Organization and Expression of the Immediate Early Genes of Human Cytomegalovirus

MARK F. STINSKI,^{1*} DARRELL R. THOMSEN,¹ RICHARD M. STENBERG,¹ AND LYNN C. GOLDSTEIN²

Department of Microbiology, School of Medicine, University of Iowa, Iowa City, Iowa 52242,¹ and Genetic Systems Corp., Seattle, Washington 98121²

Received 8 October 1982/Accepted 15 December 1982

The immediate early genes of human cytomegalovirus were characterized according to map location, RNA transcripts, and translation products. Three regions in the large unique component (0.709 to 0.751 map units) were transcribed in the presence of an inhibitor of protein synthesis (anisomycin). A single size class of polyadenylated mRNA, 1.95 kilobases (kb), was transcribed abundantly relative to the other size classes. The predominant 1.95-kb viral RNA was transcribed from right to left on the prototype arrangement of the viral genome and spanned a region of approximately 2.8 kb (0.739 to 0.751 map units). This mRNA codes for a 75,000-dalton protein that represents the predominant immediate early protein detected in infected cells. Immunoprecipitation of viral proteins synthesized *in vitro* as well as *in vivo* demonstrated that the predominant immediate early protein is synthesized as a protein of 75,000 daltons, but is presumably modified *in vivo*, resulting in a broad banding pattern ranging from 75,000 to 68,000 daltons. A different immediate early viral gene (0.732 to 0.739 map units) is transcribed from left to right at relatively low levels. The 3' ends of the above viral RNAs terminate at approximately 230 base pairs apart in the region of approximately 0.739 map units. Five RNA size classes ranging from 2.25 to 1.10 kb were detected, but the 1.75-kb and 1.40-kb RNA size classes were more abundant from this region. At least four minor proteins are coded by these mRNAs, with apparent molecular weights ranging from 56,000 to 16,500. Last, a 1.95-kb mRNA was transcribed from a third region (0.709 to 0.728 map units). This viral mRNA was present at relatively low concentration and coded for a minor protein of 68,000 daltons. Since immediate early gene expression of human cytomegalovirus is dominated by the synthesis of an mRNA from the region of 0.739 to 0.751 map units that codes for the predominant immediate early protein found in the infected cell, we hypothesize that this protein is the major regulatory protein influencing the switch from restricted to extensive transcription.

The diseases induced by human cytomegalovirus (CMV) frequently represent infections after reactivation of latent virus (20, 40). Reactivation is broadly considered as the event that allows for the replication of the viral DNA and the eventual release of infectious virus to cause overt disease. Three broad phases of CMV gene expression have been described (6, 12, 15, 46, 48, 52, 53): (i) the immediate early stage, which occurs in the absence of *de novo* protein synthesis; (ii) the early stage, which requires the action of at least one immediate early gene; and (iii) the late stage, which can be detected after viral DNA synthesis. The fraction of the genome transcribed increases as infection progresses (9).

The first viral genes expressed after reactivation or after primary infection presumably code for a viral regulatory protein(s) that controls

subsequent viral gene expression. These genes are hypothesized to be the immediate early (IE) viral genes, which are expressed independently of any preceding viral protein synthesis. Transcription of the IE viral genes of the Towne (52, 53) and the Davis (12) strains of human CMV is restricted primarily to a region between 0.66 and 0.77 map units (*Xba*I-N and -E) in the large unique section of the viral genome. IE viral RNAs that originate from the above region are associated with the polyribosomes as polyadenylated [poly(A)] RNA of 4.8, 2.2, and 1.9 kilobases (kb) (52, 53). The IE viral genes represent a restricted and readily definable class of viral genes. The viral RNA encoded by the IE viral genes constitutes approximately 0.6% of the infected cell RNA (9).

IE gene expression of CMV (Towne) is char-

acterized by the presence of an abundant mRNA of approximately 1.9 kb and a predominant protein of 75,000 to 68,000 daltons that is phosphorylated *in vivo* (18, 46, 53). The size of the viral protein varies slightly among different strains of CMV (7, 18) and migrates heterogeneously in denaturing sodium dodecyl sulfate (SDS)-polyacrylamide gels, suggesting that the protein undergoes post-translational modifications. Other IE viral mRNAs and proteins are also detectable, but they are present in the infected cell at relatively lower concentrations (46, 47, 52, 53). After synthesis of the IE viral proteins, there is a switch from restricted transcription to extensive transcription of the viral genome. Synthesis of CMV early RNAs is dependent upon the function of at least one IE viral gene product. Inhibition of viral protein synthesis with inhibitors such as cycloheximide (12, 13, 52, 53) or by treatment with interferon (47) has little to no effect on IE RNA transcription, but the switch from restricted to extensive transcription is inhibited. Therefore, it is proposed that the IE proteins of CMV have an important role in controlling viral gene expression.

This report describes the organization of the IE genes of CMV in the region of 0.709 to 0.751 map units (*Xba*I-E). The size of the transcripts in the absence of *de novo* protein synthesis, the direction of transcription, and the proteins coded for by the IE mRNAs are reported. The IE gene expression of this herpesvirus is dominated by the transcription of a single gene that allows for the translation of a predominant IE protein. The emphasis of this report is on this predominant IE gene. The adjacent IE genes are analyzed for the purpose of comparison. The transcription in the adjacent region of approximately 0.660 to 0.685 map units (*Xba*I-N) at various times after infection requires further investigation.

MATERIALS AND METHODS

Virus and tissue culture. Human fibroblast cells and the plaque purification of human CMV (Towne strain) were previously described (46). The amount of infectious virus was determined by assays for plaques (57) or tissue culture infective doses (17).

Viral infection and definitions of IE. For isolation of IE RNA, cells treated with 100 μ M anisomycin 1 h before infection were infected with CMV at a multiplicity of 10 to 20 PFU per cell in the presence of anisomycin. After 12 h in the presence of anisomycin, the cells were harvested. At the concentrations used in these experiments, anisomycin inhibited protein synthesis by 99% or greater. For pulse-labeling IE proteins, cells were treated with 200 μ g of cycloheximide per ml instead of anisomycin. After removal of the cycloheximide, infected cells were pulse-labeled with [³⁵S]methionine for 3 h in the presence of high salt and actinomycin D as previously described (46). In addition, infected cells as well as uninfected cells were

pulse-labeled with [³⁵S]methionine for 3 h without treatment with cycloheximide or actinomycin D. Infected cells were pulse-labeled in high salt, whereas uninfected cells were pulse-labeled in normal salt as previously described (46).

Isolation of RNA. All reagents, plasticware, and glassware were treated with 0.1% diethyl pyrocarbonate (Sigma Chemical Co., St. Louis, Mo.) and autoclaved before use. The cells were suspended in 25 mM Tris-hydrochloride (pH 7.5) containing 25 mM NaCl, 5 mM MgCl₂, 2% Triton X-100, 5% RNase-free sucrose, 100 μ g of heparin per ml, and 40 U of RNasin (Biotec, Madison, Wis.) and disrupted with a Dounce homogenizer (B pestle). Polyome-associated RNA was isolated by the magnesium precipitation method of Palmiter (35) as previously described (52, 53). Poly(A) RNA was selected from total polysome-associated RNA by oligodeoxythymidylic acid-cellulose chromatography (type 2; Collaborative Research, Inc., Waltham, Mass.) as previously described (53).

Physical map of the *Xba*I-E region. The cloning, purification, and characterization of recombinant plasmids containing insertions of CMV DNA have been described previously (51). Recombinant plasmids pCB45 and pCB42 were a gift from R. La Femina and G. Hayward. The recombinant plasmid containing the *Xba*I-E insert was mapped by the end-labeling and partial digestion method of Smith and Birnstiel (42) and by double restriction enzyme digestions of various fragments of the *Xba*I-E region. Restriction endonucleases were obtained from Bethesda Research Laboratories, Inc., Rockville, Md. The conditions were as described by the supplier. DNA fragments were fractionated by electrophoresis in 1.0 to 1.5% agarose gels by the method of Bachi and Arber (3). Various DNA fragments were isolated from agarose gels by electroelution and ethanol precipitated.

