Analysis of the Major Transcripts Encoded by the Long Repeat of Human Cytomegalovirus Strain AD169

SHERROL H. McDONOUGH, SILVIJA I. STAPRANS, AND DEBORAH H. SPECTOR*

Department of Biology, University of California, San Diego, La Jolla, California 92093

Received 22 August 1984/Accepted 30 November 1984

In this report, we describe the size and kinetics of appearance of RNAs from the long repeat of human cytomegalovirus. The most abundant RNA from this region was a 2.7-kilobase (kb) species that was detected throughout the infection and was most abundant at 27 and 72 h after infection. The 2.7-kb RNA was the only major species detected with a probe that included the terminus of the long repeat and the heterogeneous L-S junction region. Other transcripts were detected with probes from the internal portion of the long repeat, including an immediate-early RNA of 1.3 kb, early and late RNAs of 1.2 kb, and minor late transcripts of 4.4, 3.6, 3.3, and 1.8 kb. S1 nuclease and exonuclease VII protection analyses of RNA from immediate-early, early, midpoint, and late times in the infection indicated that the major 2.7-kb RNA was not spliced and that the RNA mapped within the long repeat, 1.6 kb from the heterogeneous region. No evidence for temporally regulated changes in transcription initiation, splicing, or choice of 3' end of this RNA was observed. Nuclease protection analysis also demonstrated that the second most abundant late RNA from this region, the 1.2-kb species, was not spliced and had the same polarity as the 2.7-kb RNA. The 1.2-kb RNA also mapped entirely within the long repeat, with its 3' terminus 1.7 kb upstream from the 5' terminus of the 2.7-kb RNA.

Human cytomegalovirus (HCMV) has been implicated in several medical disorders, which vary with the age and immunocompetence of the hosts. Few of the host or viral functions involved in productive, latent, or postulated oncogenic interactions of cytomegalovirus (CMV) with its host have been identified. The characterization of many specific viral functions is necessary before CMV-host interactions can be fully understood.

Studies on the molecular biology of HCMV have revealed that the genome consists of one long (L) and one short (S) segment, covalently attached in any of four possible orientations at a point called the L-S junction (2, 4, 5, 13, 17, 25). Both segments have terminal inverted repeats that for the AD169 strain vary in size by 0.2 to 0.5 kilobase pairs (kbp), resulting in heterogeneity in the virus population (19). Characterization of several cloned L-S junction fragments of the AD169 strain of HCMV has resulted in the identification of three classes of heterogeneity, all mapping at the L-S junction itself (18).

We have examined the transcription pattern of HCMV by dot blot analysis and pulse-labeling studies. These experiments localized the major immediate-early (IE) transcription site to the long unique segment and the major early transcription site to the repeated sequences of the HCMV genome. These repeated sequences continue to be actively transcribed at the midpoint and late in the infection (9). A similar transcription pattern has been described for the Towne and Davis strains of HCMV, except that the major early transcription site of the Davis strain is located in the long segment adjacent to the repeat (2, 3, 23).

We investigated the nature of this site of abundant transcription of the AD169 strain by Northern blot and nuclease protection analyses. The RNAs synthesized from this site were identified, including two major transcripts of 2.7 and 1.2 kilobases (kb). The map position of these RNAs and the kinetics of their accumulation were determined. The RNAs were found to be noncontiguous and to map completely

within the long repeat, outside the region of heterogeneity at the L-S junction. The 2.7-kb transcript could be detected throughout the infection and was most abundant at late times. The 1.2-kb RNA was detected in early, midpoint, and late RNAs. We conclude that the abundant transcripts synthesized from the long repeat do not extend to sequences at the L-S junction, which include the heterogeneous region.

MATERIALS AND METHODS

Virus and cells. Human cytomegalovirus strain AD169 and human fibroblasts were obtained and propagated as previously described (19).

Preparation of infected cell RNA. RNA was prepared 3 days after full cytopathic effect after an infection of 80% confluent human foreskin fibroblasts with 0.1 to 0.2 PFU per cell or 8, 27, 28, or 72 h after infection with 3 PFU per cell. Cytoplasmic RNA was isolated after lysis in an isotonic buffer containing Nonidet P-40 and vanadyl ribonucleoside complexes and was enriched for polyadenylated species by oligodeoxythymidylic acid-cellulose chromatography as previously described (9).

