RNase digestion is indicated by lightly stippled bars in the "AS RNA probes and RNase protection" section. Open boxes designate the UL119–115 ORFs. Viral RNAs detected at IE, early (E), or late (L) times after infection are indicated (bottom left). Heavily stippled boxes represent exons, and lines connecting them represent introns. Whether one or both donor splice sites in UL118 are used was not determined. TATA boxes, donor splice (DS) sites, acceptor splice (AS) sites, and polyadenylation signals are shown at the bottom. The size of the transcription unit in kilobases and map units is shown. The base pair locations, according to the AD169 strain DNA sequence, of the various transcription, RNA processing, and translation signals are indicated and are discussed in the text.

recruited by the enhancer-containing major IE promoter also recognizes the UL119 promoter. However, the UL115 promoter located just downstream also has a typical TATA box motif and is not recognized by the RNA polymerase II complex, even though this entire region is active for transcription under IE conditions or at early times after infection. The UL116 and UL115 regions are transcribed at IE and early times, and consequently, they are part of a large mRNA. Because the 4.2-kb RNA is the approximate size of the transcription unit, it was interpreted to be a precursor RNA molecule. The 3.1-kb RNA was interpreted to be a processed mRNA. The 3.1-kb mRNA from the UL119–115 transcription unit has the potential for expression of UL119–117, UL116, and UL115. In this mRNA, the translation initiation sites for UL116 and UL115 are located far downstream of the 3.1-kb mRNA cap site. The promoters upstream of UL116 and UL115 are not efficiently recognized until after viral DNA replication. Recognition of these promoters at late times after infection designates new 5′ initiation sites. The selection of a new cap site downstream of a true late promoter may be a mechanism to ensure efficient expression of these ORFs, which are normally positioned far downstream of the cap site of the IE-early mRNA.

The viral gene products predicted to be encoded by the 3.1-, 2.1-, and 1.2-kb mRNAs have the features of membrane-bound glycoproteins. Toward the amino termini of the predicted proteins are stretches of hydrophobic residues which may function in a secretory pathway. Toward the carboxyl termini are stretches of hydrophobic amino acid residues sufficient to span the cell membrane and anchor the viral glycoproteins. UL119–117 and UL116 have a high number of potential N-glycosylation sites (6). UL115 has only one potential N-glycosylation site, but UL115 has the properties of a membrane-bound glycoprotein and the potential for extensive O glycosylation.

HCMV appears unusual among the characterized herpesviruses because it has several enhancer elements upstream of different potential glycoprotein genes and these glycoprotein genes are transcribed at IE and early times after infection (14, 46). We propose that the UL119–117 gene may also be an IE-early viral glycoprotein gene. In contrast, maximal expression of UL116 and UL115 may occur at late times after infection. As described by Chee et al. (6), HCMV has the potential to code for a total of 54 reading frames that have the characteristics of glycoprotein genes. Expression of these viral glycoproteins is temporally regulated, with some being expressed early and others at late times (39). Activation of true late promoters that designate new downstream cap sites for mRNAs may play a key role in the temporal regulation of HCMV gene expression. The roles of the temporally regulated expression of viral glycoproteins and of the complicated array of viral glycoproteins in the biology and pathogenicity of HCMV infections pose difficult research questions about a species-specific virus.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant HD-19937 from the National Institute of Child Health and Human Development and in part by Public Health Service grant AI-13562 from the National Institute of Allergy and Infectious Diseases.

