Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 11137–11142, October 1996
Microbiology

Human cytomegalovirus latent gene expression in granulocyte–macrophage progenitors in culture and in seropositive individuals

(herpesvirus/cDNA/leukocytes/fetal liver/peripheral blood)

KAZUHIRO KONDO*, JIAKE XU, AND EDWARD S. MOCARSKI†

Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305-5402

Communicated by I. Robert Lehman, Stanford University Medical Center, Stanford, CA, July 22, 1996 (received for review April 19, 1996)

ABSTRACT **Following infection with cytomegalovirus, human granulocyte–macrophage progenitors carry the viral genome but fail to support productive replication. Viral transcripts arise from a region encompassing the major regulatory gene locus; however, their structure differs significantly from productive phase transcripts. One class, sense transcripts, is encoded in the same direction as productive phase transcripts but uses two novel start sites in the *ie*1/*ie*2 promoter/enhancer region. These transcripts have the potential to encode a novel 94 aa protein. The other class, antisense transcript, is unspliced and complimentary to *ie*1 exons 2–4, and has the potential to encode novel 154 and 152 aa proteins. Consistent with a role in latency, these transcripts are present in bone marrow aspirates from naturally infected, healthy seropositive donors but are not present in seronegative controls. Sense latent transcripts are present in a majority of seropositive individuals. Consistent with the expression of latent transcripts, antibody to the 94 aa and 152 aa proteins is detectable in the serum of seropositive individuals. Thus, latent infection by cytomegalovirus is accompanied by the presence of latency-associated transcripts and expression of immunogenic proteins. Overall, these results suggest that bone marrow-derived myeloid progenitors are an important natural site of viral latency.**

Human cytomegalovirus (CMV) is a significant pathogen in immunocompromised individuals and neonates (1). Latency, a hallmark of all herpesviruses (2, 3), remains poorly understood for CMV (4). We previously investigated maintenance and expression of the viral genome in an experimental latent infection using granulocyte–macrophage lineage cells (5). Bone marrow (BM)-derived hematopoietic cells (6), granulocyte–macrophage progenitors (GM-Ps; ref. 7), and peripheral blood monocytes (8–10) from healthy seropositive carriers contain viral DNA, suggesting that GM-Ps may be a natural site of latency.

Previous investigations of viral gene expression in leukocytes and BM-derived hematopoietic progenitors of normal healthy seropositive individuals demonstrated CMV α gene transcripts by *in situ* hybridization (11) or by reverse transcription–PCR (RT-PCR) following the induction of cellular differentiation (6, 12), suggesting that viral gene expression is restricted during latency. To better understand CMV latency, we investigated the extent of viral gene expression in primary GM-Ps (5) and found two novel classes of CMV latency-associated transcripts (CLTs; ref. 13). Here, we characterize these transcripts and show that they can be detected in BM aspirates from healthy CMV seropositive adults. Furthermore, we show that naturally infected individuals mount a readily detectable serological response to proteins encoded by CLTs, suggesting that these products are encoded during natural infection.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Cell and Virus Culture. Human fetal liver hematopoietic cells, cultured as GM-Ps in suspension (14), were exposed to CMV strain RC256, a *lac*Z-tagged recombinant virus (5, 15) at a multiplicity of infection of 3. The initial density of cells was kept high (10^7 per ml) to maintain cell viability. Human foreskin fibroblasts (HFs) were used for virus propagation and plaque assay (16).