Preparation of molecularly cloned fragments. Cloned viral fragments were prepared, cleaved from the vector plasmid, and isolated by Seaplaque agarose gel electrophoresis and benzoyl-naphthol-DEAE (BND)-cellulose chromatography as described (19).

To obtain a recombinant clone of the *PvuII* O-WN fusion fragment, gel-purified *PvuII*-*EcoRI* subfragments of *EcoRI* bands O and WN and an equimolar amount of dephosphorylated, *PvuII*-cleaved pACYC184 were ligated together with T4 DNA ligase (Bethesda Research Laboratories) and used to transform *Escherichia coli* HB101 as previously described (19). Tetracycline-resistant, chloramphenicol-sensitive colonies were screened by restriction enzyme analysis.

Radioactive labeling of DNA. DNA fragments were 3' labeled with [α -³²P]deoxynucleotides and T4 DNA polymerase (P-L Biochemicals) as described by Maniatis et al. (8).

* Corresponding author.

To label 5' termini, DNA fragments were treated with calf intestinal phosphatase (Boehringer Mannheim Biochemicals) and then incubated with [γ - 32 P]dATP (ICN) and T4 polynucleotide kinase by the method of Maniatis et al. (8).

Nick translations were performed with high-specific-activity [α - 32 P]dCTP (3,000 Ci/mmol) by the method of Rigby et al. (15).

Nuclease protection analyses. S1 nuclease reactions were performed by using modifications of the method described by Berk and Sharp (1). Cloned DNA fragments (0.1 to 1.0 μ g) were ethanol precipitated with cytoplasmic polyadenylated RNA and yeast tRNA carrier. Samples were suspended in 10 to 20 μ l of 80% deionized formamide–0.04 M PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid) (pH 6.4)]–0.4 M NaCl–0.001 M EDTA (pH 8.0), denatured at 70°C for 10 min, and hybridized for 3 to 6 h at 60°C. The hybrids were diluted in 10 volumes of S1 buffer (0.28 M NaCl, 0.0045 M ZnSO₄, 0.03 M sodium acetate [pH 4.6], 5% glycerol, 20 μ g of denatured salmon DNA per ml) containing 1 U of S1 nuclease (Bethesda Research Laboratories) per μ l and incubated for 30 min at the temperatures indicated below. After the reaction was terminated by the addition of 0.5 M ammonium sulfate and 0.01 M EDTA, the nucleic acid was extracted with phenol and chloroform containing 4% isoamyl alcohol and was precipitated in isopropanol with yeast tRNA as carrier. In some cases, the hybridization reaction was treated with exonuclease VII in 0.01 M Tris-hydrochloride (pH 7.4)–0.03 M KCl–0.01 M EDTA–0.01 M β -mercaptoethanol at 37°C for 30 min and precipitated with ethanol.

Gel analysis. Neutral 1.0% agarose gels were cast and run in 0.04 M Tris-hydrochloride–0.02 M sodium acetate–0.018 M NaCl–0.002 M EDTA (pH 8.0). Alkaline 1.0, 1.5, or 2.0% agarose gels were cast in 0.04 M Tris-hydrochloride–0.005 M sodium acetate–0.001 M EDTA (pH 7.8) and equilibrated in alkaline running buffer (0.03 M NaOH, 0.001 M EDTA) for at least 1 h before electrophoresis. Samples were suspended in 0.01 M Tris-hydrochloride–0.001 M EDTA (pH 7.2) and brought to 0.05 M NaOH before electrophoresis on alkaline gels.

Alkaline gels were neutralized in 7% trichloroacetic acid before drying. In some cases, DNA was transferred from gels by the Southern procedure (16). In these experiments, the alkaline gels were neutralized directly in 3 M NaCl–0.5 M Tris-hydrochloride (pH 7.2) and transferred to nitrocellulose. The filters were hybridized to 32 P-labeled probes in 50% formamide–0.05 M HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [pH 7.0])–5 $\times$ Denhardt reagent (1 $\times$ Denhardt reagent is 0.02% each of Ficoll, polyvinylpyrrolidone, and bovine serum albumin)–3 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–salmon DNA (0.1 to 0.2 mg/ml)–0.01 M EDTA at 37°C overnight. The blots were washed four times in 2 $\times$ SSC–0.1% sodium dodecyl sulfate (SDS) at room temperature for 5 min each and two times for 30 min each in 0.1 $\times$ SSC–0.1% SDS at 50°C and rinsed in 0.1 $\times$ SSC at room temperature. All gels or blots were subjected to autoradiography at –70°C in the presence of intensifying screens.