REFERENCES

1. **Alford, C. A., and W. J. Britt.** 1990. Cytomegalovirus, p. 1981–2010. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology. Raven Press, Ltd., New York.
2. **Bale, J. F., M. E. O'Neil, M. N. Hart, J. D. Harris, and M. F. Stinski.** 1989. Human cytomegalovirus nucleic acids in tissues from congenitally infected infants. Pediatr. Neurol. **5**:216–220.
3. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12**:721–732.
4. **Chang, C.-P., C. L. Malone, and M. F. Stinski.** 1989. A human cytomegalovirus early gene has three inducible promoters that are regulated differentially at various times after infection. J. Virol. **63**:281–290.
5. **Chang, C.-P., D. H. Vesole, J. A. Nelson, M. B. A. Oldstone, and M. F. Stinski.** 1989. Identification and expression of a human cytomegalovirus early glycoprotein. J. Virol. **63**:3330–3337.
6. **Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchison III, T. Kouzarides, J. A. Martignetti, E. Preddi, S. C. Satchwell, P. Tomlinson, K. M. Weston, and B. G. Barrell.** 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169, p. 125–170. *In* J. K. McDougall (ed.), Cytomegaloviruses. Springer-Verlag, New York.
7. **DeMarchi, J. M.** 1981. Human cytomegalovirus DNA: restriction enzyme cleavage maps and map location for immediate early, early, and late RNAs. Virology **124**:390–402.
8. **Einhorn, L., and A. Ost.** 1984. CMV infection of human blood cells. J. Infect. Dis. **149**:207–214.
9. **Gretch, D. R., and M. F. Stinski.** 1990. Transcription of the human cytomegalovirus glycoprotein gene family in the short unique component of the viral genome. Virology **174**:522–532.
10. **Hardwicke, M. A., P. J. Vaughan, R. E. Sekulovich, R. O'Conner, and R. M. Sandri-Goldin.** 1989. The regions important for the activator and repressor functions of herpes simplex virus type 1 α protein ICP27 map to the C-terminal half of the molecule. J. Virol. **63**:4590–4602.
11. **Hermiston, T. W., C. L. Malone, P. R. 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.
12. **Ho, M.** 1982. Cytomegalovirus biology and infection, p. 1–275. Plenum Publishing Corp., New York.
13. **Keck, J. G., C. J. Baldick, Jr., and B. Moss.** 1990. Role of DNA replication in vaccinia virus gene expression: a naked template is required for transcription of three late trans-activator genes. Cell **61**:801–809.
14. **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.
15. **Krieg, P. A., and D. A. Melton.** 1987. In vitro RNA synthesis with SP6 RNA polymerase. Methods Enzymol. **155**:397–414.
16. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
17. **Lahijani, R. S., E. W. Otteson, J. D. Adlish, and S. C. St. Jeor.** 1991. Characterization of a human cytomegalovirus 1.6-kilobase late mRNA and identification of its putative protein product. J. Virol. **65**:373–381.
18. **Leach, F. S., and E. S. Mocarski.** 1989. Regulation of cytomegalovirus late-gene expression: differential use of three start sites in the transcriptional activation of ICP36 gene expression. J. Virol. **63**:1783–1791.
19. **Mach, M., T. Stamminger, and G. Jahn.** 1989. Human cytomegalovirus: recent aspects from molecular biology. J. Gen. Virol. **70**:3117–3146.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
22. **McCarthy, A. M., L. McMahan, and P. A. Schaffer.** 1989. Herpes simplex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA deficient. J. Virol. **63**:18–27.
23. **McDonough, S. H., and D. H. Spector.** 1983. Transcription in