RT-PCR Analysis. RNA was prepared (17), cDNA was synthesized with SuperScript II reverse transcriptase (GIBCO/BRL) or thermostable rTth polymerase (Perkin–Elmer), and cDNA samples were subjected to PCR using conditions and primers described previously (ref. 13; see Fig. 1*A*). Briefly, the cycle parameters used here were: 30 cycles of 94°C for 1 min, 65°C for 1 min, and 72°C for 2 min (A cycle parameters); 30 cycles of 94°C for 1 min, 62°C for 1 min, and 72°C for 2 min (B cycle parameters); 40 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 3 min (C cycle parameters); 30 cycles of 94°C for 15 sec, 60°C for 1 min, and 72°C for 3 min (D cycle parameters); 30 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 2 min (E cycle parameters). PCR reactions contained 50 mM KCl, 10 mM Tris·HCl (pH 8.5), 2 mM $MgCl_2$, 1 mM of each primer, 200 mM of each dNTP, and 1.25 units of *Taq* polymerase (Boehringer Mannheim). Where indicated, GeneAmp XL PCR kit (Perkin–Elmer) was used under the manufacturer's specified conditions. All PCR products were separated by electrophoresis on 2.5% agarose gels, transferred to Hybond-N^+ (Amersham) membranes and hybridized with [γ-^{32}P]ATP (Amersham) end-labeled oligonucleotide probes as described (5).

Rapid Amplification of cDNA Ends (RACE). RACE (18) procedures were performed with primers (13) on RNA from 10^4 to 10^5 GM-Ps at 4 weeks postinfection (pi). RACE products were T-A cloned into the pGEM-T vector (Promega) according to the manufacturer's protocol, and colony blot hybridization was performed using [α-^{32}P]dCTP-random-primed or [γ-^{32}P]dATP-end-labeled probes (13). Nucleotide sequence was determined using Sequenase (Amersham) and the fmole DNA Sequencing System (Promega).

RNase Protection. Probes 1 and 2 were prepared from clones pON2233 and pON2234 respectively, (13) using the MAXIscript *in vitro* transcription kit (Ambion) in conjunction with T7 RNA polymerase and [α-^{32}P]UTP (Amersham).

RT-PCR Analysis of BM-Derived Cells. Needle aspirates of BM were collected using a trocar to prevent extraneous

Abbreviations: CMV, cytomegalovirus; BM, bone marrow; GM-P, granulocyte–macrophage progenitor; RT-PCR, reverse transcription–PCR; CLT, CMV latency-associated transcript; HF, human foreskin fibroblast; RACE, rapid amplification of cDNA ends; pi, postinfection; GST, glutathione *S*-transferase; LSS, latent infection transcription start site; PSS, productive infection transcription start site.

*Present address: Department of Microbiology, Osaka University Medical School, 2–2 Yamada-oka, Suita, Osaka 565 Japan.
†To whom reprint requests should be addressed. e-mail: mocarski@leland.stanford.edu.

contamination from skin or other tissues. Approximately 3×10^7 BM-derived mononuclear cells were prepared on Lymphoprep (GIBCO). RNA was isolated (19), treated with RNase-free RQ1 DNase (13), and further purified on an RNeasy total RNA kit (Qiagen, Chatsworth, CA). cDNA synthesis was primed with random hexamers (for sense transcripts) or with IEP2E (for antisense transcripts) using Superscript II. Nested amplification (sense transcript) was carried out with primers IEP1K and IEP3D followed by primers IEP1G and IEP2D (predicted product of 206 bp) using B cycle parameters. For the antisense transcript, PCR with primers IEP2AII and IEP4J was followed by primers IEP3C and IEP4BII (predicted product of 387 bp) using B cycle parameters (13). PCR products were detected by hybridization with γ^{32}P-ATP (Amersham) end-labeled oligonucleotide probes, IEP1M (sense) or IEP4AP (antisense). As a positive control, all cDNA samples were shown to be positive for interleukin 1α using a specific primer set, DM151/DM152 (Perkin–Elmer).

