JOURNAL OF VIROLOGY, July 1989, p. 3117–3127
0022-538X/89/073117-11$02.00/0

Vol. 63, No. 7

Posttranscriptional Regulation of a Class of Human Cytomegalovirus Phosphoproteins Encoded by an Early Transcription Unit

DOWAIN A. WRIGHT AND DEBORAH H. SPECTOR*

Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093

Received 17 January 1989/Accepted 2 April 1989

We have further characterized and determined the origins of a family of nuclear phosphoproteins of 84, 50, 43, and 34 kilodaltons (kDa) encoded by a class of early transcripts arising from the adjacent *Eco*RI fragments R and d (map units 0.682 to 0.713) of the strain AD169 human cytomegalovirus genome. These RNAs have a complex spliced structure with common 5′ and internal exons and alternative 3′ exons with coterminal 3′ ends. At early times, two fully processed species of 2.1 and 2.2 kilobases (kb) predominated. As the infection progressed to late times, there was a decrease in splicing of the RNA, generating larger transcripts 2.5 to 2.65 kb in size, which corresponded to the species which had spliced out only the first intron, as well as the completely unspliced transcript. We previously reported that the 34-kDa protein could be derived from a transcript which had failed to splice out the first intron (D. A. Wright, S. I. Staprans, and D. H. Spector, J. Virol. 62:331–340, 1988), but the origin of the other proteins was unclear. cDNA cloning has shown that the 2.1-, 2.2-, and 2.5-kb RNAs encode the 50-, 43-, and 84-kDa proteins, respectively. The shift in the splicing pattern of these RNAs with time revealed a posttranscriptional control mechanism which results in the differential accumulation of individual proteins within this family of nuclear phosphoproteins. Expression of the 84-, 43-, and 34-kDa proteins correlated well with the steady-state concentrations of their respective mRNAs. The 50-kDa protein, however, was not expressed in abundance until late times, despite the presence of the 2.1-kb mRNA in the cytoplasm at early times, suggesting a secondary level of posttranscriptional regulation for this protein. Full expression of the RNAs and proteins was dependent on continuing viral DNA synthesis. Accumulation of the 50-kDa protein was found to be particularly sensitive to the state of viral DNA replication and could not be detected after inhibition of replication. Further analysis of these proteins revealed that each one had a unique pattern of serine phosphorylation. Although there was one common site of phosphorylation, most likely located within the amino-terminal shared region, even this site showed quantitative differences in the level of phosphorylation for each of the proteins. Analysis of the Towne strain and two recent independent clinical isolates of human cytomegalovirus has shown that this family of proteins is highly conserved among human cytomegaloviruses.

Human cytomegalovirus (HCMV), a member of the herpesvirus group, infects a majority of the adult population, is recognized as the leading viral cause of birth defects, and poses a serious health threat to immunocompromised individuals (18, 25). HCMV can establish both persistent and latent infections and may have oncogenic potential (28, 33).

After infection of permissive cells, HCMV displays a complex sequential pattern of gene expression (6, 24, 37, 38, 40, 41). Production of herpesvirus-specific gene products has been categorized into three broad temporal classes: immediate-early (IE) (alpha), early (beta), and late (gamma) (19). IE proteins are synthesized immediately upon infection, and the induction of their transcription units does not require de novo protein synthesis. Early-protein synthesis directly follows and is dependent on prior IE-protein expression. Finally, the late proteins are expressed after the onset of viral DNA replication and function predominantly in the packaging of viral DNA and the formation of mature virus particles. For HCMV, this temporal pattern of gene expression appears to depend on both transcriptional and posttranscriptional control mechanisms (7, 9–11, 14, 17, 23, 30, 34, 36, 40).

The HCMV genome is a large double-stranded DNA molecule, 240 kilobase pairs (kbp) in length, with a theoretical coding capacity of more than 100 proteins (for a review, see reference 31). Identification of specific protein products and determination of their biochemical characteristics will provide a foundation for understanding the multiple virus-host interactions involved in the replication cycle of HCMV. Herpesvirus early proteins are of interest because they have been implicated in the regulation of viral gene expression and DNA replication (3, 13, 20, 26), and a number of antiherpesvirus chemotherapeutic agents act by interfering with the function of these proteins (5, 8, 16, 21, 29).

In our previous studies (35), we characterized a class of early transcripts (originally designated the 2.2-kilobase [kb] class) arising from the adjacent *Eco*RI fragments R and d (map units 0.682 to 0.713), located within the long unique segment of the HCMV strain AD169 genome. The fully processed RNAs consisted of invariable 5′ and internal exons and alternative 3′ exons with coterminal 3′ ends. This class of transcripts was found to encode four nuclear phosphoproteins of 84, 50, 43, and 34 kilodaltons (kDa) which share a common amino terminus (44). The 43-kDa protein was the most abundant of the four proteins, and its level of expression remained relatively constant throughout the infection, while expression of the other proteins increased as the infection progressed to late times. There was no evidence for a precursor-product relationship between any of the proteins, and although in vitro synthesis studies suggested that the 34-kDa protein could be generated by mRNAs

* Corresponding author.

which had failed to splice out the first intron, the derivation of the other proteins was unclear.

In this study, we have utilized cDNA cloning and transient expression assays to assign the 84-, 50-, 43-, and 34-kDa proteins to their corresponding mRNAs. We have further analyzed the patterns of expression of the mRNAs and proteins during the course of the HCMV infection both in the presence and absence of viral DNA replication. Our data indicate that this family of gene products is highly regulated by both transcriptional and posttranscriptional mechanisms. We also show that the phosphorylation patterns of the individual proteins differ both quantitatively and qualitatively, even for shared domains. Finally, analysis of different HCMV isolates reveals that this family of proteins is highly conserved among HCMVs.

MATERIALS AND METHODS

Virus and cells. HCMV strain AD169 and human fibroblasts were obtained and propagated as previously described (39). In some experiments, cells were infected and maintained in the presence of 300 μg of phosphonoacetic acid (PAA) per ml or the PAA was added at various times postinfection (p.i.). HCMV strain Towne was obtained from Stephen Spector and propagated as described for AD169. The clinical isolates Co and Ro were obtained from Stephen Spector and propagated in human fibroblasts by the addition of uninfected fibroblasts during passage of the cell cultures infected with these cell-associated HCMV isolates.

Generation of clones. All enzymes were obtained from Bethesda Research Laboratories, Inc., Gaithersburg, Md., and were used according to the recommendations of the manufacturer, unless otherwise noted. pIERdΔP was generated as follows. pGemRd-1 (44) was linearized with *Xba*I and partially restricted with *Bam*HI. The 3-kbp *Xba*I-*Bam*HI partial fragments were separated by agarose gel electrophoresis and isolated by GeneClean (Bio 101, Inc., San Diego, Calif.). pGem-1 (Promega Biotec, Madison, Wis.) was linearized with *Hin*dIII, treated with Klenow polymerase to generate blunt ends, followed by the addition of *Bgl*II linkers, and finally cleaved with both *Bgl*II and *Xba*I. The resulting vector was purified from excess linker by agarose gel electrophoresis, isolated by GeneClean, and ligated to the *Xba*I-*Bam*HI fragment of pGemRd-1 described above. The resulting clone which had eliminated the downstream *Pst*I site in the original pGemRd-1 polylinker was called pGemRdΔP. This construct was then restricted with *Xba*I, extracted with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol), and precipitated with ethanol. The linearized plasmid was dephosphorylated by treatment with calf intestinal alkaline phosphatase (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) according to the recommendations of the supplier. The phosphatase-treated DNA and the 600-bp HCMV major IE promoter from *Xba*I-cleaved pIEP2 (pIEP2 is a pUC9-based vector with the 600-bp *Hin*cII-to-*Sau*3AI fragment of the HCMV strain AD169 major IE promoter [1] cloned into a polylinker with both 5′ and 3′ *Xba*I sites) were electrophoresed on agarose gels, isolated by GeneClean, and ligated together. The resulting clone was then called pIERdΔP. pΔPSV, pΔ6SV, pΔ45SV, and pΔ60SV were generated as follows. pIERdΔP, pRdΔ6, pRdΔ45, and pRdΔ60 (the genomic and cDNA clones whose production is described below) were linearized with *Sma*I (upstream of the HCMV IE promoter). The DNA was extracted with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol) and precipitated with ethanol. The linearized plasmids were dephosphorylated by treatment with calf intestinal alkaline phosphatase (Boehringer Mannheim) according to the recommendations of the supplier. The phosphatase-treated DNAs were electrophoresed on agarose gels, isolated by GeneClean, and ligated to the 325-bp simian virus 40 (SV40) origin of replication. The SV40 origin was generated from pSVtkneo3 (kindly provided by Jay Nelson) by sequential treatment with *Hin*dIII, Klenow, and *Pvu*II and then isolated by polyacrylamide gel electrophoresis.