For Northern RNA blots, 1.0% agarose gels were cast in 0.01 M NaPO₄ (pH 7.4)–2.2 M formaldehyde and electrophoresed with 0.01 M NaPO₄ as the running buffer. RNAs were denatured in a solution containing 0.01 M NaPO₄, 50% formamide, 2.2 M formaldehyde, and 0.5 mM EDTA at 65°C for 5 min and cooled at 20°C, and the solution was adjusted to 0.1% SDS–5% glycerol–0.005 M EDTA–0.005% bromophenol blue before loading. After electrophoresis, the gels were stained with acridine orange, photographed, treated

with 0.05 M NaOH for 1 h, and neutralized in 0.2 M sodium acetate (pH 4.0). RNAs were transferred to diazotized paper (Schleicher & Schuell, Inc.) in the neutralizing solution and prehybridized in 50% formamide–5 $\times$ SSC–0.05 M HEPES (pH 7.0)–5 $\times$ Denhardt reagent–salmon DNA (0.25 mg/ml)–0.005 M EDTA–0.1% SDS–1% glycine. Hybridization was performed in the same solution lacking glycine at 46°C for 3 days. The blots were washed as described above.

RESULTS

Size and kinetics of appearance of transcripts from the major transcription unit in the long repeat. In earlier experiments with a dot-blot analysis, we found that fragments near the L-S junction were major sites of transcription as early as 8 h after infection. A schematic diagram of the AD169 genome and a portion of L-S junction fragment *EcoRI*-WN (WN is a fusion of terminal fragments W and N, also referred to as *EcoRI*-H) and the adjacent long repeat fragment *EcoRI*-O is shown in Fig. 1. The brackets show the position of the adjacent classes of heterogeneity that have been mapped to the L-S junction (18).

The size and kinetics of accumulation of transcripts from this region were determined by Northern blot analysis. In these experiments, cytoplasmic, polyadenylated RNA from several time points was examined. To distinguish between IE and early transcripts, RNA was prepared at 8 h after infection either in the presence (IE RNA) or absence (early RNA) of cycloheximide (200 μ g/ml). RNA from the midpoint of infection (27 h) or from late times after infection (72 h) was also examined. Uninfected and infected cell RNAs were subjected to electrophoresis through denaturing formaldehyde gels and transferred to diazotized paper. The filters were hybridized to fragments spanning the long repeat. Figure 1A shows that a 2.7-kb RNA could be detected with *EcoRI* fragment O (probe A) at all times in the infection. The RNA was present at low levels in IE and early RNA and was most abundant at the midpoint and late in the infection. In addition, the probe hybridized to an IE RNA of 1.3 kb and to early, midpoint, and late RNAs of 1.2 kb. The 1.2-kb RNA was greatly amplified between 27 and 72 h after infection and appeared, at the level of steady-state accumulation, to be regulated differently from the 2.7-kb RNA. Minor late transcripts of 4.4, 3.6, 3.3, and 1.8 kb were also detected with this probe. A *PvuII*-*EcoRI* subfragment of *EcoRI*-O (Fig. 1B, Probe B) also hybridized to the 1.3- and 2.7-kb IE transcripts, to the 1.2- and 2.7-kb early, middle, and late RNAs, and to the minor late 3.6- and 1.8-kb RNAs. This probe did not detect the minor 4.4- or 3.3-kb late RNAs. *EcoRI*-h, a fragment containing the most internal portion of the long repeat, did not hybridize to any of the transcripts detected by *EcoRI*-O (data not shown).

Using a 32 P-labeled *EcoRI*-*PvuII* subfragment of *EcoRI*-WN as the hybridization probe (Probe C; Fig. 1C), we detected the 2.7-kb RNA at all times after infection. This RNA showed the same kinetics of appearance as the 2.7-kb RNA that was detected by *EcoRI*-O (probe A). RNA from 27 h contained a minor 1.8-kb transcript, and RNA from 72 h after infection contained minor 1.8- and 3.6-kb transcripts, which hybridized to this probe.