Preparation of radioactive probes. Recombinant plasmid DNAs were labeled with [α -³²P]dCTP by nick translation as described by Rigby et al. (41). Radioactive cDNA probes to viral RNA were prepared as follows. Approximately 35 μ g of IE poly(A) RNA was used for selective hybridization to *Xba*I-E DNA covalently bound to diaminobenzyloxymethyl (DBM)-cellulose as previously described (53). After elution from the DNA-cellulose, the RNAs were rechromatographed on oligodeoxythymidylic acid-cellulose and then used as templates for the synthesis of ³²P-labeled cDNA. Synthesis of cDNA to the complete sequence of the mRNA (total cDNA) by reverse transcriptase (avian myeloblastosis virus) was primed with 2.5 μ g of random oligodeoxynucleotide from calf thymus DNA in 50 μ l of reaction mixture. Approximately 200 ng of viral RNA and 12.5 U of reverse transcriptase in 25 mM Tris (pH 8.3) containing 35 mM KCl, 5 mM MgCl₂, 2.0 mM dithiothreitol, 15 μ M each dATP, dGTP, and TTP, 3 μ M dCTP, and 100 μ Ci of [α -³²P]dCTP (3,000 Ci/mmol; Amersham Corp, Arlington Heights, Ill.) was incubated at 37°C for 1 h. The reaction was stopped by the addition of SDS to 0.5% and EDTA to 20 mM. The reaction mixture was treated with 25 μ g of proteinase K at 37°C for 45 min and then phenol-chloroform extracted. RNA template was degraded by incubation at 37°C in 0.3 M NaOH for 16 h. The cDNA was separated from unincorporated nucleotide triphosphates by gel filtration on Sephadex G-50.

The cDNA to the 3' end of the RNAs (3' cDNA) was synthesized as described above, except the RNA template was alkaline degraded in 0.1 N NaOH for 1 h at room temperature before chromatography on oligodeoxythymidylic acid-cellulose and 0.33 μ g of oligodeoxythymidylic acid (Collaborative Research, Inc., Waltham, Mass.) was used as primer. The average size of 3' cDNA was determined to be 180 nucleotides by methylmercury gel electrophoresis.

Gel electrophoresis of denatured poly(A) RNA. Poly(A) RNA was fractionated by electrophoresis in a 1.5% agarose slab gel containing 10 mM methylmercury hydroxide (Alpha, Danvers, Mo.) as described by Bailey and Davidson (4). Approximately 10 μ g of IE poly(A) RNA per cm² of gel slot was used for electrophoresis in preparative slab gels. Electrophoresis was at 11.5 mA/cm² at 15°C for 4.5 h. Molecular weight standards were 23S (3.3-kb) and 16S (1.7-kb) *Escherichia coli* rRNA (5) and 28S (5.3-kb) and 18S (2.0-kb) human cell rRNA (32). The standards were visualized by staining in 1 μ g of ethidium bromide per ml and 0.5 M ammonium acetate. The sizes of the viral mRNAs were interpolated from a standard curve.

Northern blot hybridization. Gels containing IE polysome-associated, poly(A) RNA were prepared for transfer to DBM-paper (Schleicher & Schuell Co., Keene, N.H.) as previously described (53). The DBM-paper was activated as described by Schleicher & Schuell Co. The RNA blots were incubated at 42°C for 24 h in prehybridization buffer containing 5 $\times$ SSPE (1 $\times$ SSPE is 10 mM NaPO₄ [pH 7.7], 0.18 M NaCl, 1 mM EDTA), 50% formamide, 0.08% (wt/vol) each bovine serum albumin, Ficoll, and polyvinylpyrrolidone, 1% glycine, 0.1% SDS, and 200 μ g of denatured calf thymus DNA per ml. Radioactive probes representing various sections of the *Xba*I-E region and having a specific activity of 1×10^8 to 2×10^8 cpm/ μ g were heated to 100°C for 3 min and then added to hybridization buffer prepared as described above, except the bovine serum albumin, Ficoll, and polyvinylpyrrolidone were 0.04% each. Hybridizations were with an equal molar quantity of DNA and radioactivity, depending on the size of the DNA fragment. Hybridization was at 42°C for 72 h. After hybridization, the RNA blots were washed three times with 5 $\times$ SSPE containing 0.1% SDS at room temperature and three times with 0.1 $\times$ SSPE containing 0.1% SDS at 50°C. Hybridization of the ³²P-labeled DNA probes to CMV-specific RNA was detected by autoradiography with X-Omat AR film (Eastman Kodak Co., Rochester, N.Y.).

Southern blot hybridization. Recombinant plasmid pMW34 (*Bam*HI-A fragment of the *Xba*I-E region) (Fig. 1) was digested with restriction enzymes *Bam*HI and *Pst*I (1 U/ μ g; Bethesda Research Laboratories). Recombinant plasmid pCB42 (*Bam*HI-B fragment) (Fig. 1) was digested with restriction enzymes *Bam*HI, *Pst*I, and *Sal*I. The DNA fragments were fractionated in 1.0% agarose gels by the method of Bach and Arber (3) and then immobilized onto nitrocellulose filters by the method of Southern (43). Filters containing immobilized DNA were pretreated at 42°C for 24 h with prehybridization buffer containing 5 $\times$ SSPE, 50% formamide, 1.0% each bovine serum albumin, Ficoll, and polyvinylpyrrolidone, and 200 μ g of denatured calf thymus DNA per ml. ³²P-labeled recombinant plasmid *Xba*I-E DNA, ³²P-labeled total cDNA synthesized

from viral RNA template complementary to *Xba*I-E, and ³²P-labeled 3' cDNA synthesized from the 3' ends of viral RNA template complementary to *Xba*I-E were prepared as described above. Approximately 500,000 cpm of each ³²P-labeled probe was suspended in water, heated to 100°C for 3 min, and then added to hybridization buffer prepared as described above, except the bovine serum albumin, Ficoll, and polyvinylpyrrolidone were 0.04% each. Hybridization was at 42°C for 48 h. The amount of immobilized recombinant plasmid DNA was at least 50-fold greater than the amount of input ³²P-labeled probe. After hybridization, the filters were washed twice with 2 $\times$ SSPE containing 0.1% SDS at room temperature and twice with 0.1 $\times$ SSPE containing 0.1% SDS at 50°C. After final washing with 2 $\times$ SSPE, hybridization of ³²P-labeled probe was detected by autoradiography.

Hybridization of viral RNA to DNA bound to DBM-paper. Recombinant plasmid DNAs representing the *Pst*I-D, *Sac*I-A, and *Sal*I-C (Fig. 1) sections of the *Xba*I-E region or the entire *Xba*I-E region of the viral genome were digested with restriction enzyme *Pst*I or *Bam*HI, denatured, and covalently linked to DBM-paper by the method of Stark and Williams (44). Each DNA preparation had approximately 10⁵ cpm of ³²P-labeled recombinant plasmid for estimating the percent DNA linkage to the DBM-paper. Generally, 40% of the input radioactivity remained bound to the DBM-paper, which was extrapolated to approximately 15 μ g of recombinant plasmid DNA linkage. The linked DNA was at approximately 150-fold excess for each microgram of RNA hybridized.