cDNA cloning. Oligonucleotides (25-mers) corresponding to sequences from the 5′ and 3′ exons of the 2.2-kb RNAs were synthesized at the Peptide Oligonucleotide Synthesis Facility at the University of California, San Diego. The oligonucleotides were purified by polyacrylamide gel electrophoresis and utilized to prime both first- and second-strand cDNA reactions. Synthesis of first-strand cDNA was carried out in a reaction volume of 40 μl containing 50 mM Tris hydrochloride (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM TTP, 100 μg of oligonucleotide (L622) per ml, 60 U of RNAsin (Promega), 200 U of Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories), and 4 μg of AD169-infected (72 h p.i.) cell poly(A)$^+$ mRNA at 37°C for 2 h. The reaction was extracted with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol) adjusted to 2 M ammonium acetate and precipitated twice with ethanol. The RNA-cDNA hybrids were suspended in 20 μl of H_2O, heated to 100°C for 3 min, and quickly placed on ice. Second-strand synthesis was performed in a reaction volume of 100 μl containing 100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 6.9), 10 mM $MgCl_2$, 2 mM dithiothreitol, 70 mM KCl, 0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dGTP, 0.5 mM TTP, 100 μg of oligonucleotide (L621) per ml, 100 U of Klenow polymerase (Bethesda Research Laboratories), and denatured RNA-DNA from the first-strand synthesis at 15°C for 4 h. The reaction was extracted with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol) and precipitated twice following the addition of 2 M ammonium acetate and ethanol. Second-strand synthesis was completed in a 50-μl reaction volume containing 100 mM Tris hydrochloride (pH 8.3), 140 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM TTP, Klenow-treated cDNA from above, and 20 U of avian myeloblastosis virus reverse transcriptase (Bethesda Research Laboratories) at 42°C for 2 h. The reaction was extracted with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol) and precipitated twice following the addition of 2 M ammonium acetate and ethanol. The cDNAs were then cut with *Cla*I and *Pst*I restriction endonucleases (Bethesda Research Laboratories) in the buffer supplied by the manufacturer. The reaction was extracted with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol), passed over a Sephadex G-50 spin column, and precipitated twice with 2 M ammonium acetate and ethanol as described above. The final cDNAs were ligated into *Cla*I-*Pst*I-cleaved pIERdΔP (as described above) which had been propagated in the Dam methylase-negative strain LEΔ10 (kindly provided by Douglas Smith). The resulting clones representing each class of mRNA were called pRdΔ6, pRdΔ45, and pRdΔ60.

DNA sequencing. Sequence analysis of the splice donor-acceptor regions of the genomic (pΔPSV) as well as cDNA clones (pΔ6SV, pΔ45SV, and pΔ60SV) was performed using the GemSeq K/RT sequencing system (Promega) essentially as recommended by the supplier, except that the sequencing primers utilized were 20-mer oligonucleotides which corre-

sponded to the sequences extending from 94 to 74 bp upstream of the putative splice donor of exon 1 and from 67 to 47 bp upstream of the putative splice donor of exon 2 (sequence analysis not shown). The primers were synthesized and purified as described above for the cDNA cloning and were used at the same concentration as was recommended for the T7 or SP6 primers supplied with the sequencing kit.

Transient expression. Subconfluent monolayers of COS-7 cells (5×10^5 cells per 60-mm dish) were washed twice with phosphate-buffered saline followed by the addition of 4 ml of Dulbecco modified Eagle medium (Irvine Scientific, Irvine, Calif.) without serum containing 400 μg of DEAE-dextran per ml and 5 μg of DNA (pΔPSV, pΔ6SV, pΔ45SV, or pΔ60SV) and incubated for 2.5 h at 37°C. The cells were then washed twice with phosphate-buffered saline and once with Dulbecco modified Eagle medium containing 10% calf serum, followed by the addition of 5 ml of the medium described above and incubation at 37°C. At 48 h posttransfection, the monolayers were trypsinized, pelleted in a clinical centrifuge, suspended in 200 μl of sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer (2% SDS, 5 mM Na_2HPO_4 [pH 7.0], 10% glycerol, 5% 2-mercaptoethanol, 0.1 M dithiothreitol, 0.01% bromophenol blue), heated to 100°C for 5 min, and analyzed along with AD169-infected cell proteins by immunoblotting as previously described (44).

Preparation of infected-cell RNA. RNA was prepared from AD169-infected cells at 2, 8, 26, and 72 h p.i. or from infected cells incubated in the presence of 300 μg of PAA per ml from 26 to 72 or from 0 to 72 h p.i. Cytoplasmic RNA was isolated after lysis in an isotonic buffer containing 0.5% Nonidet P-40 and vanadyl-ribonucleoside complexes and was enriched for polyadenylated species by oligo(dT)-cellulose chromatography as previously described (24).

Labeling of RNA with ^{32}P. Radioactive single-stranded probes were transcribed from the *Hin*dIII-linearized clone pG0.9XBR. pG0.9XBR contains the 900-bp *Xho*I-to-*Bam*HI fragment of *Eco*RI fragment R cloned into an *Xho*I-to-*Bam*HI-cleaved pGem-1 vector (Promega). Antisense RNA probe was generated by using 1 μg of the linear DNA template in a 20-μl reaction volume containing the following: 40 mM Tris hydrochloride (pH 7.5); 6 mM $MgCl_2$; 10 mM dithiothreitol; 2 mM spermidine; 100 μg of bovine serum albumin per ml; 50 U of RNAsin (Promega); 500 μM each ATP, CTP, and GTP; and 10 μM UTP plus 100 μCi of [α-^{32}P]UTP. After the addition of 4 U of T7 polymerase (Promega), the reaction mixture was incubated at 37°C for 10 min. This was followed by the addition of RNase-free DNase to 50 μg per ml, and the reaction mixture was incubated at 37°C for 10 min. After extraction with phenol-chloroform-isoamyl alcohol (50:48:2, vol/vol/vol), the labeled RNA was purified from unincorporated nucleotides by Sephadex G-50 chromatography and used directly as a probe for the specific RNAs.

Partial proteolysis. Partial proteolysis with *Staphylococcus aureus* V-8 protease was performed essentially as described by Cleveland et al. (2). Briefly, infected-cell lysates, labeled with [^{35}S]methionine or $^{32}P_i$, were immunoprecipitated and subjected to SDS-PAGE as previously described (44). The gels were dried and autoradiographed at −70°C with an intensifying screen. Protein bands were excised and rehydrated in a small volume of overlay buffer (125 mM Tris hydrochloride [pH 6.8], 1 mM EDTA, 0.1% SDS, 1 mM 2-mercaptoethanol, 30% glycerol, 0.005% bromophenol blue). Gel pieces were placed in the wells of an SDS-PAGE gel; 10 μl of overlay buffer was added to each well; either 0, 50, or 200 ng of V-8 protease (Pierce Chemical Co., Rockford, Ill.) in 10 μl of enzyme buffer (125 mM Tris hydrochloride [pH 6.8], 1 mM EDTA, 0.1% SDS, 1 mM 2-mercaptoethanol, 10% glycerol, 0.01% phenol red) was added; and the samples were electrophoresed until the dye front was 3 mm from the top of the resolving gel. The power was turned off for 30 min, and then electrophoresis was resumed. Gels containing ^{35}S-labeled samples were impregnated with PPO (2,5-diphenyloxazole; Sigma Chemical Co., St. Louis, Mo.), dried, and fluorographed at −70°C. Gels containing ^{32}P-labeled samples were dried and autoradiographed at −70°C with an intensifying screen.

^{32}P-tryptic peptide mapping. Tryptic peptide analysis of $^{32}P_i$-labeled proteins was performed essentially as described previously for [^{35}S]methionine-labeled proteins (44). Electrophoresis was performed at 1.0 kV for 27 min with *n*-butanol-pyridine-acetic acid-water (2:1:1:36, vol/vol) (pH 4.72), followed by chromatography in the second dimension in buffer containing *n*-butanol, pyridine, acetic acid, and water (75:50:15:60, vol/vol). The plates were autoradiographed at −70°C with an intensifying screen.

Phosphoamino acid analysis. Phosphoamino acid content of the proteins was determined as described previously (4, 22). Briefly, labeled proteins were immunoprecipitated and resolved by SDS-PAGE, extracted from the gel, suspended in 0.4 ml of constant-boiling hydrochloric acid (Pierce), and hydrolyzed at 110°C for 1 h. After lyophilization, the samples were suspended in water and spotted on thin-layer cellulose chromatography plates (EM Science, Cherry Hill, N.J.) along with 0.5 μg each of phosphoserine, phosphothreonine, and phosphotyrosine (Sigma). Electrophoresis in the first dimension was in pH 1.9 buffer (2.2% formic acid, 7.8% acetic acid) at 1.5 kV for 20 min, followed by rotation of 90° and electrophoresis in pH 3.5 buffer (0.5% pyridine, 5% acetic acid, 0.5 mM EDTA) at 1.3 kV for 16 min. The internal markers were visualized with ninhydrin, and the plate was autoradiographed at −70°C with an intensifying screen.