Glutathione *S*-Transferase (GST) Fusion Proteins and Immunoblot Analysis. RNA was isolated from CMV (AD169)-infected HFs at 8 h pi (19), cDNA was synthesized from 1 μg of RNA using primer HIE4A 5′-CCTCGAAAGGCTCATGAACC-3′ in a 20 μl reaction with Superscript II, PCR amplification was carried out on 5 μl with primers HIE4A and HIE1A (5′-ATCCACGCTGTTTTGACCTC-3′) using E cycle parameters, and a 458-bp *Eco*RV–*Sac*II fragment PCR product was cloned into *Eco*RV–*Sac*II-digested pBluescript (KS+) and designated pON2500 to prepare *ie*1 region cDNA. To add the 5′ sequences found on sense CLTs, a 470-bp *Cla*I–*Sac*II fragment from pON2500 and a 656-bp *Sac*II–*Spe*I fragment from pON308G (20) were ligated into *Spe*I–*Cla*I-digested pBluescript (KS+) and designated pON2501. To construct GST fusions (21), the following fragments were cloned after blunt-ending with Klenow or T4 polymerase into the appropriate *Sma*I-digested pGEX vectors: ORF94$_{1-94}$, a 336-bp *Hin*cII fragment from pON2501 into pGEX-2T (pON2303); ORF154$_{6-120}$, a 351-bp *Acc*I–*Apa*I fragment from pON308G into pGEX-3X (pON2304); ORF152$_{1-152}$, a 764-bp *Nco*I–*Sph*I fragment from pON308G into pGEX-2T (pON2305); and, IE1$_{232-400}$, a 444-bp *Nco*I–*Ear*I fragment from pON308G into pGEX-2T (pON2307). All clones were sequenced using the fmol DNA Sequencing System (Promega) with the pGEM-specific primer 5′-ATAGCATGGCCTTTGCAGGG-3′.

GST fusion proteins were produced in *Escherichia coli* DH5α (21) and affinity purified on glutathione agarose beads (Pharmacia; ref. 22). Elution of GST-ORF94 required boiling the beads for 5 min in 1% SDS and PBS, and further purification by elution from polyacrylamide gels following electrophoresis (23). Protein concentration was determined by Bradford assay (Bio-Rad). Serum samples (50 μl) were adsorbed with 1 ml of *E. coli* cell lysate at 4°C overnight, clarified by centrifugation for 5 min in a microcentrifuge, and used for immunoblot analysis at 1/100 final dilution with horseradish peroxidase-conjugated goat anti-human IgG (Vector Laboratories) as a secondary antibody and development using Enhanced Chemiluminescence system (Amersham).

RESULTS

Novel Transcripts in CMV-Infected GM-Ps. We previously showed that *ie*1/*ie*2 region is transcribed in infected GM-Ps (5, 13). To determine the template strand giving rise to these transcripts, IEP4BII (Fig. 1*A*) was used to prime synthesis of cDNA from the same (sense) and IEP2AII was used to prime synthesis from the opposite (antisense) strand relative to the productive transcripts (24–26). PCR with primers IEP2AII and IEP3D (*ie*1 exons 2 and 3) using A cycle parameters or IEP4BII and IEP3C (*ie*1 exons 3 and 4) using B cycle parameters demonstrated that transcripts were encoded by both DNA strands in infected GM-Ps. When sense cDNA was amplified, the exon 2–3 primer set yielded a 151-bp PCR product and the exon 3–4 primer set yielded a 217-bp product (Fig. 1*B*, lanes 2 and 7), both of which appeared similar in size to spliced transcripts (24–26) made in productively infected HFs (Fig. 1*B*, lanes 5 and 10). When antisense cDNA was amplified, the exon 2–3 primer set yielded a 263-bp product and exon 3–4 primers yielded a 387-bp product (Fig. 1*B*, lanes