The input IE poly(A) RNA was 20 μ g for *Pst*I-D DNA, 80 μ g for *Sac*I-A DNA, and 100 μ g for *Sal*I-C DNA. After the poly(A) RNA was heated at 80°C for 2 min and cooled on ice, it was added to a hybridization buffer of 20 mM PIPES (piperazine-*N,N'*-bis(2-ethanesulfonic acid), pH 6.4, containing 0.4 M NaCl, 5 mM EDTA, 50% deionized formamide, 0.2% SDS, and 1 μ g of tRNA from calf liver (Boehringer Mannheim Corp., Indianapolis, Ind.) per ml. Hybridization was at 57°C for 5 h. The DBM-paper was then washed three times with 2 $\times$ SSPE containing 0.2% SDS at room temperature and six times with 0.1 $\times$ SSPE containing 0.2% SDS at 60°C for 15-min periods. The hybridized RNA was eluted in 99% deionized formamide containing 10 mM PIPES (pH 6.4) at 70°C. Under these conditions approximately 20 to 100 ng of virus-specific RNA was isolated. The viral RNA was ethanol precipitated with 10 μ g of calf liver tRNA as carrier.

R-loops and electron microscopy. IE polysome-associated, poly(A) RNA was selected by hybridization to recombinant plasmid *Xba*I-E DNA bound to DBM-cellulose by the method of Noyes and Stark (34) as previously described (1, 53). Approximately 50 μ g of IE poly(A) RNA was hybridized to approximately 50 μ g of recombinant plasmid *Xba*I-E that was cleaved by digestion with the restriction endonuclease *Bam*HI and bound to cellulose. Hybridization was for 16 h at 57°C in 0.1 M HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), pH 8.0, containing 70% deionized formamide, 0.5 M Na⁺, and 0.005 M EDTA. After hybridization, the cellulose was rinsed twice with 2 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at room temperature, twice with hybridization buffer at room temperature, and three times with hybridization buffer at 60°C. The hybrid-

ized RNA was eluted in 99% formamide containing 10 mM HEPES (pH 8.0) at 60°C. The eluent was diluted to a final formamide concentration of 22% and adjusted to 0.2 M sodium acetate, and then the RNA was ethanol precipitated. Approximately 5% of the input RNA was recovered. The selected IE RNA (2 to 3 µg) was incubated at 80°C for 2 min, cooled on ice, and then incubated at 65°C for 10 min with 1 µg of *Xba*I-digested, recombinant plasmid *Xba*I-E DNA in hybridization buffer prepared as described above. The mixture was then incubated at 57°C for 16 h. Hybridization of RNA to double-stranded DNA and the formation of R-loops were detected by electron microscopy. The samples were mounted for electron microscopy by the formamide technique, stained with uranyl acetate, and shadowed with platinum-palladium by the method of Thomas et al. (50). Vector plasmid pACYC 184 (4.2 kb) and M₁₃ DNA (7.2 kb) were used as size standards. R-loops present in 18 representative molecules were analyzed.

In vitro translation of RNA and analysis of polypeptide products. IE poly(A) RNA selected by hybridization was ethanol precipitated twice in 2% sodium acetate and once in 2% potassium acetate. The viral RNA was dissolved in a minimal volume of sterile water, freeze-dried, and suspended in 10 µl of sterile water. Viral RNA was translated in vitro by using a micrococcal nuclease-treated rabbit reticulocyte lysate prepared by the method of Pelham and Jackson (37). Each 22-µl assay mixture contained 10 µl of reticulocyte lysate, 25 µCi of [³⁵S]methionine (1,225 Ci/mmol; Amersham Corp, Arlington Heights, Ill.), and 10 µl of mRNA in water. After incubation at 37°C for 60 min, 5 µl of pancreatic RNase (100 µg/ml) and 20 U of T₁ RNase per ml in 25 mM EDTA were added, and the mixture was incubated for an additional 10 min at 37°C.

Infected cells were pulse-labeled with [³⁵S]methionine under IE conditions as previously described (46).

Polypeptides from in vitro translations, pulse-labeled infected cells, or immunoprecipitates were fractionated by electrophoresis in discontinuous SDS-polyacrylamide slab gels by the method of Laemmli (26) as previously described (46). After the gels were stained and destained as previously described (45), they were soaked in water, and then 1 M sodium salicylate was added for fluorography (8); the gels were dried and exposed to Kodak X-Omat AR film. Molecular weights were determined from their migration relative to unlabeled molecular weight markers (45, 46) and labeled adenovirus polypeptides from in vitro translation (31).

Immunoprecipitation of virus-specified polypeptides. Infected and uninfected cells were pulse-labeled with [³⁵S]methionine as described above. Extracts of [³⁵S]methionine-labeled cells were prepared for immunoprecipitation as described by Goldstein et al. (19). The extraction buffer contained 10 µg of phenylmethylsulfonyl fluoride per ml to inhibit proteolytic enzyme activity. Immunoprecipitation of immune complexes by Formalin-fixed *Staphylococcus aureus* Cowan I strain was by the method of Kessler (25). Appropriate samples of [³⁵S]methionine-labeled proteins containing 10⁵ cpm were incubated at 4°C for 60 min with a 1:40 dilution of normal mouse serum and an equal volume of a 10% suspension of *S. aureus*. The *S. aureus* was pelleted by centrifugation, and the super-

natants were removed. Monoclonal antibody E-3 (19) or normal mouse serum was added to the supernatants at a final dilution of 1:100. After 60 min at 4°C, an equal volume of a 10% suspension of *S. aureus* was added, and the mixture was incubated at 4°C for 60 min. The *S. aureus* was pelleted by centrifugation, and the pellet was washed five times in 0.1 M Tris-hydrochloride (pH 8.0) containing 0.5 M LiCl and 1.0% beta-mercaptoethanol. The sample was then suspended in dissociating solution for SDS-polyacrylamide gel electrophoresis (26) and boiled for 3 min, and the *S. aureus* was pelleted by centrifugation. Bromophenol blue and glycerol were added to the supernatants, and the samples were fractionated by SDS-polyacrylamide gel electrophoresis as previously described (45, 46).

The gels were stained, destained, and prepared for fluorography as described above. Molecular weight standards were also as described above.

RESULTS

Physical map of the *Xba*I-E DNA. Abundant IE RNA of human CMV originates from the *Xba*I-N (6.9-kb) and -E (20.0-kb) region (0.660 to 0.770 map units) of the viral genome (52, 53). Physical maps of the *Xba*I-E DNA were generated by the method of Smith and Birnstiel (42) and by double restriction enzyme digestions of various fragments of the *Xba*I-E region. Figure 1 illustrates the various physical maps of the *Xba*I-E region and the relationship of this region to the *Xba*I physical map of the entire viral genome determined by La Femina and Hayward (27; submitted for publication). The *Ava*I, *Sma*I, and *Hinc*II sites were mapped by double restriction enzyme digestion of defined subclones of the *Xba*I-E region; consequently, these sites are not mapped for the entire *Xba*I-E region. The recombinant plasmid designations for various subclones of the *Xba*I-E region and relevant map units of the region are also indicated in Fig. 1.

Mapping and size distribution of IE RNA. The IE polysome-associated poly(A) RNA synthesized in the presence of 100 µM anisomycin was isolated and separated according to molecular weight in a denaturing methylmercury hydroxide gel and immobilized on DBM-paper as described above. Various recombinant plasmids of the *Xba*I-E region (Fig. 1) or various restriction enzyme DNA fragments from recombinant plasmids were ³²P labeled by nick translation (41) and used as probes to locate the regions of IE RNA transcription. The mRNA size classes were localized to the nearest restriction fragment or junction of two fragments. Figure 2 illustrates the size classes of IE RNA detected by the various sections of the *Xba*I-E region. The exposure period of the autoradiogram varied according to the relative concentration of the immobilized viral RNA.