ADP-ribosylation. Incorporation of [^{32}P]NAD by uninfected or AD169-infected (48 h p.i.) cell nuclei was performed as described by Preston and Notarianni (27). After termination of the reaction, the labeled nuclei were split into two fractions and pelleted at 2,000 × *g* for 4 min. One fraction was suspended in 80 μl of SDS-PAGE sample buffer, heated to 100°C for 5 min, and stored at −20°C. The other fraction was suspended in 100 μl of RIPA buffer (10 mM sodium phosphate [pH 7.2], 150 mM NaCl, 0.5 trypsin-inhibiting U of aprotinin [Sigma] per ml, 1% sodium deoxycholate, 1% Nonidet P-40, 0.1% SDS) and incubated at 4°C for 30 min. The labeled nuclear lysate was cleared by centrifugation at 4°C for 15 min in an Eppendorf centrifuge. The proteins were immunoprecipitated with antipeptide antiserum as described previously (44) and analyzed by SDS-PAGE along with the total nuclear lysate.

RESULTS

Generation of cDNAs and transient expression. We previously identified four nuclear phosphoproteins of 84, 50, 43, and 34 kDa which share a common amino terminus and are encoded by an HCMV early transcription unit which generates a complex set of spliced transcripts (35, 44). To precisely determine the origin of these proteins, we prepared cDNAs corresponding to each class of transcript (see Materials and Methods and Fig. 1). Cytoplasmic poly(A)$^+$ RNA from AD169-infected cells at 72 h p.i. was used as the

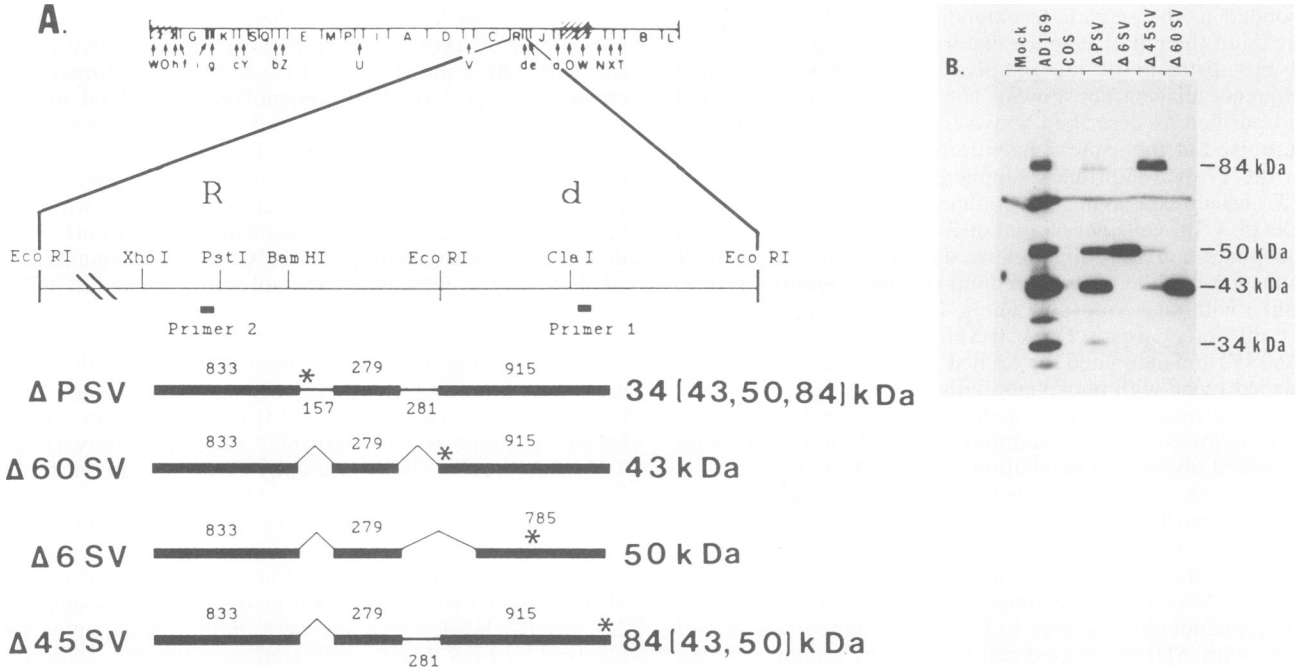

FIG. 1. cDNA cloning and transient expression. (A) The *Eco*RI restriction profile of HCMV strain AD169 (32) is shown, along with an expanded map of the R-d region encoding the family of early RNAs. The positions of the component exons and introns of the RNAs are summarized below, with sizes given in nucleotides. ΔPSV is a plasmid construct containing the SV40 origin of replication and the HCMV major IE promoter (1) followed by the genomic sequences encoding the HCMV family of early proteins beginning 25 bp 5′ of the start site of translation and extending to the *Bam*HI site in fragment d (44). cDNAs were generated from cytoplasmic poly(A)$^+$ RNA, obtained from AD169-infected cells at 72 h p.i. Oligonucleotide primers 1 and 2 were utilized for first- and second-strand synthesis, respectively, and representative cDNA clones for each class of mRNA were isolated (Δ60SV, Δ6SV, and Δ45SV). In-frame stop codons for each cDNA (*) are shown. (B) Immunoblot analysis of transiently expressed cDNAs. The genomic (ΔPSV) and cDNA expression vectors (Δ6SV, Δ45SV, and Δ60SV) were transfected into COS-7 cells, and total cellular proteins were isolated at 48 h posttransfection, along with untransfected COS-7, mock-infected fibroblasts, and AD169-infected (48 h p.i.) cell proteins, separated by SDS-PAGE, transferred to nitrocellulose, probed with immune antiserum, and labeled with ^{125}I-protein A.

template for the reaction. First-strand synthesis was primed by a 25-base synthetic oligonucleotide (L622), designated primer 1, corresponding to a sequence located near the 3′ end of the RNA coding region (Fig. 1A). Second-strand synthesis was primed by another 25-base oligonucleotide (L621), designated primer 2, corresponding to a sequence near the 5′ end of the first exon. The nascent cDNAs were restricted with *Pst*I and *Cla*I and cloned directly into a *Pst*I-*Cla*I-cleaved pIERdΔP expression vector, as described in Materials and Methods. This expression vector contains the major HCMV IE promoter but uses both the natural ATG translation initiation site and the 3′-terminal polyadenylation site for these transcripts. The reaction generated 336 clones, of which 82 were positive for insertion of the viral-specific sequences. Of the positive cDNAs, 62 clones corresponded to full-length unspliced transcripts, 8 clones had spliced out the first intron and the shorter of the two alternative second introns and are represented by pRdΔ60, 7 clones had spliced out the first intron and the longer second intron and are represented by pRdΔ6, and 3 clones had spliced out only the first intron and are represented by pRdΔ45. Because we could detect no difference between cDNAs for the full-length unspliced transcript and the genomic DNA, the clone corresponding to the genomic DNA was used for all subsequent experiments. The SV40 origin of replication was cloned into the genomic clone (pIERdΔP) as well as the representative cDNAs (pRdΔ6, pRdΔ45, and pRdΔ60), and the resulting clones were called pΔPSV, pΔ6SV, pΔ45SV, and pΔ60SV, respectively.

The constructs pΔPSV, pΔ6SV, pΔ45SV, and pΔ60SV were transfected into COS-7 cells by the DEAE-dextran method, and at 48 h posttransfection transient proteins were analyzed by immunoblot (Fig. 1B). The genomic clone, pΔPSV, generated all four proteins in proportions similar to those found in AD169-infected cells at late times. pΔ6SV, which had spliced out the first intron and had spliced the second exon to the 785-nucleotide 3′ exon, encoded only the 50-kDa protein. pΔ45SV, which had spliced out only the first intron, encoded predominantly the 84-kDa protein along with traces of the 50- and 43-kDa proteins, most likely from secondary splicing events. pΔ60SV, which had spliced out the first intron and had spliced the second exon to the 915-nucleotide 3′ exon, generated only the 43-kDa protein. Sequence analysis of the splice donor-acceptor sites for the genomic clone (pΔPSV) as well as the cDNAs (pΔ6SV, pΔ45SV, and pΔ60SV) was performed by dideoxy sequencing, using synthetic oligonucleotide primers 5′ to each of the putative splice donors (data not shown). These data allowed for more precise assignment of the sizes of the common exons as well as for the first intron and the alternative 3′ introns (Fig. 1A). The sizes were comparable to those found by Staprans and Spector (35), using exon-mapping techniques. The sizes assigned to the alternative 3′ exons remain based on these previous studies.