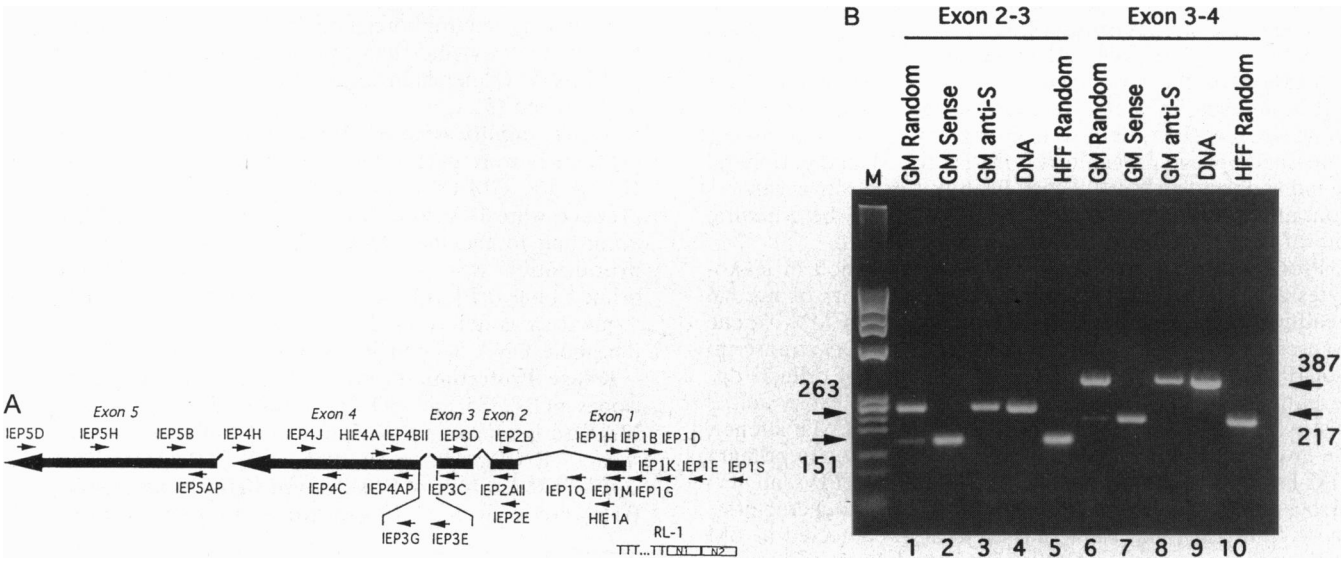

FIG. 1. Sense and antisense *ie*1/*ie*2 region transcripts in latently infected GM-Ps. (*A*) Position of primers (13) used in this study. (*B*) RT-PCR analysis of human CMV *ie*1 region gene expression in GM-Ps at four weeks pi. Expression between exons 2 and 3 (lanes 1–5) and between exons 3 and 4 (lanes 6–10) are shown. RNA from 10^4 latently infected GM-Ps (lanes 1–3 and 6–8) is compared with RNA from 10 productively infected (4 h pi) HFs (lanes 5 and 10). For samples in lanes 1, 5, 6, and 10, cDNA was synthesized using random hexamer primers and SuperScript II. For samples in lanes 2, 3, 7, and 8, five units of thermostable rTth polymerase (Perkin–Elmer) were used with primer IEP4BII to copy the sense strand (lanes 2 and 7) and with primer IEP2AII to copy the antisense (anti-S) strand (lanes 3 and 8) before PCR. Approximately 10^5 copies of viral DNA (lanes 4 and 9) were subjected to the same PCR conditions (13) for comparison. An ethidium bromide-stained 2.5% agarose gel is shown. Arrows adjacent to the lanes indicate the position of the predicted 151 bp spliced and 263 bp unspliced exon 2–3 products (lanes 1–5), and 217 bp spliced and the 387 bp unspliced exon 3–4 products (lanes 6–10). Size markers (M): *Hae*III-digested ϕX174 DNA.

3 and 8), both of which appeared similar in size to PCR products from viral DNA (Fig. 1*B*, lanes 4 and 9). Only spliced, sense transcript was detected in productively infected HFs (4 h pi) using either set of primers (Fig. 1*B*, lanes 5 and 10).