Initial mapping of the 2.7- and 1.2-kb RNAs. From the above studies, it appeared that the 2.7-kb RNA was transcribed from sequences in *EcoRI* fragment O as well as from sequences within the *PvuII*-*EcoRI* fragment of *EcoRI*-WN in which we had mapped the region of heterogeneity. To delineate further the map position of the 2.7-kb RNA, we

FIG. 1. (Top) Northern analysis of transcripts synthesized from the long repeat of HCMV. Cytoplasmic polyadenylated RNA extracted from ca. 5×10^6 uninfected (U) or infected cells 8 h (IE and E) and 27 h (M) after infection and from ca. 10^6 cells 72 h (L) after infection was run on a denaturing formaldehyde gel, transferred to diazotized paper, and hybridized to ^{32}P -labeled fragments (IE RNA was prepared in the presence of cycloheximide). RNAs in panels A, B, and C were hybridized to probe A, *EcoRI*-O; probe B, a 3.2-kbp *PvuII*-*EcoRI* subfragment of *EcoRI*-O; and probe C, the 4.2-kbp *EcoRI*-*PvuII* subfragment of *EcoRI*-WN, respectively. In panel B, the autoradiograms represent an 11-h exposure (left side) and a 4-h exposure (right side). In panel C, the left-hand autoradiogram was exposed for 14 h, and the right-hand autoradiogram was exposed for 4 h. The markers were human and *E. coli* ribosomal RNAs, and sizes are shown in kb. (Bottom) Diagram of the genome structure and *EcoRI* restriction endonuclease sites (17) and a more detailed map of the fragments used for hybridization. The brackets denote the position of the three classes of heterogeneity mapped to the L-S junction (18). Not all *SmaI* sites are shown.

used the Berk-Sharp technique (1). In this procedure, cloned viral DNA fragments are hybridized to RNA under conditions that favor RNA-DNA hybrid formation and then are treated with specific nucleases. The enzyme S1 nuclease digests any single-stranded nucleic acid resulting in DNA strands equal in length to transcribed exons, and exonuclease VII is unable to remove internal single-stranded loops and leaves DNA strands equal in length to transcribed exons plus intervening sequences.

Figure 2 shows the results of S1 nuclease protection assays with a 7.4-kbp *PvuII* fragment from the L-S junction. This fragment, represented by the bar at the bottom of the figure, was constructed by ligation of the *PvuII*-*EcoRI* subfragment of *EcoRI*-O to the *PvuII*-*EcoRI* subfragment of *EcoRI*-WN as described above. In this experiment, the unlabeled O-WN fusion *PvuII* fragment was hybridized to RNA from late in the infection, treated with S1 nuclease at 20, 30, 37, or 45°C, and analyzed on neutral and alkaline gels. The separated fragments were transferred to nitrocellulose and hybridized to a *PvuII*-*EcoRI* subfragment of *EcoRI*-O (probe A; Fig. 2A), or *EcoRI*-WN (probe B; Fig. 2B). When the S1 nuclease digestion was performed at 20, 30, or 37°C, we detected with both probes a 2.5-kb protected fragment on neutral and alkaline gels. This protected fragment is of the size expected for a 2.7-kb RNA with a polyadenylated tail of 200 residues. When the RNA-DNA hybrids were treated with S1 nuclease at 45°C, the 2.5-kb fragment was no longer detected. In its place, a 2.0-kb

fragment was detected with *EcoRI*-WN (Fig. 2B), and a 0.5-kb fragment was detected with a subfragment of O (Fig. 2A). The fact that fragments of identical size were observed after neutral or alkaline gel analysis implied that digestion at 45°C resulted in clipping of both DNA and RNA strands of the hybrid. This defined a site of S1 nuclease sensitivity, rather than a splice site, 0.5 kb from the terminus of the RNA-DNA hybrid in *EcoRI*-O (see below). This S1 nuclease protection experiment also resulted in a 0.9-kb protected fragment that hybridized only to *EcoRI*-O (Fig. 2A). The protected fragment was clearly visible at all temperatures of digestion in the original autoradiograph and most likely represented the DNA fragment protected by the 1.2-kb RNA.