From the data described above, we conclude that the 43- and 50-kDa proteins are encoded by the fully processed 2.2- and 2.1-kb transcripts, respectively. The 84-kDa protein could then be encoded by a larger transcript which had

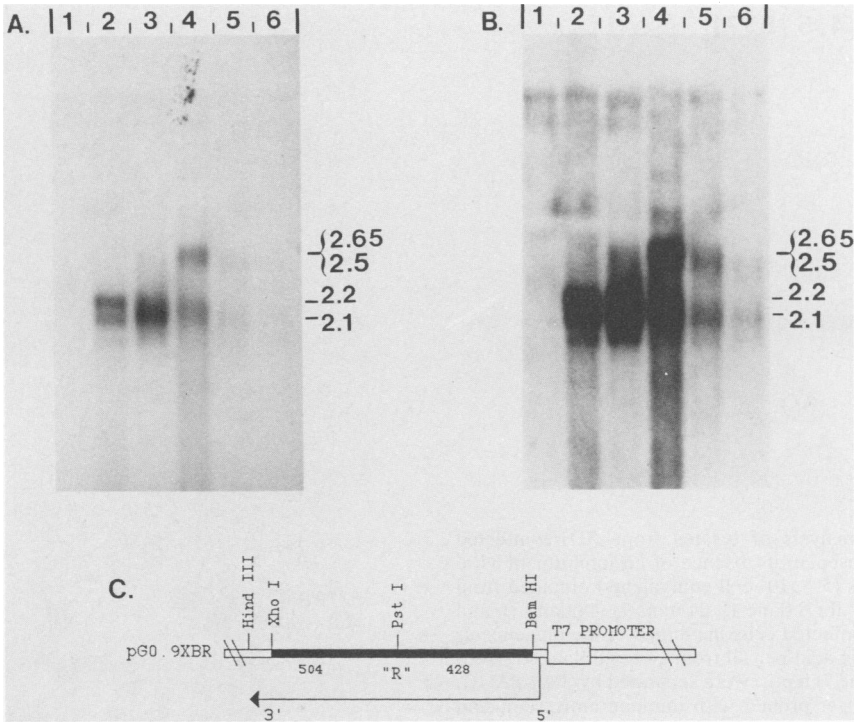

FIG. 2. Northern (RNA) analysis of the HCMV class of early transcripts synthesized in the presence or absence of an inhibitor of viral DNA replication. (A and B) Cytoplasmic polyadenylated RNA (3.6 μg per lane) extracted from HCMV-infected cells at 2 (lane 1), 8 (lane 2), 26 (lane 3), and 72 (lane 4) h p.i., along with infected cells incubated in the presence of 300 μg of phosphonoacetic acid per ml from 26 to 72 (lane 5) and 0 to 72 (lane 6) h p.i., was electrophoresed on a denaturing formaldehyde gel, transferred to NYTRAN (Schleicher & Schuell), and hybridized to ^{32}P-labeled single-stranded probe derived from the first exon of the RNAs (C). Panels A and B are 12- and 60-h exposures of the same blot. Fragment sizes are given (in kilobases) on the right of the gels.

spliced out only the first intron, leaving the 34-kDa protein to be encoded by an unspliced transcript, as we had suggested previously (44). It is clear that these proteins are generated by differential splicing of their respective RNA transcripts. Thus, production of these gene products is regulated not only at the level of transcription, as shown previously (34), but also posttranscriptionally by alternative RNA processing.

Pattern of expression of the family of early RNAs. Previous analysis of the kinetics of expression of the members of this family of proteins showed temporal differences in their pattern of expression (44). The 43-kDa protein was the most abundant of the four proteins, and its level of expression remained relatively constant throughout the infection. Expression of the other proteins increased as the infection progressed. In an earlier study (35), it was also noted that the complexity of transcription from the general region of the genome encoding these proteins increased greatly as the infection progressed, but these transcripts were not characterized. To more accurately determine the temporal pattern of accumulation of the transcripts encoding this family of proteins and to correlate each with their proposed protein product, we analyzed the steady-state level of RNAs from infected cells, both in the presence and absence of an inhibitor of viral DNA replication (Fig. 2). Cytoplasmic poly(A)$^+$ RNA was prepared from AD169-infected cells at 2, 8, 26, or 72 h p.i. and from infected cells incubated in the presence of 300 μg of PAA per ml from 26 to 72 and from 0 to 72 h p.i. The infected-cell RNAs were subjected to electrophoresis through denaturing formaldehyde gels and transferred to NYTRAN (Schleicher & Schuell, Inc., Keene, N.H.). The filters were hybridized with a ^{32}P-labeled RNA probe corresponding to the 5′ exon of these RNAs (see Materials and Methods; Fig. 2C). We identified the fully processed species of 2.1 and 2.2 kb, as well as larger transcripts migrating between 2.5 and 2.65 kb. The 2.1- and 2.2-kb RNAs, encoding the 50- and 43-kDa proteins, respectively, reached near maximum abundance by 8 h p.i., remained relatively constant through 26 h p.i., and decreased slightly at late times. As the infection progressed to late times, an increase in complexity was seen, and the larger transcripts, which correspond in size to the species which has spliced out only the first intron and the completely unspliced transcript, could then be detected. In the presence of PAA, from 26 to 72 h p.i., there was a four- to fivefold decrease in the steady-state levels of all of the RNAs, with a slightly greater decrease when the drug was present from time zero. Under these conditions, accumulation of the 2.2-kb RNA species was inhibited to a greater extent than was the 2.1-kb RNA. We also noted that in the presence of PAA the 2.5- to 2.65-kb species appeared to be migrating faster on the gels, indicating a slight decrease in size.

Pattern of expression of the proteins in the absence of viral DNA replication. To determine whether the accumulation of this family of proteins was also dependent on the state of DNA replication and mirrored that of the RNAs, we isolated total cellular proteins from AD169-infected cells at 8, 24, 48, or 72 h p.i. and from infected cells incubated in the presence of 300 μg of PAA per ml from 24 to 48, from 24 to 72, or from 0 to 72 h p.i. The proteins were separated by SDS-PAGE, electroblotted, and probed with antiserum generated against a peptide corresponding to the amino terminus of these

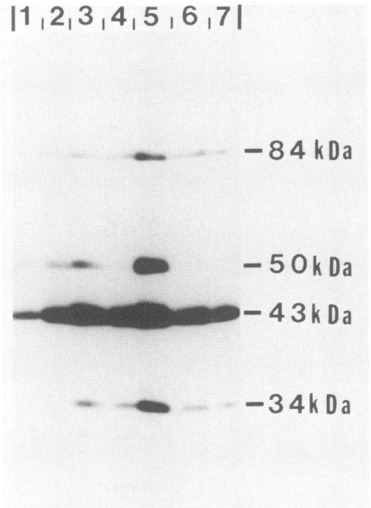

FIG. 3. Immunoblot analysis of lysates from AD169-infected cells incubated in the presence or absence of an inhibitor of viral DNA replication. Lysates (5×10^5 cell equivalents) obtained from cells infected with HCMV for 8 (lane 1), 24 (lane 2), 48 (lane 3), and 72 (lane 5) h, along with infected cells incubated in the presence of 300 μg of phosphonoacetic acid per ml from 24 to 48 (lane 4), 24 to 72 (lane 6), and 0 to 72 (lane 7) h p.i., were separated by SDS-PAGE, transferred to nitrocellulose, probed with immune antiserum, and labeled with ^{125}I-protein A.

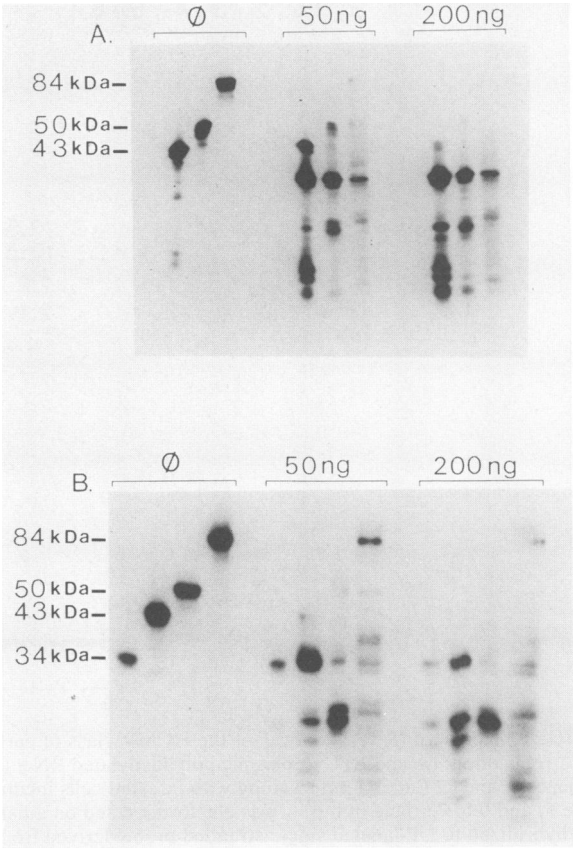

FIG. 4. Partial proteolytic cleavage with *S. aureus* V-8 protease. (A) [^{35}S]methionine-labeled infected-cell proteins were immunoprecipitated with immune antiserum, isolated by preparative SDS-PAGE, subjected to partial proteolysis with V-8 protease (0, 50, or 200 ng), and resolved by SDS-PAGE (15% acrylamide). (B) $^{32}P_i$-labeled infected-cell proteins were isolated as described above, subjected to partial proteolysis with V-8 protease (0, 50, or 200 ng), and resolved by SDS-PAGE (15% acrylamide).