Structural Analysis of Sense Latent Transcripts. To evaluate the structure of sense transcripts, we used 5′- and 3′-RACE procedures (13, 18). Preliminary analysis positioned latent infection transcription start sites (LSSs) upstream of the productive infection transcription start site (PSS) in this region (Fig. 2*A*), and these were denoted CLTs (13). To map LSSs more precisely, we performed PCR on the 5′-RACE products and isolated clones by hybridization with a probe representing *ie*1/*ie*2 promoter-enhancer sequences. The 5′-RACE product was subjected to nested PCR with primers IEP2D and RL-1 using C cycle parameters followed by primers IEP1D and N1 (and GeneAmp XL) and D cycle parameters. The products from three preparations of GM-Ps, all from different donors, were analyzed. cDNAs were cloned into pGEM-T, and clones representing two different size classes were identified by restriction enzyme analysis. Two examples of longer clones (pON2218 and pON2219) contained one identical 5′ end (5′-TTTTTTTGTATCATATGCCAAGTCCG-3′) and three shorter clones (pON2222, pON2223, and pON2224) exhibited another identical 5′ end (5′-TTTTTTTATGCCCAGTACATGACCTT-3′). One other clone (pON2220) showed a additional nontemplate G (5′-TTTTTTTGGTATCATATGCCAAGTCCG-3′) but was otherwise similar to the longer clones. None of the clones isolated from infected GM-Ps initiated at the PSS. These results identified two LSSs: LSS1 (+1-GUAUCAUAUGCCAAGUACG-3′), corresponding to nt 174,086 on the CMV genome (GenBank accession no. X17403) 356 bp upstream of the PSS, and LSS2 (+1-