Transcription from the *Bam*HI-D region (data

FIG. 1. Physical maps for the *Xba*I cleavage sites for human CMV (Towne) DNA and for the *Bam*HI, *Sal*I, *Sac*I, and *Pst*I cleavage sites for *Xba*I-E DNA. The CMV genome (240 kb) consists of 197 kb of long unique (U_L) sequences and 42 kb of short unique (U_S) sequences. At each end of the long unique sequences, there are repeat sequences of about 11 kb that are inverted relative to each other. Likewise, at each end of the short unique sequences, there are repeat sequences of about 2.0 to 2.5 kb that are inverted relative to each other. Such a structure results in four equimolar populations of the viral DNA differing in the relative orientations of the long and short segments. The *Bam*HI letters in parentheses represent fragment designations of the entire *Bam*HI physical map (52) determined by La Femina and Hayward (submitted for publication). Regions of *Xba*I-E subcloned into various recombinant plasmids are designated. Cleavage sites for *Ava*I, *Sma*I, and *Hinc*II are designated only for defined regions within the *Xba*I-E DNA. The data for the *Xba*I physical map of the entire viral genome are from La Femina and Hayward (27; submitted for publication).

not shown), the *Bam*HI-C region (Fig. 2, lane 1) and the *Pst*I-C region (data not shown) of *Xba*I-E was not detected. Three separate regions of IE RNA transcription within the *Xba*I-E region were detected. The approximate limits of each region were defined as follows. A 1.95-kb IE RNA size class was detected with fragments of viral DNA extending from the *Bam*HI-E fragment into the *Pst*I-B region, but not extending beyond the *Ava*I site (Fig. 2, lanes 2, 3, and 4). This region encoded for a single size class of IE RNA that was present in the infected cell at relatively low concentrations requiring 16 h for detection by autoradiography. This region was designated as IE coding region 3 (Fig. 2) and is located between 0.709 to 0.728 map units (Fig. 1). Finer restriction enzyme maps are necessary to narrow the limits for coding region 3.

The region from the *Ava*I site in *Pst*I-B to the left end of the *Sal*I-A region was either not transcribed or partially transcribed (Fig. 2, lane 5). It is possible that this region contains a small amount of leader sequence for DNA coding region 2 described below.

The region from the left end of the *Sal*I-A fragment to the left end of the *Pst*I-E fragment detected five IE RNAs ranging from 2.25 to 1.10

kb (Fig. 2, lane 6). Two of the RNAs, 1.75 and 1.40 kb, were slightly more abundant than the rest. The RNAs from this coding region were also present in the infected cell at relatively low concentrations requiring 16 h for detection by autoradiography. The same RNA size classes were detected with 32 P-labeled *Bam*HI-B DNA, except that with this probe a 1.95-kb size class of RNA was highly abundant (Fig. 2, lane 7). A probe extending from the left end of *Pst*I-E to the *Sma*I site in the *Pst*I-E region also detected a 1.95-kb RNA, but not in high abundance (Fig. 2, lane 8). Therefore, the region from the left end of the *Sal*I-A fragment to the *Sma*I site in the *Pst*I-E region (0.732 to 0.739 map units) was designated coding region 2 (Fig. 2).

The predominant IE RNA has a size class of 1.95 kb and was easily detected after only 2 h of autoradiography (Fig. 2). The 32 P-labeled probe from the *Sma*I site in *Pst*I-E to the right end of *Pst*I-E (Fig. 2, lane 9) as well as the entire *Pst*I-E DNA (Fig. 2, lane 10) and *Pst*I-F DNA (Fig. 2, lane 11) detected the predominant 1.95-kb RNA. Although the same size class of RNA was detected with *Pst*I-G, the relative amount of hybridization was lower (Fig. 2, lane 12). In addition, the predominant IE RNA was not easily

FIG. 2. Autoradiograph of Northern blot hybridizations of IE polysome-associated poly(A) RNA. Virus-specific RNA was detected by hybridization of ^{32}P -labeled viral DNA prepared as described in the text. The sizes of the viral RNAs are indicated in kb. The autoradiograms were incubated for various periods of time as indicated. The following recombinant plasmids or DNA sections were used as ^{32}P -labeled probes in the indicated lanes: 1, *Bam*HI-C; 2, *Sal*I-C; 3, left end of *Sal*I-B to right end of *Pst*I-A; 4, left end of *Pst*I-B to right end of *Sal*I-B; 5, *Ava*I site in *Pst*I-B to right end of *Sal*I-B; 6, left end of *Sal*I-A to right end of *Pst*I-B; 7, *Bam*HI-B; 8, left end of *Pst*I-E to *Sma*I site; 9, *Sma*I site in *Pst*I-E to right end of *Pst*I-E; 10, *Pst*I-E; 11, *Pst*I-F; 12, *Pst*I-G; 13, left end of *Pst*I-D to left end of *Sac*I-E; 14, *Sac*I-E; 15, right end of *Sac*I-E to right end of *Pst*I-D; 16, *Pst*I-D. The various DNA-coding regions are designated as 1, 2, and 3.

detected by a ^{32}P -labeled probe from the right end of *Pst*I-G to the right end of *Sac*I-D (Fig. 2, lane 13). Because these probes are relatively small, have intron sequences, and are relatively rich in adenine and thymine (Stenberg et al., unpublished data), they may have hybridized less efficiently with the predominant IE RNA. A ^{32}P -labeled *Sac*I-E DNA probe (Fig. 2, lane 14) detected the predominant IE RNA within a 2-h incubation period for autoradiography. Likewise, the ^{32}P -labeled probe representing the *Pst*I-D fragment detected the predominant IE RNA (Fig. 2, lane 16). However, the region extending from the right end of *Sac*I-E to the right end of *Pst*I-D was not transcribed under IE conditions (Fig. 2, lane 15). Therefore, the re-

gion extending from the *Sma*I site in *Pst*I-E to the right end of *Sac*I-E (0.739 to 0.751 map units) was designated as coding region 1 (Fig. 2). The map units of the various IE regions are based on determining the size of the various DNA fragments within the *Xba*I-E region by using a wide range of DNA size standards.

Electron microscopy of R-loops with the *Xba*I-E region. Electron microscopy of R-loops was utilized primarily to visualize the three coding regions within *Xba*I-E and secondarily to qualitatively estimate coding distances. IE polysome-associated, poly(A) RNA was selected by hybridization with *Xba*I-E DNA linked to DBM-cellulose. After elution of the RNA, the RNA was ethanol precipitated and suspended in buffer

FIG. 3. Electron micrograph of the *Xba*I-E DNA of CMV bearing R-loops formed with IE RNA. RNA complementary to *Xba*I-E DNA was selected by hybridization to *Xba*I-E DNA linked to DBM-cellulose. The selected IE RNA was hybridized to 1 μ g of *Xba*I-digested, recombinant plasmid *Xba*I-E DNA as described in the text. Size markers were double-stranded (ds) vector plasmid pACYC 184 and single-stranded (ss) phage M2 DNAs. The interpretative drawing illustrates the R-loops formed which are designated as coding regions 1, 2, and 3. The hybrids are represented diagrammatically by the thicker line, and the displaced single-stranded DNA strands are represented diagrammatically by the thinner line. The orientation for the molecule was facilitated by the large 3.0-kb R-loop which was positioned to the left. Magnification, $\times 12,500$.

for the formation of R-loops with *Xba*I-E DNA as described above. Figure 3 shows three R-loops within the *Xba*I-E DNA that are designated coding regions 1, 2, and 3. From the left-hand *Xba*I site to coding region 3, there was approximately 4.6 kb of DNA not transcribed under IE conditions. The R-loop associated with coding region 3 extended for approximately 3.0 kb. To the right of coding region 3, there was a region of

approximately 2.8 kb that was not transcribed under IE conditions. Coding region 2 extended for approximately 1.7 kb in a direction opposite that of coding region 1. The poly(A) tails of the mRNAs of coding regions 1 and 2 were estimated to be approximately 230 base pairs apart (Fig. 3). Coding region 1 extended for approximately 2.8 kb, but the exact measurement was frequently complicated by loops found to be near the 5' end

of the mRNA. Finally, there was again a relatively large region of approximately 4.7 kb to the right of region 1 that was not transcribed under IE conditions. The measurements presented above represent approximations subject to the limitations of electron microscopy and are based on double-stranded and single-stranded DNA size standards described above. The extent of the three encoding regions and the lack of IE transcription at the left as well as the right ends of the *Xba*I-E region were in good agreement with the Northern blot analysis presented in Fig. 2.