proteins, as previously described (44). In the absence of PAA, the pattern of expression of the proteins was similar to that seen previously (44) (Fig. 3). The 43-, 50-, and 84-kDa proteins could be detected by 8 h p.i. As the infection progressed, the 34-kDa protein could be detected by 24 h p.i., was clearly seen at 48 h p.i., and along with the 50- and 84-kDa proteins increased greatly between 48 and 72 h p.i. When the infected cells were harvested at 48 h p.i. after incubation with PAA from 24 to 48 h p.i., the 43-kDa protein decreased slightly in abundance, compared with that at 48 h p.i. in the absence of the drug, but remained at essentially the same concentration as that found at 24 h p.i. The 84-kDa protein remained at levels similar to those found at 24 and 48 h p.i. in the absence of the drug, while the 34-kDa protein increased slightly in abundance over the levels seen at 24 h p.i. in the absence of the drug. However, the 50-kDa protein was greatly affected and could not be detected under these conditions. In cells incubated from 24 to 72 or from 0 to 72 h p.i. in the presence of PAA, the 84-, 43-, and 34-kDa proteins remained at concentrations which were slightly reduced but similar to that found when the drug was present from 24 to 48 h p.i., as described above. However, again the 50-kDa protein could not be detected. In all cases the final levels of the proteins were decreased in the presence of PAA, compared with those at 72 h p.i. in the absence of the drug.

The accumulation of the 84-, 43-, and 34-kDa proteins (Fig. 3) coincided with the pattern of accumulation of the transcripts (Fig. 2) which we have proposed to encode each of these proteins. However, the abundance of the 50-kDa protein did not correlate with the level of its corresponding RNA. The steady-state level of the 2.1-kb RNA species, which encodes the 50-kDa protein, was only slightly lower than that of the 2.2-kb RNA species, yet the 50-kDa protein was much less abundant than the 43-kDa protein and increased in abundance only at late times. This effect was more apparent in the presence of PAA, when the 2.2-kb RNA species decreased in abundance to levels below detection (Fig. 2), and the 43-kDa protein was still the most abundant species (Fig. 3). Under these conditions, the levels of the 2.1-kb RNA species, although decreased, were easily detected, but the 50-kDa protein could not be detected.

Partial proteolysis. From the analysis described above, we identified the origin of each of the members of this family of proteins and reconfirmed that each of the proteins shares a common amino terminus with variable carboxy termini. Previously, we reported that all of the proteins were phosphorylated in vivo but did not determine whether they were unique or common sites of phosphorylation (44). To determine the pattern of phosphorylation, we used partial proteolysis with *S. aureus* V-8 protease on infected-cell proteins labeled from 48 to 72 h p.i. with [^{35}S]methionine or $^{32}P_i$ (see Materials and Methods; Fig. 4). The V-8 proteolytic profile of [^{35}S]methionine-labeled 43-, 50-, and 84-kDa proteins is shown in Fig. 4A. Each of the proteins generated a common fragment of approximately 34 kDa which comigrated with the in vivo 34-kDa protein (data not shown; compare Fig. 4A with Fig. 4B). This shared band most likely represents the common amino-terminal portion of the proteins. The 43- and 50-kDa proteins also had distinct patterns of lower bands which had some similarities, but the bands migrated at slightly different positions (Fig. 4A). In our previous studies,

we showed that the tryptic peptide maps of the [^{35}S]methionine-labeled 43- and 50-kDa proteins were nearly identical (44). This suggested that all the methionine-containing tryptic peptides were within the shared regions corresponding to the first two exons of the proteins. Sequence analysis has confirmed that there are no methionine residues within the alternative carboxy termini of the 43- and 50-kDa proteins (data not shown). Thus, the lower-molecular-mass bands for the 43- and 50-kDa proteins represent shared methionines which contain either unique carboxy-terminal sequences or common sequences which were differentially cleaved due to an altered sensitivity to the V-8 protease conferred by the carboxy-terminal unique sequences. The 84-kDa protein shared only the amino-terminal fragment and generated a unique pattern of carboxy-terminal fragments which contain the common shared methionines from exon 2 as well as additional methionines from the carboxy terminus. We next compared the partial proteolytic pattern of $^{32}P_i$-labeled proteins with that of the [^{35}S]methionine-labeled proteins described above (Fig. 4B). The 43-kDa protein was phosphorylated predominantly on the amino-terminal shared fragment, while the 50- and 84-kDa proteins were only lightly phosphorylated on this fragment. Most of the phosphorylation of the 50-kDa protein was on the fragments derived from carboxy-terminal sequences, while the 43-kDa protein was only lightly phosphorylated in this region. The 84-kDa protein showed a pattern of carboxy-terminal phosphorylation which was unique and was similar to its ^{35}S pattern described above.

Tryptic peptide mapping. Although the analysis described above showed that each member of this family of proteins was differentially phosphorylated, we could not determine whether these differences were only quantitative or if there were unique as well as shared sites of phosphorylation for each of the proteins. To differentiate between these possibilities, we used tryptic peptide mapping of AD169-infected cell proteins labeled with $^{32}P_i$ from 48 to 72 h p.i. (see Materials and Methods). Tryptic peptide maps of these four proteins were compared (Fig. 5), along with maps of mixes of each of the proteins (data not shown). The maps of each of the proteins were highly divergent, showing unique patterns of phosphorylation with only slight overlap. This can be contrasted with the tryptic peptide maps of the [^{35}S]methionine-labeled proteins, of which the 34-kDa protein was an exact subset of the other proteins representing exon 1, the 43- and 50-kDa proteins were virtually identical, representing exons 1 and 2, and the 84-kDa protein contained the 43- and 50-kDa protein pattern as well as additional carboxy-terminal [^{35}S]methionine-containing peptides (44). The $^{32}P_i$-labeled proteins generated maps in which there appeared to be only a single common phosphopeptide (Fig. 5, spot A). Peptide A was the predominant phosphorylated peptide for the 43-kDa protein and was represented in various amounts in each of the other maps. This is consistent with the partial proteolytic mapping detailed above (Fig. 5), showing that the predominant site of phosphorylation on the 43-kDa protein is within a common region and that this region is less phosphorylated in each of the other proteins. Each of the proteins also generated unique tryptic phosphopeptides which may be derived from alternative sites of phosphorylation within the shared sequences or from unique sites within the carboxy-terminal regions.

The experiments described above demonstrated that these proteins contain both quantitative as well as qualitative differences in phosphorylation. To determine whether these differences could be accounted for, at least in part, by phosphorylation of different amino acids, the phosphoamino acid composition of the proteins was determined. $^{32}P_i$-labeled infected-cell lysates were immunoprecipitated with immune antisera, and the proteins were isolated by preparative SDS-PAGE. Each of the proteins was hydrolyzed to individual amino acids and analyzed by two-dimensional thin-layer electrophoresis (see Materials and Methods). We detected only phosphoserine in all four proteins. There was no evidence that any of the proteins contained phosphothreonine or phosphotyrosine (data not shown). Similar results were obtained when infected-cell proteins were subjected to immunoblot analysis (44) and probed with either immune serum against the proteins or antiphosphotyrosine antiserum (kindly provided by J. Y. J. Wang); again, none of the proteins was found to contain phosphotyrosine (data not shown).

Finally, to determine whether any of the phosphorylation of the proteins could be contributed by ADP-ribosylation, [^{32}P]NAD was incorporated into uninfected and AD169-infected (48 h p.i.) cell nuclei in vitro. Total nuclear lysates as well as immunoprecipitated samples were analyzed by SDS-PAGE (data not shown). Although [^{32}P]ADP-ribose was incorporated into both normal cellular and viral-induced infected-cell proteins, no incorporation was found in any member of this family of proteins. This distinguishes these proteins from the herpes simplex virus type 1 IE polypeptide Vmw175/ICP4 (27), the SV40 large T antigen (15), and the adenovirus T antigen (12), in which this type of modification has been identified.