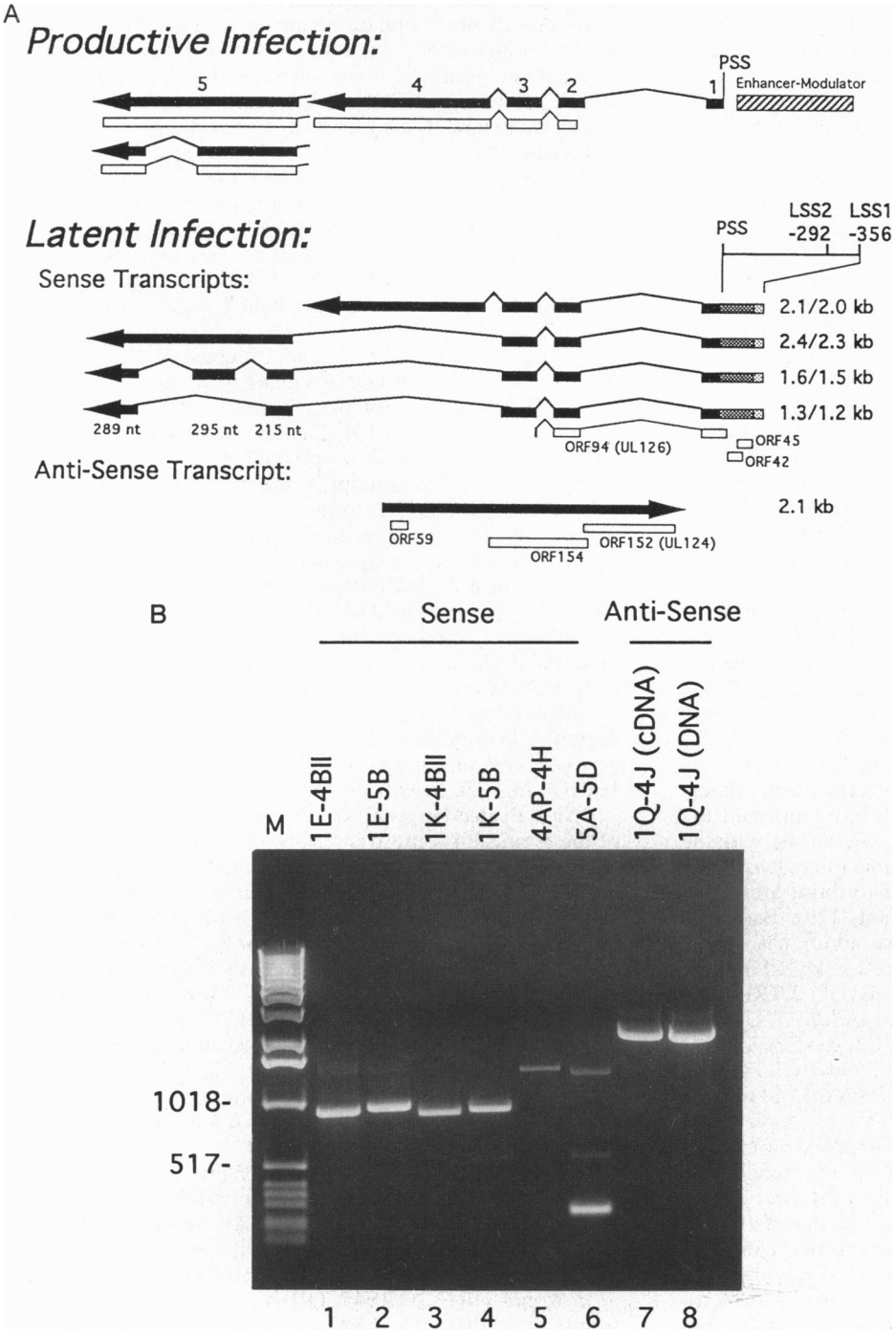

FIG. 2. (*A*) Summary of CLT structure. The upper line shows predominant α transcripts from the *ie*1/*ie*2 region (thick arrows) expressed from the PSS, +1-TCAGATCGCCTGGA-3′ (24, 25). The *ie*1 transcript (encoding a 491-aa protein) is composed of exons 1–4, and *ie*2 transcript (encoding a 579-aa protein) is composed of exons 1, 2, 3, and 5. A minor *ie*2 transcript species containing additional splicing in exon 5 is also depicted. The *ie*1/*ie*2 enhancer-modulator region (hatched box) is depicted upstream of the PSS. The lower lines show predominant sense and antisense CLTs. The expanded region depicts the PSS, as well as the additional sequences corresponding to the use of LSS1 and LSS2 with 5′ extensions of exon 1 detected in GM-Ps are depicted with differential shading. Open boxes denote ORFs that are conserved in strains Towne and AD169. The size of latency-specific exons derived by additional splicing events within exon 5 region is depicted below the exon 5 region of sense CLTs. (*B*) RT-PCR amplification of RNA obtained from latently infected GM-Ps. cDNA was made with Superscript II using either RNA from 10^4 GM-Ps at 4 weeks pi and random hexamer primers (lanes 1–6) or RNA from 10^5 GM-Ps at 4 weeks pi and oligo(dT) primer. All cDNA samples were subsequently amplified using 40 cycles PCR with GeneAmp XL PCR kit using B cycle parameters and compared with 10^5 copies of viral DNA amplified in the same manner (lane 8). Lanes: 1, IEP1E and IEP4BII; 2, IEP1E and IEP5B; 3, IEP1K and IEP4BII; 4, IEP1K and IEP5B; 5, IEP4AP and IEP4H; 6, IEP5AP and IEP5D; 7 and 8, IEP1Q and IEP4J.

AUGCCCAGUACAUGACCUU-3′), corresponding to nt 174,024 on the CMV genome 292 bp upstream of the PSS. Both LSS1 and LSS2 are located within the *ie*1/*ie*2 enhancer region (4, 27, 28), as depicted in Fig. 2*A*. Both are positioned downstream of putative TATA elements (13).