Direction of transcription by Southern blot analysis. Electron microscopy of R-loops suggested that coding regions 1 and 2 were transcribed in opposite directions. Southern blot hybridization was used to map the IE RNAs and their 3' ends. Virus-specific IE RNA was selected by hybridization to *Xba*I-E DNA linked to DBM-cellulose. 32 P-labeled cDNAs to either the entire RNA sequence (total cDNA) or to the 3'

ends of the RNA sequence (3' cDNA) were prepared as described above. Selected sub-clones of the *Xba*I-E region containing the *Bam*HI-A (pMW34) or -B (pCB42) fragments were digested with the restriction enzymes *Bam*HI and *Pst*I or *Bam*HI, *Pst*I, and *Sal*I, respectively (Fig. 1). After immobilization of DNA fragments of 2.0 kb or less by the method of Southern (43), probes of 32 P-labeled total cDNA or 3' cDNA as well as 32 P-labeled *Xba*I-E DNA were used for hybridization as described above.

The 32 P-labeled recombinant plasmid *Xba*I-E probe hybridized to all DNA bands, including vector plasmid DNA, with equal intensity (Fig. 4A). The 32 P-labeled total cDNA probe hybridized intensely to a 1.25-kb DNA (Fig. 4B, lane 1) representing region 2 from the left end of *Sal*I-A to the left end of *Pst*I-E (Fig. 1) and to a 0.75-kb DNA (Fig. 4B, lane 2) representing region 1 from the left end of *Bam*HI-A to the right end of *Pst*I-E (Fig. 1). In contrast, the relative amount of hybridization with 32 P-labeled total cDNA probe was less to a 0.70-kb DNA (Fig. 4B, lane 1) extending from the left end of *Pst*I-E to the left end of *Bam*HI-A (Fig. 1), which is the region of DNA where region 1 encounters region 2 (Fig. 2). The total cDNA probe also hybridized to a 0.60-kb DNA that is the *Pst*I-F fragment of encoding region 1 (Fig. 4B, lane 2).

When the 32 P-labeled 3' cDNA probe was used for hybridization, the relative amount of hybridization was low to the 1.5-kb and the 0.75-kb DNAs representing coding regions 1 and 2, respectively (Fig. 4C, lanes 1 and 2). This hybridization is presumably due to a small amount of 32 P-labeled cDNA synthesized from RNA template that was only partially degraded. In contrast, the relative amount of hybridization was high to the 0.70-kb DNA that is the region of DNA where coding region 1 encounters coding region 2 (Fig. 4C, lane 1). Therefore, the 3' ends of the IE RNAs in coding regions 1 and 2 map near or within the 0.70-kb DNA that extends from the left end of *Pst*I-E DNA to the left end of *Bam*HI-A DNA (Fig. 1). These results also confirm in part the location of coding regions 1 and 2 identified by Northern blot hybridization analysis. The transcription of coding DNA sequence 1 proceeds from right to left, whereas transcription of encoding DNA sequence 2 is on the complementary strand and in the opposite direction. The 3' ends of these viral RNAs are located approximately left and right of 0.739 map units. The exact location of the 5' and 3' ends of the above genes is currently being determined by DNA sequencing.

The physical maps in coding region 3 were not defined in sufficient detail to allow unambiguous orientation of this mRNA. Since this viral RNA

FIG. 4. Southern blot hybridization of 32 P-labeled total or 3' cDNA made to selected IE RNA. IE polysome-associated, poly(A) RNA was selected by hybridization to *Xba*I-E DNA linked to DBM-cellulose. 32 P-labeled total or 3' cDNA probes to the selected IE RNA were synthesized with reverse transcriptase as described in the text. 32 P-labeled recombinant plasmid *Xba*I-E was prepared by nick translation (41). Recombinant plasmid pCB42 DNA (lane 1) was digested with the restriction enzymes *Bam*HI, *Pst*I, and *Sal*I. Recombinant plasmid pMW34 (lane 2) was digested with the restriction enzymes *Bam*HI and *Pst*I. Samples of 5×10^5 cpm of 32 P-labeled *Xba*I-E DNA (A), total cDNA (B), or 3' cDNA (C) were hybridized for 48 h at 42°C as described in the text. The sizes of the viral DNA fragments are indicated in kilobases. The locations of the various restriction enzyme sites and their relationship to DNA coding regions 1 and 2 are indicated.

and its protein product are synthesized in relatively low amounts, the orientation of this mRNA was not analyzed further.

In vitro translation products of IE RNAs. The *Xba*I-N and -E regions of the viral genome are located between 0.660 to 0.770 map units and code approximately 88% of the IE whole cell RNA as well as for the abundant size classes of polysome-associated, poly(A) RNA (52). To determine the translation products of these IE RNAs, IE polysome-associated, poly(A) RNA was preparatively hybridized to *Xba*I-N or -E DNA linked to DBM-paper and subsequently eluted. The virus-specified RNAs were then tested for translation in a rabbit reticulocyte lysate as described above. RNA isolated when using the *Xba*I-N DNA did not translate in the reticulocyte lysate. RNAs originating from this region of the genome at various times after infection are currently being investigated. Preparations of IE RNA selected by hybridization to *Xba*I-E were translated. The virus-specified RNA eluted from *Xba*I-E DNA was hybridized to *Xba*I-E DNA a second time, eluted, and translated in vitro. Four polypeptides having apparent molecular weights of 75,000, 56,000, 39,000, and 16,500 were identified by SDS-polyacrylamide gel electrophoresis (Fig. 5, lane 2). Some polypeptides between 39,000 and 16,000 daltons were present at relatively low concentrations. The polypeptide of 75,000 daltons was the predominant translation product.

To determine the location on the genome of the various proteins coded for by IE RNA, three sections of the *Xba*I-E DNA representing coding regions 1, 2 plus 3, or 3 were used. Since the Northern blot hybridization (Fig. 2) and in vitro transcription with Manley extracts (Thomsen et al., unpublished data) indicated that the predominant IE RNA for coding region 1 was initiated in the *Pst*I-D region (Fig. 1), this DNA as well as the *Pst*I-F, -G, and -D DNAs (Fig. 1) were immobilized on DBM-paper for mRNA hybridization. In addition, the *Sac*I-A DNA and the *Sal*I-C DNA (Fig. 1) were immobilized for hybridization of viral RNAs originating from coding regions 2 and 3.

Since the Northern blot analysis indicated that coding region 1 was highly transcribed under IE conditions, whereas transcription from coding regions 2 and 3 was relatively low, the amounts of IE poly(A) RNA used for hybridization were 20, 80, and 100 μ g for coding regions 1, 2 plus 3, and 3, respectively. The IE RNA was eluted and translated in vitro as described above.

The IE RNA selected by coding region 1, the *Pst*I-D DNA, or the *Pst*I-F, -G, and -D DNA (data not shown) was translated into a predominant polypeptide of 75,000 daltons and a polypeptide at lower relative concentrations of

39,000 daltons (Fig. 5, lane 5). Several additional polypeptides of lower molecular weight were present that were not detected in the infected cell; consequently, these polypeptides may not be virus specified. The nonspecific polypeptides above 42,000 daltons are endogenous to the rabbit reticulocyte lysate, and some of these are shown in Fig. 5, lane 3. The polypeptides below 39,000 daltons may be premature termination products or endogenous products that are normally not detected unless the reticulocyte lysate is highly stimulated. Alternatively, these products may reflect an unusual property of the viral mRNA that causes early termination. Polypeptides of similar apparent molecular weight were not found in infected cells (Fig. 6).