human fibroblasts permissively infected by human cytomegalovirus strain AD169. Virology **125**:31–46.
24. **McMahan, L., and P. A. Schaffer.** 1990. The repressing and enhancing functions of the herpes simplex virus regulatory protein ICP27 map to C-terminal regions and are required to modulate viral gene expression very early in infection. J. Virol. **64**:3471–3485.
25. **Nelson, J. A., J. W. Gnann, Jr., and P. Ghazal.** 1990. Regulation and tissue-specific expression of human cytomegalovirus, p. 75–100. *In* J. K. McDougall (ed.), Cytomegaloviruses. Springer-Verlag, New York.
26. **Rice, G. P. A., R. D. Schrier, and M. B. A. Oldstone.** 1984. Cytomegalovirus infects human lymphocytes and monocytes: virus expression is restricted to immediate-early gene products. Proc. Natl. Acad. Sci. USA **81**:6134–6138.
27. **Rice, S. A., and D. M. Knipe.** 1990. Genetic evidence for two distinct transactivation functions of the herpes simplex virus α protein ICP27. J. Virol. **64**:1704–1715.
28. **Rice, S. A., L. Su, and D. M. Knipe.** 1989. Herpes simplex virus alpha protein ICP27 possesses separable positive and negative regulatory activities. J. Virol. **63**:3399–3407.
29. **Rudolph, S.-A., T. Stamminger, and G. Jahn.** 1990. Transcriptional analysis of the eight-kilobase mRNA encoding the major capsid protein of human cytomegalovirus. J. Virol. **64**:5167–5172.
30. **Sacks, W. R., C. C. Greene, D. P. Aschmann, and P. A. Schaffer.** 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. **55**:796–805.
31. **Schrier, R. D., J. A. Nelson, and M. B. A. Oldstone.** 1985. Detection of human cytomegalovirus in peripheral blood lymphocytes in a natural infection. Science **230**:1048–1051.
32. **Smale, S. T., and D. Baltimore.** 1989. The initiator as a transcription control element. Cell **57**:103–113.
33. **Spector, D. H., K. M. Klucher, D. K. Robert, and D. A. Wright.** 1990. Human cytomegalovirus early gene expression, p. 21–46. *In* J. K. McDougall (ed.), Cytomegaloviruses. Springer-Verlag, New York.
34. **Stamminger, T., and B. Fleckenstein.** 1990. Immediate-early transcription regulation of human cytomegalovirus, p. 3–20. *In* J. K. McDougall (ed.), Cytomegaloviruses. Springer-Verlag, New York.
35. **Stenberg, R. M., A. S. Depto, J. Fortney, and J. A. Nelson.** 1989. Regulated expression of early and late RNAs and proteins from the human cytomegalovirus immediate-early gene region. J. Virol. **63**:2699–2708.
36. **Stenberg, R. M., and M. F. Stinski.** 1985. Autoregulation of the human cytomegalovirus major immediate-early gene. J. Virol. **56**:676–682.
37. **Stenberg, R. M., D. R. Thomsen, and M. F. Stinski.** 1984. Structural analysis of the major immediate early gene of human cytomegalovirus. J. Virol. **49**:190–199.
38. **Stenberg, R. M., P. R. Witte, and M. F. Stinski.** 1985. Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate-early region 1. J. Virol. **56**:665–675.
39. **Stinski, M. F.** 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. J. Virol. **23**:751–767.
40. **Stinski, M. F.** 1990. Cytomegaloviruses and their replication, p. 1959–1980. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology. Raven Press, Ltd., New York.
41. **Stinski, M. F., C. L. Malone, T. W. Hermiston, and B. Liu.** 1991. Regulation of human cytomegalovirus transcription, p. 246–260. *In* E. K. Wagner (ed.), Herpesvirus transcription and its regulation. CRC Press, Inc., Boca Raton, Fla.
42. **Stinski, M. F., D. R. Thomsen, R. M. Stenberg, and L. C. Goldstein.** 1983. Organization and expression of the immediate early genes of human cytomegalovirus. J. Virol. **46**:1–14.
43. **Thomsen, D. R., and M. F. Stinski.** 1981. Cloning of the human cytomegalovirus genome as endonuclease XbaI fragments. Gene **16**:207–216.
44. **Vales, L. D., and J. E. Darnell.** 1989. Promoter occlusion prevents transcription of adenovirus polypeptide IX until after DNA replication. Genes Dev. **3**:49–59.
45. **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.
46. **Weston, K.** 1988. An enhancer element in the short unique region of human cytomegalovirus regulates the production of a group of abundant immediate early transcripts. Virology **162**: 406–416.