For 3′-RACE analysis of sense CLTs, first strand cDNA synthesis used primer RL-1 and SuperScript II followed by nested PCR with primers IEP2AII and N2 and primers IEP3C and N1 (and GeneAmp XL) with D cycle parameters (Fig. 1*A*). Three predominant PCR products were generated, ranging in size from under 1.0 to over 1.7 kbp (data not shown), indicating a considerable level of heterogeneity downstream of exon 3. To identify the 3′ ends of the sense transcripts, PCR products were T-A-cloned. Four clones were isolated and sequenced: one represented exon 4 and three represented different sized derivatives of exon 5. These clones revealed the same polyadenylylation sites as are used during productive infection (refs. 24 and 25; Fig. 2*A*).

Based on the positioning of the 5′ ends of latent transcripts, primers IEP1E or IEP1K specific for the 5′-ends of sense CLTs were synthesized and then used with the anchor primer RL-1 to complete the structural analysis of CLTs. RNA was isolated from 10^6 infected GM-Ps (at 4 weeks pi) and cDNA was synthesized using SuperScript II and primer RL-1. This product was subjected to PCR with primers IEP1E (specific for LSS1) and N2, or IEP1K (specific for LSS2) and N2, using GeneAmp XL and B cycle parameters. A small portion of this product was subjected to 25 additional cycles of PCR using primers IEP1G and N1 under the same conditions. Amplified products were cloned into the pGEM-T vector and candidate clones were identified by hybridization with ^{32}P-labeled exon 1 probe (from pON2347) as well as with oligonucleotide probes, IEP4AP and IEP5AP. These cDNA clones were evaluated by PCR analysis with exon-specific primer sets IEP1M–IEP2D, IEP2AII–IEP3D, IEP3C–IEP4BII, IEP3C–IEP5B, IEP4AP–IEP4H, IEP5AP–IEP5D, and IEP5AP–IEP5H, followed by separation on agarose gels (data not shown). Two different cDNA clones (pON2235 and pON2236), representing the 2.1 and 2.0 kb *ie*1 region cDNAs, and six different clones (pON2237–pON2242), representing the 2.4, 2.3, 1.6, 1.5, 1.3, and 1.2 kb *ie*2 species, were characterized by nucleotide sequence analysis, resulting in the structures depicted in Fig. 2*A*. The structure of CLTs was confirmed by RT-PCR of infected GM-P RNA using sequence specific primers. In agreement with the analysis of cDNA clones, RT-PCRs covering the region from LSS1 to exon 4 (Fig. 2*B*, lanes 1 and 3), from LSS1 to exon 5 (Fig. 2*B*, lanes 2 and 4), and the region within exon 4 (Fig. 2*B*, lane 5) were homogeneous. Three expected PCR products between primers IEP5AP and IEP5D (1300, 550, and 250 bp) confirmed the alternatively spliced forms of exon 5 (Fig. 2*B*, lane 6), with the most highly spliced form of exon 5 predominating. Sense CLT splicing patterns included species similar to those previously observed as low abundance splice variants (29). Based on sequence analysis of cDNA clones, these additional exon 5 region splice sites were within the 1.6/1.5 kb CLTs (5′-CCACGCGUCCUUUCAG/GUGAUUAUU. . .UCGUCU-UCCUCCUGCAG/UUCGGCUUC. . . .AAGAUUGAC-GAG/GUGAGCCGCA. . .UUUCCCAAACAG/GU-CAUGGUGCG-3′) and the 1.3/1.2 kb CLTs (5′-CCACGCGUCCUUUCAAG/GUGAUUAUU. . . UUCCCAAACAG/GUCAUGGUGCG-3′) as depicted in Fig. 2*A*. Thus, sense CLTs underwent complex differential splicing patterns that may be specific to the monocyte/macrophage lineage (29). When productively infected HFs (4 h pi) were analyzed, sense CLTs could not be detected using LSS-specific primers in RT-PCR (data not shown). Thus, PSS usage exceeded LSS site usage by at least 10^4-fold.