Fine mapping of the 2.7 kb RNA. The 5' end of the 2.7-kb transcript was mapped and tested for temporally regulated changes in transcription initiation or splicing by analysis with a 5'-labeled probe (Fig. 3). *EcoRI*-O was 5' labeled and cleaved with *Bam*HI. The subfragment bordering *EcoRI*-WN was gel purified, hybridized to IE, 8-, 28-, and 72-h RNA, treated with S1 nuclease at 20°C or exonuclease VII at 37°C, and subjected to alkaline gel electrophoresis. RNA from all times protected fragments of 0.63 kb, and the size of the protected fragment was the same with both S1 nuclease and exonuclease VII (Fig. 3).

To map the 3' end of the RNA, *EcoRI* fragment WN was 3' labeled with T4 DNA polymerase and cleaved with *PvuII*. The 4.2-kbp fragment nearest *EcoRI*-O, represented by a bar

FIG. 3. Localization of the 5' end of the 2.7-kb RNA. A *Bam*HI subfragment of *Eco*RI-O, uniquely 5'-labeled at the *Eco*RI site (represented by the bar below the restriction map [bottom]), was hybridized to 12 μ g of IE RNA (IE lanes), 5 μ g of 8-h RNA (E lanes), 1 μ g of 28-h RNA (M lanes), and 2 or 0.7 μ g of 72-h RNA (L lanes). Lanes marked (-) were hybridized in the absence of infected cell RNA. The reactions were divided and treated with S1 nuclease at 20°C (lanes S) or exonuclease VII (lanes X). Hybrids were run on alkaline 1.5 or 2.0% agarose gels with phage ϕ X174 DNA cut with *Hae*III as the size markers (lanes m), and sizes are shown in nucleotides. The arrow shows the position of the 2.7-kb transcript on the restriction map (bottom).

fragments were observed. Therefore, no evidence for 30- to 350-base exons was obtained by this analysis. We were also unable to find evidence for introns in this experiment, as agarose gel analysis showed no size differences between S1 nuclease- or exonuclease VII-digested fragments (data not shown).

The direction of transcription of this RNA was determined in a separate hybridization reaction with a 3'-labeled probe consisting of the *Pvu*II-*Eco*RI subfragment of *Eco*RI-O (probe B) uniquely labeled at the *Pvu*II site. The RNA-DNA hybrids were treated with S1 nuclease or exonuclease VII and subjected to electrophoresis on an alkaline gel. A 0.9-kb

FIG. 4. Localization of the 3' end of the 2.7-kb transcript. This figure shows the alkaline gel analysis of a subfragment of *Eco*RI-WN (denoted by the bar below the diagram [bottom]) uniquely 3' labeled at the *Eco*RI site and hybridized in the presence of cytoplasmic polyadenylated RNA from IE, early, midpoint, or late in the infection or in the absence of infected cell RNA (top). The hybrid was treated with S1 nuclease (S1) or exonuclease VII (Exo) as described for Fig. 3. Exposures ca. eightfold shorter are shown for M and L lanes. The 4.2-kb band is equal in length to the undigested fragment. Symbols are as described for Fig. 3. The markers were phage λ DNA, and fragment sizes are shown in kb.

FIG. 5. Map position of the 1.2-kb RNA. (Top) (A) Unlabeled *EcoRI*-O, a *PvuII*-*EcoRI* subfragment of *EcoRI*-O, and a *SmaI*-*EcoRI* subfragment of *EcoRI*-O (designated probes A, B, and C, respectively, below the map [bottom]), were hybridized in the presence (+) or absence (-) of 72-h RNA and treated with S1 nuclease at 20°C. Protected fragments were analyzed by electrophoresis on an alkaline 1.5% agarose gel, transfer to nitrocellulose, and hybridization to ³²P-labeled *EcoRI*-O. (B) Probe A, hybridized as in panel A, treated with S1 nuclease (S1) or exonuclease VII (Exo) and analyzed as in (A). (C) Probe B, uniquely 3' labeled at the *PvuII* site, was hybridized in the presence (+) or absence (-) of 72-h RNA, treated with S1 nuclease or exonuclease VII, and subjected to electrophoresis on an alkaline 1.5% agarose gel. The high-molecular-weight band comigrated with the intact labeled fragment and was also seen on longer exposures of the (-) lane. Markers (M) were phage ϕ X174 DNA cleaved with *HaeIII*, and sizes are shown in kb. The arrows on the map represent the positions and directions of the 1.2- and 2.7-kb RNAs, and the brackets denote the position of the heterogeneity at the L-S junction.