Evolutionary conservation. The studies described above demonstrated that the expression of this family of proteins is regulated both at the level of transcription (34) and posttranscriptionally by differential RNA processing, translation or stability of the proteins, and phosphorylation. Although we are not yet able to assign a function(s) to these proteins, a relative measure of their importance in the replication of HCMV could be determined by their conservation in other viral isolates. To determine whether these AD169-encoded proteins were conserved among other HCMV strains, we analyzed a common laboratory strain (Towne) and two recent independent clinical isolates (Co and Ro). Total cellular proteins from mock-infected cells or cells infected with AD169, Towne, Co, or Ro strains of HCMV were separated by SDS-PAGE and analyzed by immunoblot, as previously described (44). All four HCMV strains were shown to induce comigrating proteins of 84, 50, 43, and 34 kDa which were recognized by immune serum generated against a synthetic peptide corresponding to the amino terminus of these proteins (Fig. 6). Interestingly, two proteins with migrations slower than that of the 84-kDa protein were detected in the lysates from the Towne, Co, and Ro strains (Fig. 6). The pair from Towne and Ro comigrated, while the pair from Co had a slightly faster migration. These proteins were found to be nonspecific in the sense that in the presence of excess amino-terminal peptide (100 μg/ml), the 84-, 50-, 43-, and 34-kDa proteins could be completely eliminated by competition, whereas the pair of proteins were not and were actually enhanced (Fig. 6). Although these viral-induced proteins are nonspecific, it is interesting that they are not detected in the AD169-infected cells, and their origin is unclear. These data show that these proteins are highly conserved among various HCMV strains, suggesting a potentially important role in the life cycle of the virus.

DISCUSSION

HCMV, like other herpesviruses, exhibits a temporal pattern of gene expression which is subject to transcriptional

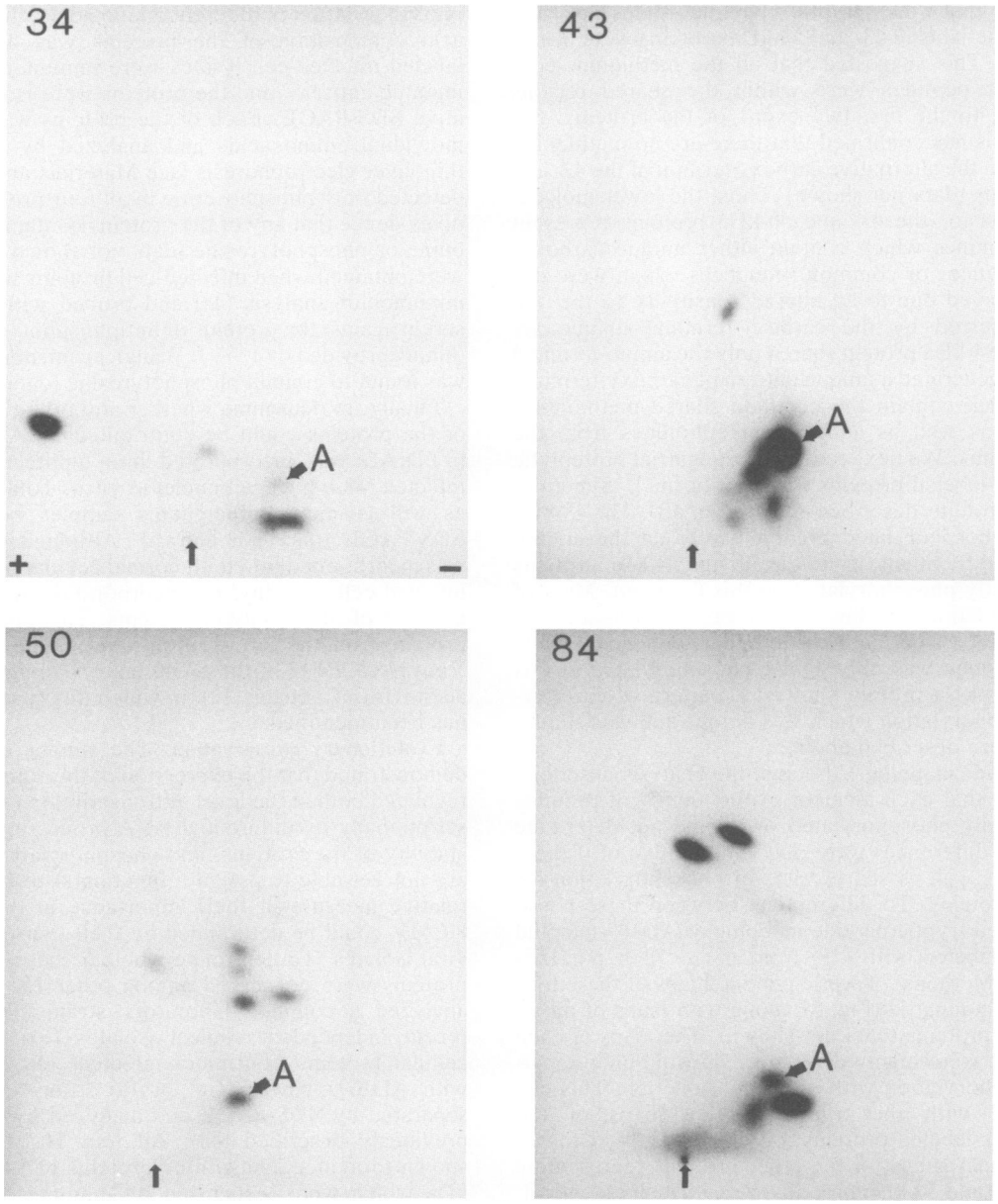

FIG. 5. Tryptic peptide mapping of $^{32}P_i$-labeled infected-cell proteins. Labeled proteins were immunoprecipitated with immune antiserum and isolated by preparative SDS-PAGE, and tryptic peptides were generated. The origin of each tryptic map is indicated (↑), along with the orientation of the anode (+) and the cathode (−). The numbers correspond to the isolated protein used to generate each map. Conditions for electrophoresis and chromatography were identical for all tryptic peptide maps. Spot A is labeled for purposes of discussion in the text.

control. However, HCMV also appears to utilize posttranscriptional regulatory mechanisms. Previous analysis of the steady-state levels of HCMV RNAs indicated that at early times in the infection, although the viral genome appeared to be extensively transcribed, many of these transcripts either failed to be transported to the cytoplasm or were unstable (7, 40). Subsequent studies have given some insight into possible mechanisms involved in the posttranscriptional regulation of HCMV genes. In one case, an alternative processing event at the 3′ end of the RNA occurring late in the infection correlated with the appearance of RNA transcripts in the cytoplasm (14). Other studies indicate that posttranscriptional regulation of certain transcripts may be occurring at the level of translation or stability of the protein products (9–11). In this report, we show that gene expression from the genetic unit contained within the adjacent HCMV strain AD169 *Eco*RI fragments R and d (map units 0.682 to 0.713) is subject to multiple levels of posttranscriptional control.

Previously, we identified a class of early transcripts arising from the adjacent HCMV *Eco*RI fragments R and d and showed that the fully processed RNAs (designated the 2.2-kb class of RNAs) consisted of invariable 5′ and internal exons and a heterogeneous 3′ exon (35). At late times in the infection, additional transcripts from this region were detected, but they were not further characterized. Subsequently, we analyzed the putative promoter regulatory region of this coding unit and determined that it was regulated at the level of transcription by IE gene products (34). With

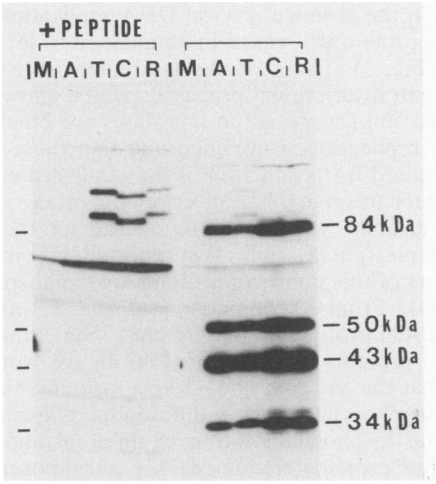

FIG. 6. Evolutionary conservation of the HCMV proteins. Total cellular proteins isolated from mock-infected cells (lanes M) and cells infected with HCMV strains AD169 (lanes A), Towne (lanes T), Co (lanes C), and Ro (lanes R) were separated by SDS-PAGE, transferred to nitrocellulose, probed with immune antiserum either in the presence of 100 μg of free peptide per ml or in the absence of peptide, and labeled with ^{125}I-protein A.

the use of antisera directed against a synthetic peptide corresponding to the predicted amino terminus of the proteins, we were able to show that this class of RNAs specified four nuclear phosphoproteins of 84, 50, 43, and 34 kDa (44). We noted that the 43-kDa protein was the most abundant of the four proteins and that its level of expression remained relatively constant throughout the infection, while expression of the other proteins increased as the infection progressed to late times. The presence of multiple proteins with differential kinetics of synthesis suggested that this transcription unit might have additional levels of regulation which occur posttranscriptionally.