The IE RNA selected by hybridization to the *Sac*I-A DNA was translated to polypeptides of 56,000, 42,000, 21,000, and 16,500 daltons (Fig. 5, lane 6) plus several additional polypeptides. Infected cell-specific polypeptides of 56,000, 42,000, 21,000, and 16,500 daltons have been detected in infected cells (Fig. 6) under IE conditions, but at relatively low levels (6, 7, 45, 46, 53). In addition, a polypeptide of 75,000 daltons was translated (Fig. 5, lane 6). Our current measurements of the predominant IE RNA from coding region 1 and our DNA sequence data indicate that the majority of transcripts are terminated before the *Sac*I-A site

FIG. 5. Fluorograms of the in vitro translation products of virus-specified IE mRNA. Lanes: 1, in vitro translation with adenovirus type 2 cytoplasmic RNA; 2, IE RNA selected twice by hybridization to *Xba*I-E DNA; 3, no added RNA; 4, adenovirus type 2 cytoplasmic RNA; 5, IE RNA selected by hybridization to *Pst*I-D DNA; 6, IE RNA selected by hybridization to *Sac*I-A DNA; 7, IE RNA selected by hybridization to *Sal*I-C DNA. The apparent molecular weights of the virus-specified polypeptides coded for by IE RNA are designated.

(Stenberg et al., unpublished data). The considerably lower amount of translation of a 75,000-dalton polypeptide by RNA from *SacI*-A DNA even though fourfold more RNA was used for hybridization may be due to (i) IE RNA from

coding region 2, (ii) occasional transcription across the *SacI*-A site from coding region 1, or (iii) contamination by the abundant IE RNA from region 1. The IE RNA from region 1 is 20- to 30-fold more abundant than the IE RNA from regions 2 and 3 (Thomsen et al., unpublished data); consequently, it is difficult to remove all of this viral RNA from the DNA-DBM filter. In addition, the predominant IE RNA translates extremely well. We are able to translate as little as 10 to 15 ng of this mRNA; consequently, the slightest amount of contamination would be detected in the rabbit reticulocyte lysate. There is currently no evidence to suggest that there are similar sequences in coding regions 1 and 2. Therefore, we favor the latter interpretation for the presence of this translation product when using *SacI*-A DNA for mRNA selection. The IE RNA selected by hybridization to DNA representing coding region 3, the *SalI*-C DNA (Fig. 1), coded for a very minor polypeptide of approximately 68,000 daltons (Fig. 5, lane 7).

Immunoprecipitation of the predominant IE protein synthesized in vitro or in vivo. A monoclonal antibody (E-3) against the 72,000-dalton IE protein (19) was used to immunoprecipitate IE protein synthesized in vitro or in vivo. The controls were normal mouse serum plus lysate from in vitro translation or infected cells. Addi-

FIG. 6. Fluorograms of IE polypeptides synthesized in vitro or in vivo and immunoprecipitated by monoclonal antibody. (A) Lanes: 1, polypeptides from in vitro translation with adenovirus type 2 cytoplasmic RNA; 2, in vitro-translated adenovirus polypeptides plus monoclonal antibody E-3; 3, in vitro-translated polypeptides of IE RNA selected by hybridization to *PstI*-D DNA; 4 and 5, in vitro-translated polypeptides of IE RNA selected by hybridization to *PstI*-D DNA plus normal mouse serum (lane 4) or monoclonal antibody E-3 (lane 5); 6, infected cell-specific polypeptides pulse-labeled with [35 S]methionine in the presence of actinomycin D after treatment of the infected cells for 12 h with cycloheximide; 7, infected cell-specific polypeptides pulse-labeled with [35 S]methionine from 1 to 4 h postinfection in the absence of drug treatment; 8, uninfected cell polypeptides. (B) Lanes: 1, polypeptides from in vitro translation with adenovirus type 2 cytoplasmic RNA; 2, uninfected cell polypeptides; 3, uninfected cell polypeptides plus monoclonal antibody E-3; 4, infected cell-specific polypeptides pulse-labeled with [35 S]methionine in the presence of actinomycin D after treatment of the infected cells for 12 h with cycloheximide; 5 and 6, infected cell-specific polypeptides pulse-labeled and treated as described above plus normal mouse serum (lane 5) or monoclonal antibody E-3 (lane 6); 7, in vitro-translated polypeptides of IE RNA selected by hybridization to *PstI*-D DNA plus monoclonal antibody E-3. The apparent molecular weights of the infected cell-specific IE polypeptides and the immunoprecipitated polypeptides are designated.

tional controls were specific monoclonal antibody plus lysate from in vitro translation of adenovirus mRNA or uninfected cells.

The IE mRNA selected by hybridization to *Pst*I-D DNA representing coding region 1 was translated in vitro and analyzed by SDS-polyacrylamide gel electrophoresis as described above. Figure 6A (lane 3) shows the typical polypeptide profile with the predominant IE protein of 75,000 daltons, the minor polypeptide of 39,000 daltons, as well as the endogenous and low-molecular-weight proteins described above (Fig. 5, lane 5). Monoclonal antibody E-3 immunoprecipitated the 75,000-dalton IE polypeptide, but not the 39,000-dalton polypeptide (Fig. 6A, lane 5). Normal mouse serum caused no immunoprecipitation (Fig. 6A, lane 4).

The immunoprecipitated IE 75,000-dalton polypeptide comigrates with some of the IE polypeptides found in the infected cell and synthesized under IE conditions. However, the proteins synthesized in vivo have an apparent molecular weight range of 75,000 to 68,000 as previously reported (46). The majority of the IE polypeptides had an apparent molecular weight of 72,000 (Fig. 6A, lane 6). A similar infected cell-specific polypeptide of 72,000 daltons can be detected in infected cells without resorting to a cycloheximide block for mRNA accumulation and pulse-labeling in the presence of actinomycin D (Fig. 6A, lane 7).

Since the predominant IE protein is 75,000 daltons when synthesized in vitro, but is found in vivo as a broad banding protein of 75,000 to 68,000 daltons, monoclonal antibody E-3 was used to immunoprecipitate IE protein synthesized in vitro or in vivo, and the precipitates were analyzed by SDS-polyacrylamide gel electrophoresis. Monoclonal antibody E-3 did not precipitate any detectable proteins from uninfected cells (Fig. 6B, lane 3). In contrast, the monoclonal antibody precipitated the predominant IE protein from lysates of cells pulse-labeled under IE conditions. The apparent molecular weight of the protein ranged from 75,000 to 68,000, with the majority of the protein at 72,000 (Fig. 6B, lane 6). The majority of the IE protein synthesized in vivo migrated slightly faster than the predominant IE protein of 75,000 daltons synthesized in vitro (Fig. 6B, compare lanes 6 and 7). The IE mRNA that coded for the IE protein of 75,000 daltons originated from coding region 1. Therefore, we propose that the IE protein of primarily 72,000 daltons is also coded by region 1. The cause of the slight shift in migration when the predominant IE protein is synthesized in vivo has not been determined.

Summary of the IE RNAs and proteins coded within the *Xba*I-E region (0.685 to 0.770 map units). The map locations for the various coding

regions in *Xba*I-E, the direction of transcription, the size classes of IE RNAs, and the in vitro translation products are summarized in Fig. 7. IE gene expression of human CMV is dominated by transcription of coding region 1. The direction of transcription, where known, is indicated by arrows pointing toward the 3' end in the prototype arrangement of the viral genome. Region 1 codes for an abundant mRNA of 1.95 kb that is translated in vitro to a polypeptide of 75,000 daltons. This protein is presumably modified in vivo to a protein of primarily 72,000 daltons. An infected cell-specific polypeptide of the same apparent molecular weight is also detectable in the cell within 4 h postinfection without manipulating the cell to accumulate IE mRNA by inhibiting de novo protein synthesis. The size of the minor polypeptides coded in regions 1, 2, and 3 are also shown. We propose that coding region 1 represents the predominant region of viral gene expression immediately after infection.