fragment was protected from S1 nuclease and exonuclease VII digestion, and as expected, the 0.63-kb fragment was not observed (Fig. 5C). These results indicate that the 1.2-kb RNA was transcribed in the same direction as the 2.7-kb RNA and that the 3' end of the 1.2-kb RNA mapped 1.7 kb upstream of the 5' end of the 2.7-kb transcript.

DISCUSSION

We have characterized the major transcripts encoded by the long repeat of HCMV strain AD169. A 2.7-kb transcript was detected in RNA from IE, early, middle, and late times in the infection but not in RNA from uninfected cells. The detection of the 2.7-kb RNA at IE and early times is consistent with our earlier analysis of transcription from this region. Using a dot-blot technique, we found low levels of transcription from *EcoRI* fragments O, WL (F), and WN (H) at these times (9). Most IE transcripts accumulate in the absence of protein synthesis but are otherwise turned off early in the infection. The 1.3-kb IE RNA that hybridized to *EcoRI*-O may represent the classic IE RNA, as its abundance was much lower in early RNA. The 2.7-kb transcript, in contrast, became more abundant as the infection progressed. Although this RNA may represent an unusual class of IE transcripts which is accumulated throughout the infection, the cycloheximide treatment used in these experiments may not block protein synthesis completely and may allow some early transcripts to be synthesized. Further experiments are required to determine conclusively the class of transcripts in which the 2.7-kb RNA belongs.

Our previous analysis of steady state and de novo RNA synthesis indicated that transcription from the long repeat

continues to increase with time, becoming particularly abundant at the midpoint and late times in the infection (9). The Northern analysis presented here showed that much of the observed midpoint and late transcription is due to the 2.7-kb RNA, although other species detected by fragment *EcoRI*-O contributed to our observations. It was of interest to note that the two most abundant RNAs in the long repeat are regulated differently. The 2.7-kb RNA is greatly amplified between 8 and 27 h after infection, whereas the 1.2-kb RNA is amplified later, between 27 and 72 h after infection. Transcripts of 2.7 and 1.2 kb have also been detected with DNA fragments containing the long repeat sequences of the Towne strain of HCMV in RNA extracted from cells at 24 and 72 h after infection (23). In contrast, the only major transcript detected with the repeated sequences of the Davis strain is a 1.7-kb RNA (3). The reason for this difference is not known but could be related to the fact that the long repeat in the Davis strain genome is ca. 7 to 8 kbp shorter than that in the Towne and AD169 genomes (2, 4, 5, 17, 25).

In experiments with the *PvuII* O-WN fusion fragment, the 2.7-kb RNA protected a DNA fragment of 2.5 kb at temperatures of S1 digestion from 20 to 37°C. This is the size predicted for a 2.7-kb RNA with a 200-nucleotide polyadenylated tract, which implies that the transcript mapped completely within the borders of the *PvuII* fragment. As expected, the protected fragment hybridized to subfragments of both *EcoRI*-O and *EcoRI*-WN.

S1 nuclease digestion of the 2.7-kb RNA-DNA hybrid at 45°C resulted in two fragments of 2.0 and 0.5 kb rather than the 2.5-kb fragment. This led to the identification of an S1 nuclease-sensitive site 0.5 kb from the 5' end of the tran-