To further address the posttranscriptional regulatory mechanisms involved in the generation of these proteins, we used cDNA cloning to correlate each protein product with a specific mRNA. Four classes of cDNAs were identified. One class of cDNAs corresponds to the unspliced 2.65-kb genomic transcript, which increases in abundance as the infection proceeds to late times (Fig. 2). Transient expression of the genomic clone in COS-7 cells generated all four of the proteins in proportions similar to that seen in AD169-infected cells at late times in the infection. Two other classes of cDNAs correspond to the fully processed transcripts of 2.1 and 2.2 kb, which were previously identified by Staprans and Spector (35) as the predominant early RNAs. The group which had spliced the two common 5′ exons to the larger of the two 3′ heterologous exons and which corresponded to the 2.2-kb RNA encoded the 43-kDa protein. The other class, which had spliced the two common exons to the shorter 3′ exon and thus corresponds to the 2.1-kb RNA, encoded the 50-kDa protein. The final group of cDNAs had spliced out only the first intron and corresponds to a 2.5-kb transcript which does not appear in abundance until late times in the infection. This cDNA class specified predominantly the 84-kDa protein and small amounts of the 43- and 50-kDa proteins, which were probably generated by secondary splicing events in COS-7 cells. The 34-kDa protein was not generated by any clone which had removed the first intron. This further supports our previous suggestion, which was based on a comparison of tryptic peptide maps of an in vitro-synthesized polypeptide with the in vivo 34-kDa protein, that a full-length transcript could encode the 34-kDa protein (44).

The data presented above, as well as previous analysis of the kinetics of accumulation and relative rates of synthesis of this family of proteins as a function of time (44), suggest that this transcription unit is regulated posttranscriptionally by alternative RNA splicing as the infection progresses. This type of posttranscriptional regulatory mechanism generating alternatively spliced RNAs, although not commonly used by herpesviruses, has been noted for several other HCMV genes. For example, the multiple HCMV IE RNAs originating from the IE1- and IE2-coding units likely arise, at least in part, from alternative splicing (17, 36). Alternative splicing has also been seen for the IE RNAs originating from the short unique region of HCMV at an open reading frame designated HQLF1 (42, 43), as well as for the IE RNAs arising from the *Hin*dIII fragments Z and J located in the long unique segment of the genome (23). The general trend in these alternative splicing mechanisms is the transition from highly spliced RNAs to less processed RNAs as the infection progresses to late times, a pattern we also see for the family of early transcripts analyzed in this study. One exception to this trend has been seen with a class of RNAs which are transcribed at early times but are not represented in the cytoplasm until late in the infection (14). In this case, the appearance of the transcripts in the cytoplasm correlates with differential processing of the 3′ ends either by alternative 3′-terminal cleavage and polyadenylation or by differential splicing. The mechanisms responsible for the differential splicing patterns of HCMV RNAs have yet to be elucidated but likely involve modifications of the splicing machinery as the infection progresses. Further studies aimed at the identification and characterization of these modifications should provide important insights into the host cell spliceosome as well as the posttranscriptional regulatory mechanisms used by HCMV.

With assignment of the 84-, 50-, 43-, and 34-kDa proteins to individual transcripts, we were able to compare the patterns of expression of the RNAs with those of the proteins. We showed that as the infection progressed to late times, the concentration of the 2.5- to 2.65-kb RNAs, which specify the 84- and 34-kDa proteins, respectively, increased greatly. This is consistent with our previous report that the 84- and 34-kDa proteins also are produced in abundance only at late times (44). In contrast, the concentration of the fully processed 2.1- and 2.2-kb transcripts was nearly maximal by 8 h p.i., with the 2.2-kb transcript, encoding the 43-kDa protein, being slightly more abundant. Previous analysis of the 3′ exons of these RNAs by quantitative S1 nuclease protection also demonstrated that the 2.2-kb transcript was slightly more abundant than the 2.1-kb transcript and that both species decreased in abundance by 28 h p.i., but their relative proportions did not change (35). These results were surprising in view of our previous data, which showed by pulse-labeling that the 43-kDa protein was the most abundantly translated at all times throughout the infection. Its rate of synthesis was highest at 8 h p.i. and decreased slightly at late times in the infection. On the other hand, the synthesis of the 50-kDa protein was barely detectable at 8 h p.i., and its rate of synthesis increased as the infection progressed (44). Thus, it appears that expression of the 50-kDa protein is subject to additional posttranscriptional control. The mechanisms underlying this posttranscriptional control are unclear. Given the small difference in the 3′

exons of the 2.2- and 2.1-kb mRNAs, one possible explanation for the lack of correlation between the concentration of the 2.1-kb RNA and the synthesis of the 50-kDa protein is that at early times in the infection the 2.1-kb RNA, although located in the cytoplasm, is blocked in transport to polysomes or not translated efficiently. Then, as the infection progresses to late times, these RNAs are translated more efficiently. Alternatively, since the carboxy-terminal regions of the 50- and 43-kDa proteins differ, these proteins may exhibit differential stability. This latter explanation is difficult to reconcile with our previous pulse-chase analysis, which showed that when infected-cell monolayers at 8 h p.i. were pulsed for 10 min with [^{35}S]methionine and chased for up to 120 min, the 50-kDa protein was as stable as the 43-kDa protein (44). It is possible, however, that the 50-kDa protein is transiently unstable either during or immediately after translation and that after localization to the nucleus it becomes stabilized.

We have also determined that the accumulation of this class of HCMV early RNAs encoded by *Eco*RI fragments R and d is dependent on the state of viral DNA replication. In particular, when viral DNA replication is inhibited, the steady-state concentration of these transcripts is reduced. The dependence of the accumulation of these transcripts on continuing viral DNA replication is in some ways reminiscent of that seen for leaky-late RNAs, which may be present at early times but increase in abundance at late times following amplification of the viral template. However, there is an important distinction. In the absence of inhibitors, the overall abundance of the class of RNAs analyzed in this study does not increase at late times, despite a large amplification of template. Therefore, this pseudodependence on DNA replication is more likely the result of this true early transcription unit being greatly down regulated as viral template levels increase, giving the appearance of a relatively constant level of RNA production. This down regulation could be the result of either a decrease in some positive regulatory factor(s) or the production of a negative regulatory factor(s), the expression of which is not dependent on the state of viral DNA replication.

Although the overall production of this family of RNAs and their protein products was diminished when viral DNA replication was inhibited, we also observed differential susceptibility of the proteins to this inhibition, as well as changes in the mobility of some of the RNAs. Specifically, we noted that when the 2.5- to 2.65-kb RNAs, which specify the 84- and 34-kDa proteins, respectively, were isolated in the presence of an inhibitor of viral DNA replication, they migrated slightly faster than those isolated in the absence of the drug. Since Staprans and Spector (35) reported that at late times an alternative start site, which is approximately 60 bp upstream of the 5′ end of the RNAs, is used, it is possible that the shift to the upstream start site may be inhibited. Nevertheless, the levels of these transcripts paralleled the accumulation of their protein products. However, this was not the case for the 43- and 50-kDa proteins. In the absence of viral DNA replication, there was a disproportional loss of the 2.2-kb transcript, encoding the 43-kDa protein (Fig. 2). Despite the decrease in abundance of the 2.2-kb RNA, the 43-kDa protein still accumulated to higher levels than did the other proteins (Fig. 3). These data support the previous results, which indicated that the 2.2-kb RNA was most efficiently translated at early times, and imply that this mechanism is retained when the infection is held in an early state by inhibition of viral DNA replication. In contrast, although the concentration of the 2.1-kb RNA decreased somewhat in the absence of viral DNA replication (Fig. 2), the 50-kDa protein decreased in abundance to levels below detection (Fig. 3). These results lend support to the ideas that the posttranscriptional processes which allow accumulation of the 50-kDa protein at late times are blocked when viral DNA replication is inhibited and that these functions are not required for production of the other proteins.

A final mechanism of posttranscriptional processing of this family of proteins was found to operate at the level of differential phosphorylation. We reported previously that the members of this family of proteins were phosphorylated in vivo (44). From a comparison of the *S. aureus* V-8 protease partial proteolytic patterns of [^{35}S]methionine- and $^{32}P_i$-labeled proteins shown here (Fig. 4), we were able to establish that the 84-, 50-, and 43-kDa proteins were differentially phosphorylated. This differential phosphorylation was not due to phosphorylation of different amino acids, since all four proteins contained only phosphoserine (data not shown). The 43-kDa protein was predominantly phosphorylated in the amino-terminal region encoded by the first exon, while the 50- and 84-kDa proteins were only lightly phosphorylated in this region. Tryptic peptide mapping of $^{32}P_i$-labeled proteins confirmed that the differences in the phosphorylation patterns of the proteins were not only quantitative but were qualitative as well. The tryptic maps of all four proteins showed unique patterns of phosphorylation, with only a single common phosphopeptide (Fig. 5, spot A). The shared phosphopeptide contained the most abundant site of phosphorylation for the 43-kDa protein and was present in lesser amounts in each of the other proteins. This supports the partial proteolytic data described above, which showed that the predominant site of phosphorylation for the 43-kDa protein was in a common region and that this region was quantitatively less phosphorylated in the other three proteins. The functional significance of the alternative sites of phosphorylation in the HCMV proteins is not clear. Further characterization of the biochemical properties of the proteins will be required before the role alternative phosphorylation plays in the activity of the proteins can be addressed.