DISCUSSION

One or more of the viral proteins coded by the IE mRNAs of herpesviruses are necessary for efficient transcription of the other viral genes. This has been documented for herpesviruses by using wild-type virus in cells treated with cycloheximide, an inhibitor of protein synthesis (2, 9, 11-16, 21, 23, 28-30, 39, 48, 52, 53), or with interferon (47). In addition, mutants of herpes simplex virus that are temperature sensitive in an IE gene produce a similar restricted transcriptional pattern at the nonpermissive temperature (36, 38, 54).

Five major IE mRNA species of herpes simplex virus originate primarily, but not exclusively, from the inversely repeated sequences of both the long and short components of the viral genome (10, 11, 14, 22, 56). These mRNA species accumulate in the presence of protein synthesis inhibitors at relatively similar concentrations. In contrast, in the human CMV-infected cell, IE mRNA originates from a region (0.660 to 0.770 map units) in the large unique component of the viral genome. The region (0.739 to 0.751 map units) designated as DNA coding region 1 is highly transcribed relative to regions 2 (0.732 to 0.739 map units) and 3 (0.709 to 0.728 map units). The region of 0.660 to 0.680 map units (*Xba*I-N) requires further investigation. It is assumed that RNA polymerase II recognizes the promoters for IE RNA synthesis. We propose that the upstream regulatory sequence of DNA coding region 1 competes more efficiently for RNA polymerase II and that this constitutes the first step in the regulation of human CMV gene expression.

In the herpes simplex virus-infected cell, at

FIG. 7. Summary of the IE RNAs and proteins coded within the *XbaI*-E DNA region. The map units of coding regions 1, 2, and 3 depict the limits of the probes used to detect viral RNAs. The direction of transcription is indicated for coding regions 1 and 2. The thickness of the bar represents the relative abundance of the IE RNAs originating from the various coding regions, estimated by the relative amount of hybridization and the incubation time of the Northern blot autoradiograms. The size classes of the viral RNAs are indicated in kilobases. The in vitro-translated polypeptides from IE RNA selected by the various coding regions are designated by apparent molecular weights ($\times 10^3$). The abundant IE proteins synthesized in vitro or in vivo and immunoprecipitated by monoclonal antibody E-3 are designated by apparent molecular weight ($\times 10^3$).

least one IE protein has been implicated as a regulatory protein necessary for the efficient transcription of early as well as late viral RNA (38, 55). Since the 72,000-dalton protein of CMV is the predominant IE protein, we hypothesize that this protein plays the major role in influencing transcription of the other viral genes and, consequently, that the protein presumably has an influence in determining whether the infection is latent, persistent, or productive. Immunoprecipitation analysis of this protein synthesized in vitro or in vivo suggests that the predominant IE protein is synthesized as a 75,000-dalton polypeptide that is modified in vivo in a way that increases migration in a denaturing gel. Gibson (18) has suggested that the defused nature of the predominant IE protein is suggestive of structural or conformational heterogeneity. This could be due to inter- or intramolecular interactions (e.g., disulfide bonding) or rapid post-translational modification (e.g., phosphorylation). The predominant IE protein of CMV (Towne) is phosphorylated (18). The predominant IE gene does not appear to be located in an active genetic element resulting in different locations or orientations within the viral chromosome. In addition, the majority of the viral RNA from the predominant IE gene follows a repeatable splicing pattern (Stenberg et al., unpublished data). Since the translation of the predominant IE mRNA in vitro renders a

protein that migrates more homogeneously, we propose that the broad migrating protein seen in vivo is due to post-translational modifications.

IE antigens of CMV accumulate in the nucleus of infected cells (33, 49), but if the cells are treated with cycloheximide the antigens also accumulate in the cytoplasm (Landini and Stinski, unpublished data). In the nucleus, the predominant IE protein is associated preferentially with chromatin (S. Michelson, personal communication). At very early stages of infection, CMV induces a protein that is responsible for the stimulation of chromatin template activity as measured by the incorporation of [3 H]UMP in the presence of *Escherichia coli* RNA polymerase (24). Therefore, we proposed that the predominant IE protein of CMV is a regulatory protein that influences transcription. This protein may serve obligatory functions that are required throughout the replication cycle of the virus.

The mRNAs and proteins from coding regions 2 and 3 were detected under the conditions defined as IE. However, it is possible that these are early genes that function after the synthesis of the predominant IE protein. There is a switch from restricted to extensive transcription after synthesis of the predominant IE protein. The region of the viral genome most affected is the large repeat sequence and adjacent sequences (12, 52). Since viral RNA appears on polyribo-

somes as functional mRNA from the above regions at early times or in the presence of phosphonoacetic acid, an inhibitor of viral DNA synthesis, we have proposed that this region codes for early viral proteins (52, 53). The data suggest that the predominant IE protein of CMV regulates transcription from the above region as well as other regions of the viral genome. The predominant IE protein may enhance promoter recognition by interacting with the viral chromosome or RNA polymerase II. Even though the predominant IE protein may serve to enhance early or late gene transcription, additional regulatory events are presumably necessary for transport of the viral RNA to the cytoplasm. Most of the viral RNA originating from the large and small unique sequences is retained in the nucleus at early times and not transported to the cytoplasm until after viral DNA replication (52). The above regulatory events presumably play an important role in the protracted replicative cycle of human CMV and in the species restrictions on replication.

ACKNOWLEDGMENTS

This investigation was supported by Public Health Service grant 2R01AI13526 from the National Institute of Allergy and Infectious Diseases and by grant 1-697 from the National Foundation March of Dimes. M.F.S. is the recipient of Public Health Service career development award 1K04AI00373 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