script. This S1 nuclease sensitivity could be caused by a spliced-RNA-DNA hybrid, which would contain a single-stranded DNA loop, or by a thermosensitive site in the RNA-DNA hybrid, which would contain single-stranded RNA and DNA strands. In the first case, the single-stranded DNA loop should be cleaved by S1 nuclease, resulting in a hybrid that migrates as a 2.5-kb band on neutral gels and as 2.0- and 0.5-kb bands on alkaline gels. This was not observed. The shorter fragments were seen on both neutral and alkaline gels and only at the highest temperature of S1 nuclease digestion. This indicated that the RNA-DNA hybrid was completely contiguous but became S1 nuclease sensitive at 45°C due to strand denaturation. Treatment of the full-length RNA-DNA hybrid with a 10-fold excess of S1 nuclease, up to 10,000 U/ml, at 20°C did not generate any of the 2.0-kb fragment (data not shown), making it unlikely that an increase in S1 nuclease activity between 37 and 45°C was responsible for the production of the novel fragments. Preliminary DNA sequencing of this region has revealed a string of 22 deoxyadenosines punctuated by four thymidine residues in the transcribed strand (data not shown). The resulting RNA-DNA hybrid would be thermodynamically unstable and therefore sensitive to S1 nuclease at elevated temperatures. The comparison of exonuclease VII- and S1 nuclease-digested fragments shown in Fig. 3 is consistent with our interpretation that no splicing occurs in the 5' end of the 2.7-kb RNA.

The 5' end of the 2.7-kb transcript from all times in the infection mapped 0.63 kb from the border of *EcoRI*-WN in *EcoRI*-O. The length of the 5'-labeled protected fragment was the same after either S1 nuclease or exonuclease VII digestion, indicating that there were no large spliced segments in the 5' end. When the experiments were repeated at lower temperatures to facilitate formation of small unstable RNA-DNA hybrids, the length of the 5' protected fragment remained unchanged and there was no evidence for smaller exons (data not shown). Although our experiments suggest that the 2.7-kb RNA contains no additional exons, further experiments are required to position precisely the 5' end of this RNA.

The 3' end of the 2.7-kb transcript mapped 1.85 kb to the right of the O-WN boundary in *EcoRI*-WN and showed no evidence for S1 nuclease sensitivity or splicing. The 3' terminus was the same for RNA prepared at IE, early, midpoint, or late times. This placed the transcript completely within the long repeat, with its 3' end proximal to the terminus of the L segment. The DNA sequence from the region encoding the 3' terminus of the RNA has been determined (18), and AATAAA, the putative signal for polyadenylation or cleavage or both (14), occurs in the nontranscribed strand starting 238 nucleotides from a *Bam*HI site, or ca. 1.84 kb from the *EcoRI* site. The position corresponds well to the location of this highly conserved 3' regulatory sequence in most eucaryotic mRNAs.

The 1.2-kb RNA was mapped within *EcoRI*-O, crossing a *Pvu*II site 3.2 kb from the O-WN border. The 2.7- and 1.2-kb RNAs were transcribed in the same direction and were contained within the long repeat, with their 3' ends mapping 2 and 6.2 kb from the L-S junction, respectively. The three classes of heterogeneity that have been described for the AD169 strain of HCMV consist of a single copy of a 0.55-, 0.95-, or 1.35-kbp fragment (class I heterogeneity), multiple tandem duplications of one of these fragments (class II heterogeneity), or multiple tandem copies of a 0.22-kbp fragment (class III heterogeneity). The characterized L-S junction fragments vary in size by as much as 3 kbp (18). Our

analysis of transcription from this region indicates that the major transcripts are separated from the region of heterogeneity by at least 1.6 kbp of constant DNA. The accumulation of tandemly repeated sequences, such as those found at the L-S junction, might have an effect on the efficiency of transcription of genes in the immediate vicinity, but this has not been investigated.

Most herpesvirus transcripts examined to date are unspliced, as the 2.7- and 1.2-kb RNAs of HCMV appear to be. The 2.7-kb RNA has the same signal AAUAAA at its 3' end as other herpesvirus RNAs. It will be of interest to examine the sequences upstream of these RNAs for α and β regulatory elements already described for HSV-1 (6, 7, 10, 11, 12, 21, 22, 24) and upstream sequences described for the major HCMV IE RNA (20).

ACKNOWLEDGMENTS

We thank S. Shaw, J. Tamashiro, and J. Marks for helpful discussions.

This research was supported by Public Health Service grant CA 34729 and training grant 5T32 GM07240 from the National Institutes of Health.