In summary, these studies have revealed that this family of proteins has multiple levels of regulation. Transcription from this early gene is highly regulated and is dependent on IE gene expression for induction (34) and on an unknown mechanism for down regulation late in the infection. Production of the proteins is regulated by alternative RNA splicing, which varies as the infection progresses to late times by decreasing the overall levels of splicing. The relative abundance of the proteins is also determined by differential translation or protein stability, which allows increased levels of the 50-kDa protein to be produced at late times. Finally, the identification of alternative phosphorylation patterns for each of the proteins, despite significant regions of homology, reveals yet another level of posttranscriptional regulation involved in their synthesis. Further characterization of the functions of this complex class of highly conserved HCMV nuclear phosphoproteins and the posttranscriptional mechanisms involved in their production should help elucidate their role in the replication cycle of HCMV.

ACKNOWLEDGMENTS

We thank D. Jaquish for excellent technical assistance and D. David for preparation of the manuscript.

This investigation was supported by Public Health Service grant 34729 awarded by the National Cancer Institute to D. H. Spector

and by Public Health Service training grant GM07198 from the National Institutes of Health.

LITERATURE CITED

1. **Boshart, M., F. Weber, G. Jahn, K. Dorsch-Hasler, B. Fleckenstein, and W. Schaffner.** 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. Cell **41**:521–530.
2. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1002–1006.
3. **Conley, A. J., D. M. Knipe, P. C. Jones, and B. Roizman.** 1981. Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of γ polypeptides. J. Virol. **37**:191–206.
4. **Cooper, J. A., B. M. Sefton, and T. Hunter.** 1983. Detection and quantification of phosphotyrosine in proteins. Methods Enzymol. **99**:387–402.
5. **DeClerq, E.** 1982. Specific targets for antiviral drugs. Biochem. J. **205**:1–13.
6. **DeMarchi, J. M.** 1981. Human cytomegalovirus DNA: restriction enzyme cleavage and map locations for immediate-early, early and late RNAs. Virology **114**:23–28.
7. **DeMarchi, J. M.** 1983. Posttranscriptional control of human cytomegalovirus gene expression. Virology **124**:390–402.
8. **Elion, G. B.** 1983. The biochemistry and mechanism of acyclovir. J. Antimicrob. Chemother. **12**(Suppl. B):9–17.
9. **Geballe, A. P., F. S. Leach, and E. S. Mocarski.** 1986. Regulation of cytomegalovirus late gene expression: γ genes are controlled by posttranscriptional events. J. Virol. **57**:864–874.
10. **Geballe, A. P., and E. S. Mocarski.** 1988. Translational control of cytomegalovirus gene expression is mediated by upstream AUG codons. J. Virol. **62**:3334–3340.
11. **Geballe, A. P., R. R. Spaete, and E. S. Mocarski.** 1986. A cis-acting element within the 5′ leader of a cytomegalovirus β transcript determines kinetic class. Cell **46**:865–872.
12. **Goding, C. R., C. H. Shaw, G. E. Blair, and W. C. Russell.** 1983. ADP-ribosylation in *in vitro* systems synthesizing adenovirus DNA. J. Gen. Virol. **64**:477–483.
13. **Godowski, P. J., and D. M. Knipe.** 1983. Mutations in the major DNA-binding protein gene of herpes simplex virus type 1 result in increased levels of viral gene expression. Virology **47**:478–486.
14. **Goins, W. F., and M. F. Stinski.** 1986. Expression of a human cytomegalovirus late gene is posttranscriptionally regulated by a 3′-end-processing event occurring exclusively late after infection. Mol. Cell. Biol. **6**:4202–4213.
15. **Goldman, N., M. Brown, and G. Khoury.** 1981. Modification of SV40 T antigen by poly(ADP-ribosyl)ation. Cell **24**:567–572.
16. **Helgstrand, E., B. Eriksson, N. G. Johansson, B. Lannero, A. Larsson, A. Misiorny, J. O. Noren, B. Sjoberg, K. Stenberg, G. Stenig, S. Stridh, B. Oberg, S. Alenius, and L. Philipson.** 1978. Trisodium phosphonoformate, a new antiviral compound. Science **201**:819–821.
17. **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.
18. **Ho, M.** 1982. Cytomegalovirus: biology and infection. Plenum Publishing Corp., New York.
19. **Honess, R. W., and B. Roizman.** 1974. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. **14**:8–19.
20. **Huang, E.-S.** 1975. Human cytomegalovirus. III. Virus-induced DNA polymerase. J. Virol. **16**:298–310.
21. **Huang, E.-S.** 1975. Human cytomegalovirus. IV. Specific inhibition of virus-induced DNA polymerase activity and viral DNA replication by phosphonoacetic acid. J. Virol. **16**:1560–1565.
22. **Hunter, T., and B. M. Sefton.** 1980. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc. Natl. Acad. Sci. USA **77**:1311–1315.
23. **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.
24. **McDonough, S. H., and D. H. Spector.** 1983. Transcription in human fibroblasts permissively infected by human cytomegalovirus strain AD169. Virology **125**:31–46.
25. **Nankervis, G. A., and M. L. Kumar.** 1978. Diseases produced by cytomegaloviruses. Med. Clin. North Am. **62**:1021–1035.
26. **Nishiyama, Y., K. Maeno, and S. Yoshia.** 1983. Characterization of human cytomegalovirus-induced DNA polymerase and associated 3′ and 5′ exonuclease. Virology **124**:221–231.
27. **Preston, C. M., and E. L. Notarianni.** 1983. Poly(ADP-ribosyl)-ation of a herpes simplex virus immediate early polypeptide. Virology **131**:492–501.
28. **Rapp, F.** 1983. The biology of cytomegaloviruses, p. 1–66. *In* B. Roizman (ed.), The herpesviruses. Plenum Publishing Corp., New York.
29. **Shipman, C., Jr., S. H. Smith, J. C. Drach, and D. L. Klayman.** 1981. Antiviral activity of 2-acetylpyridine against herpes simplex virus. Antimicrob. Agents Chemother. **19**:682–685.
30. **Spaete, R. R., and E. S. Mocarski.** 1985. Regulation of cytomegalovirus gene expression: α and β promoters are *trans* activated by viral functions in permissive human fibroblasts. J. Virol. **56**:135–143.
31. **Spector, D. H.** 1985. Molecular studies on the cytomegaloviruses of mice and men, p. 199–234. *In* J. K. Setlow and A. Hollaender (ed.), Genetic engineering: principles and methods, vol. 7. Plenum Publishing Corp., New York.
32. **Spector, D. H., L. Hock, and J. C. Tamashiro.** 1982. Cleavage maps for human cytomegalovirus DNA strain AD169 for restriction endonucleases *Eco*RI, *Bgl*II, and *Hin*dIII. J. Virol. **42**: 558–582.
33. **Spector, D. H., and S. A. Spector.** 1984. The oncogenic potential of human cytomegalovirus. Prog. Med. Virol. **29**:45–89.
34. **Staprans, S. I., D. K. Rabert, and D. H. Spector.** 1988. Identification of sequence requirements and *trans*-acting functions necessary for regulated expression of a human cytomegalovirus early gene. J. Virol. **62**:3463–3473.
35. **Staprans, S. I., and D. H. Spector.** 1986. 2.2-Kilobase class of early transcripts encoded by cell-related sequences in human cytomegalovirus strain AD169. J. Virol. **57**:591–602.
36. **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.
37. **Stinski, M. F.** 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. J. Virol. **23**:751–767.
38. **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.
39. **Tamashiro, J. C., L. J. Hock, and D. H. Spector.** 1982. Construction of a cloned library of the *Eco*RI fragments from the human cytomegalovirus genome (strain AD169). J. Virol. **42**: 547–557.
40. **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.
41. **Wathen, M. W., D. R. Thomsen, and M. F. Stinski.** 1981. Temporal regulation of human cytomegalovirus transcription at immediate early and early times after infection. J. Virol. **38**: 446–451.
42. **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.
43. **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.
44. **Wright, D. A., S. I. Staprans, and D. H. Spector.** 1988. Four phosphoproteins with common amino termini are encoded by human cytomegalovirus AD169. J. Virol. **62**:331–340.