- Anderson, K. P., R. H. Costa, L. E. Holland, and E. K. Wagner. 1979. Isolation and localization of herpes simplex type 1 mRNA. *J. Virol.* 30:805-820.
- Anderson, K. P., R. H. Costa, L. E. Holland, and E. K. Wagner. 1980. Characterization of herpes simplex virus type 1 RNA present in the absence of de novo protein synthesis. *J. Virol.* 34:9-27.
- Bachi, B., and W. Arber. 1977. Physical mapping of BglII, BamHI, EcoRI, HindIII, and PstI restriction fragments of bacteriophage P1 DNA. *Mol. Gen. Genet.* 153:311-324.
- Bailey, J. M., and N. Davidson. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. *Anal. Biochem.* 70:75-85.
- Bishop, D. H. L., J. R. Claybrook, and S. Spiegelman. 1967. Electrophoretic separation of viral nucleic acid on polyacrylamide gels. *J. Mol. Biol.* 26:373-387.
- Blanton, R. A., and M. J. Tevethia. 1981. Immunoprecipitation of virus-specific immediate-early and early polypeptides from cells lytically infected with human cytomegalovirus strain AD169. *Virology* 112:262-273.
- Cameron, J. M., and C. M. Preston. 1981. Comparison of the immediate early polypeptides of human cytomegalovirus isolates. *J. Gen. Virol.* 54:421-424.
- Chamberlain, J. P. 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. *Anal. Biochem.* 98:132-135.
- Chua, C. C., T. H. Carter, and S. St. Jeor. 1981. Transcription of the human cytomegalovirus genome in productively infected cells. *J. Gen. Virol.* 56:1-11.
- Clements, J. B., J. McLauchlan, and D. J. McGeoch. 1979. Orientation of herpes simplex virus type 1 immediate early mRNA's. *Nucleic Acids Res.* 7:77-91.
- Clements, J. B., R. J. Watson, and N. M. Wilkie. 1977. Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. *Cell* 12:483-496.
- De Marchi, J. M. 1981. Human cytomegalovirus DNA: restriction enzyme cleavage maps and map locations for immediate-early, early, and late RNAs. *Virology* 114:23-38.
- De Marchi, J. M., C. A. Schmidt, and A. S. Kaplan. 1980. Patterns of transcription of human cytomegalovirus in permissively infected cells. *J. Virol.* 35:277-286.
- Easton, A. J., and J. B. Clements. 1980. Temporal regulation of herpes simplex virus type 2 transcription and characterization of virus immediate early mRNA's. *Nucleic Acid Res.* 8:2627-2645.
- Feldman, L., F. Rixon, J. H. Jean, T. Ben-Porat, and A. S. Kaplan. 1979. Transcription of the genome of pseudorabies virus (a herpesvirus) is stringently controlled. *Virology* 97:316-327.
- Feldman, L. T., J. M. DeMarchi, T. Ben-Porat, and A. S. Kaplan. 1982. Control of abundance of immediate-early mRNA in herpesvirus (pseudorabies)-infected cells. *Virology* 116:250-263.
- Furukawa, T., A. Floretti, and S. Plotkin. 1973. Growth characteristics of cytomegalovirus in human fibroblast with demonstration of protein synthesis early in viral replication. *J. Virol.* 11:991-997.
- Gibson, W. 1981. Immediate-early protein of human cytomegalovirus strains AD169, Davis, and Towne differ in electrophoretic mobility. *Virology* 112:350-354.
- Goldstein, L. C., J. McDougall, R. Hackman, J. D. Meyers, E. D. Thomas, and R. C. Nowinski. 1982. Monoclonal antibodies to cytomegalovirus: rapid identification of clinical isolates and preliminary use in diagnosis of CMV pneumonia. *Infect. Immun.* 38:273-281.
- Ho, M. 1982. Cytomegalovirus biology and infection, p. 131-212. *In* W. B. Greenough and T. C. Merigan (ed.), *Current topics in infectious disease*. Plenum Publishing Corp., New York.
- Jean, J. H., T. Ben-Porat, and A. S. Kaplan. 1974. Early functions of the genome of herpesvirus. III. Inhibition of the transcription of the viral genome in cells treated with cycloheximide early during the infective process. *Virology* 59:516-523.
- Jones, P. C., G. S. Hayward, and B. Roizman. 1977. Anatomy of herpes simplex virus DNA. VII. RNA is homologous to noncontiguous sites in both the L and S components of viral DNA. *J. Virol.* 21:268-276.
- Jones, P. C., and B. Roizman. 1979. Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. *J. Virol.* 31:299-314.
- Kamata, T., S. Tanaka, and Y. Watanabe. 1978. Human cytomegalovirus induced chromatin factors responsible for changes in template activity and structure of infected cell chromatin. *Virology* 90:197-208.
- Kessler, S. W. 1975. Rapid isolation of antigens from cells with a staphylococcal protein A-antibody adsorbent: parameters of the interaction of antibody-antigen complexes with protein A. *J. Immunol.* 115:1617-1624.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
- La Femina, R. L., and G. S. Hayward. 1980. Structural organization of the DNA molecules from human cytomegalovirus, p. 39-55. *In* B. N. Fieldo and R. Jaenish (ed.), *Animal virus genetics*. Academic Press, Inc., New York.
- Leung, W. C., K. Dimock, J. R. Smiley, and S. Bacchetti. 1980. Herpes simplex virus thymidine kinase transcripts are absent from both nucleus and cytoplasm during infection in the presence of cycloheximide. *J. Virol.* 36:361-365.
- Mackem, S., and B. Roizman. 1980. Regulation of herpesvirus macromolecular synthesis: transcription-initiation sites and domains of α genes. *Proc. Natl. Acad. Sci. U.S.A.* 77:7122-7126.
- Macken, S., and B. Roizman. 1981. Regulation of herpes-

- virus macromolecular synthesis: temporal order of transcription of α genes is not dependent on the stringency of inhibition of protein synthesis. *J. Virol.* 40:319-322.
31. McGrogan, M., D. J. Spector, C. J. Goldenberg, D. Halbert, and H. J. Raskas. 1979. Purification of specific adenovirus 2 RNAs by preparative hybridization and selective thermal elution. *Nucleic Acids Res.* 6:593-606.
 32. McLauchlan, J., and J. B. Clements. 1982. A 3' co-terminus of two early herpes simplex virus type 1 mRNAs. *Nucleic Acids Res.* 10:501-512.
 33. Michelson-Fiske, S., F. Horodniceanu, and J. C. Guillon. 1977. Immediate early antigens in human cytomegalovirus-infected cells. *Nature (London)* 270:615-617.
 34. Noyes, B. E., and G. R. Stark. 1975. Nucleic acid hybridization using DNA covalently coupled to cellulose. *Cell* 5:301-310.
 35. Palmiter, R. D. 1974. Mg^{++} precipitation of ribonucleo-protein complexes. Expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. *Biochemistry* 13:3606-3614.
 36. Parris, D. S., R. J. Courtney, and P. A. Schaffer. 1978. Temperature sensitive mutants of herpes simplex virus type 1 defective in transcriptional and post transcription functions required for viral DNA synthesis. *Virology* 90:177-186.
 37. Pelham, H. R. B., and R. T. Jackson. 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.* 67:247-256.
 38. Preston, C. M. 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant *tsK*. *J. Virol.* 29:275-284.
 39. Rokusanova, T., T. Ben-Porat, M. Mimeno, and A. S. Kaplan. 1971. Early functions of the genome of herpesvirus. I. Characterization of the RNA synthesized in cycloheximide-treated infected cells. *Virology* 46:877-889.
 40. Rapp, F. 1980. Persistence and transmission of cytomegalovirus, p. 193-232. In H. Fraenkel-Conrat and R. R. Wagner (ed.), *Virus-host interactions. Viral invasion, persistence, and disease*, vol. 16. Plenum Publishing Corp., New York.
 41. Rigby, P. W. J., M. Dieckman, C. Rhoades, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
 42. Smith, H. O., and M. L. Birnstiel. 1976. A simple method for DNA restriction site mapping. *Nucleic Acids Res.* 3:2387-2399.
 43. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
 44. Stark, G. R., and J. G. Williams. 1979. Quantitative analysis of specific labelled RNAs using DNA covalently linked to diazobenzyloxymethyl-paper. *Nucleic Acids Res.* 6:195-203.
 45. Stinski, M. F. 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. *J. Virol.* 23:751-767.
 46. 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.
 47. Stinski, M. F., D. R. Thomsen, and J. E. Rodriguez. 1982. Synthesis of human cytomegalovirus-specified RNA and protein in interferon-treated cells at early times after infection. *J. Gen. Virol.* 60:261-270.
 48. Stinski, M. F., D. R. Thomsen, and M. W. Wathen. 1980. Structure and function of the cytomegalovirus genome, p. 72-84. In A. Nahmias, W. R. Dowdle, R. F. Schinazi (ed.), *Human herpesviruses*. Elsevier-North Holland Publishing Co., New York.
 49. Tanaka, S., M. Otsuka, S. Ihara, F. Maeda, and Y. Watanabe. 1979. Introduction of pre-early nuclear antigen(s) in HEL cells infected with human cytomegalovirus. *Microbiol. Immunol.* 23:263-271.
 50. Thomas, M., R. L. White, and R. W. Davis. 1976. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc. Natl. Acad. Sci. U.S.A.* 73:2294-2298.
 51. Thomsen, D. R., and M. F. Stinski. 1981. Cloning of the human cytomegalovirus genome as endonuclease XbaI fragments. *Gene* 16:207-216.
 52. 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.
 53. 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.* 38:446-459.
 54. Watson, R. J., and J. B. Clements. 1978. Characterization of transcription deficient temperature sensitive mutants of HSV-1. *Virology* 91:364-379.
 55. Watson, R. J. and J. B. Clements. 1980. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. *Nature (London)* 285:329-330.
 56. Watson, R. J., C. M. Preston, and J. B. Clements. 1979. Separation and characterization of herpes simplex virus type 1 immediate-early mRNA's. *J. Virol.* 31:42-52.
 57. Wentworth, B. B., and L. French. 1970. Plaque assay of cytomegalovirus strains of human origin. *Proc. Soc. Exp. Biol. Med.* 135:253-259.