LITERATURE CITED

1. Berk, A. J., and P. A. Sharp. 1978. Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 75:1274-1278.
2. DeMarchi, J. M. 1981. Human cytomegalovirus DNA. Restriction enzyme cleavage maps and map locations for immediate-early, early, and late RNAs. *Virology* 114:23-38.
3. DeMarchi, J. M. 1983. Nature of the block in the expression of some early virus genes in cells abortively infected with human cytomegalovirus. *Virology* 129:287-297.
4. Fleckenstein, B., I. Müller, and J. Collins. 1982. Cloning of the complete human cytomegalovirus genome in cosmid. *Gene* 18:39-46.
5. LaFemina, R. L., and G. S. Hayward. 1980. Structural organization of the DNA molecules from human cytomegalovirus. *ICN-UCLA Symp. Mol. Cell. Biol.* 18:39-55.
6. Mackem, S., and B. Roizman. 1982. Differentiation between α promoter and regulator regions of herpes simplex virus 1: the functional domains and sequence of a movable α regulator. *Proc. Natl. Acad. Sci. U.S.A.* 79:4917-4921.
7. Mackem, S., and B. Roizman. 1982. Structural features of the herpes simplex virus α gene 4, 0, and 27 promoter-regulatory sequences which confer α regulation on chimeric thymidine kinase genes. *J. Virol.* 44:939-949.
8. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. McDonough, S. H., and D. H. Spector. 1983. Transcription in human fibroblasts permissively infected by human cytomegalovirus strain AD169. *Virology* 125:31-46.
10. McKnight, S. L. 1980. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. *Nucleic Acids Res.* 8:5949-5964.
11. McLauchlan, J., and J. B. Clements. 1983. Organization of the herpes simplex virus type 1 transcription unit encoding two early proteins with molecular weights of 140,000 and 40,000. *J. Gen. Virol.* 64:997-1006.
12. Murchie, M. J., and D. J. McGeoch. 1982. DNA sequence analysis of an immediate-early gene region of the herpes simplex virus type 1 genome (map coordinates 0.950 to 0.978). *J. Gen. Virol.* 62:1-15.
13. Oram, J. D., R. G. Downing, A. Akrigg, A. A. Dollery, C. J. Duggleby, G. W. G. Wilkinson, and P. J. Greenaway. 1982. Use of recombinant plasmids to investigate the structure of the human cytomegalovirus genome. *J. Gen. Virol.* 59:111-129.
14. Proudfoot, N. J., and G. G. Brownlee. 1976. 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature (London)* 263:211-214.

15. Rigby, P. W. J., M. Dieckmann, C. Rhodes, 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.
16. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503-517.
17. Spector, D. H., L. Hock, and J. C. Tamashiro. 1982. Cleavage maps for human cytomegalovirus DNA strain AD169 for restriction endonucleases *EcoRI*, *BglII*, and *HindIII*. *J. Virol.* **42**: 558-582.
18. Tamashiro, J. C., D. Filpula, T. Friedmann, and D. Spector. 1984. Structure of the heterogeneous L-S junction region of human cytomegalovirus strain AD169 DNA. *J. Virol.* **52**: 541-548.
19. Tamashiro, J. C., L. J. Hock, and D. H. Spector. 1982. Construction of a cloned library of the *EcoRI* fragments from the human cytomegalovirus genome (strain AD169). *J. Virol.* **42**:547-557.
20. Thomsen, D. R., R. M. Stenberg, W. F. Goins, and M. F. Stinski. 1984. Promoter-regulatory region of the major immediate early gene of human cytomegalovirus. *Proc. Natl. Acad. Sci. U.S.A.* **81**:659-663.
21. Wagner, E. K. 1983. Transcription patterns in HSV infections. *Adv. Viral Oncol.* **3**:239-270.
22. Wagner, M. J., J. A. Sharp, and W. C. Summers. 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. *Proc. Natl. Acad. Sci. U.S.A.* **78**: 1441-1445.
23. 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.
24. Watson, R. J., M. Sullivan, and G. F. Vande Woude. 1981. Structures of two spliced herpes simplex virus type 1 immediate-early mRNA's which map at the junctions of the unique and reiterated regions of the virus DNA S component. *J. Virol.* **37**:431-444.
25. Weststrate, M. W., J. L. M. C. Geelen, and J. van der Noordaa. 1980. Human cytomegalovirus DNA: physical maps for the restriction endonucleases *BglII*, *HindIII*, and *XbaI*. *J. Gen. Virol.* **49**:1-21.