Structural Analysis of Antisense Latent Transcripts. Antisense CLTs were detected in infected GM-Ps at 3 or 4 weeks pi, but were not detected in productively infected HFs at 4 h pi (Fig. 1*B*). To map the 5′ ends of antisense CLTs, first strand cDNA was made using SuperScript II and primer IEP2E (Fig. 1*A*) and was 3′ tailed with poly(A). Nested PCR was carried out with primers RL-1 and IEP3E using C cycle parameters followed by primers N1 and IEP3G using D cycle parameters. The 5′-end of antisense CLTs was shown to be positioned ≈1.1 kbp upstream of the IEP3G annealing site (data not shown). To further characterize the 5′ ends, these 5′-RACE products were cloned in a T-A vector. Four clones hybridizing to ^{32}P-labeled *ie*1 exon 4 probe were sequenced. Two of these clones (pON2227 and pON2228) exhibited identical ends (+1-GTGACACCAGAGAATCAGAGG-3′; where +1 would correspond to nt 171,256 on AD169), and two clones (pON2225 and pON2226) were shorter by two (+1-GACACCAGAGAAT-3′) or six (+1-CCAGAGAAT-3′) bp. Such heterogeneity makes it difficult to predict a uniform start site of transcription for antisense CLTs.

To identify the 3′ end of the antisense transcript, RL-1 was used to prime cDNA synthesis on GM-P RNA using SuperScript II and nested PCR was carried out with primers IEP3D and N2 followed by IEP2D and N1 using GeneAmp XL and D cycle parameters. By analysis on agarose gels, the 3′-end of this transcript was shown to map ≈0.7 kbp downstream of the IEP2D primer annealing site (data not shown). These 3′-RACE products were T-A-cloned and all four examined (pON2229, pON2230, pON2231, and pON2232) exhibited an identical sequence consistent with a single polyadenylylation site (5′-AAATAATAAATGAGACCCCATCCTGTA-AAAAAA-3′; where the 3′ proximal T would correspond to nt 173,331 on AD169).

Having identified the 5′ and 3′ ends of the antisense transcripts, full-length cDNA clones were synthesized using primer IEP4J specific for the 5′ end of antisense transcripts and the anchor primer RL-1 (Fig. 1*A*). RNA was isolated from 10^6 infected GM-Ps (4 weeks pi), cDNA was synthesized with primer RL-1 and SuperScript II, and PCR was performed with primers IEP4J and N2 using GeneAmp XL and B cycle parameters. The antisense transcript was unspliced and initiated within a region complementary to *ie*1 exon 4 and terminated within a region complementary to the first intron of *ie*1/*ie*2 transcripts (Fig. 2*A*). Although the region upstream of the antisense transcript lacks a consensus TATA element, potential initiator sequences (5′-CGGGGACTCTGGGGGT-GACACCAGAGAAT-3′) similar to those found in the human terminal deoxynucleotidyl transferase gene (30, 31) were present. The antisense transcript was confirmed to be homogeneous and unspliced by PCR using primers IEP4J and IEP1Q (Fig. 2*B*, lanes 7 and 8).

RNase Protection Analysis. To confirm 5′-RACE mapping of the 5′ ends of sense transcripts, probe 1 (572 nt) was used and shown to protect species of 470 and 420 nt (Fig. 3, lane 2), with the longer (LSS1) species predominating. As expected, a 120 nt protected species consistent with PSS usage was in productively infected cells (Fig. 3, lane 3). Transcripts starting from PSS were not detected in latently infected GM-Ps, and transcripts starting from LSS1 or LSS2 were not observed in productively infected HFs (at 2 h pi). To confirm 5′-RACE mapping of the 5′ ends of the antisense transcript, probe 2 (603 nt) was used and shown to protect of a 220 nt species (Fig. 3, lane 5). These data also confirm that the use of LSS1 and LSS2 as well as the presence of the antisense CLT are latency associated.

Predicted ORFs on Latent Transcripts. Sequence analysis revealed the presence of several novel ORFs greater than 40 codons on latent transcripts (Fig. 2*A*). Several short ORFs were 5′ proximal on the sense transcripts initiating at LSS1 or LSS2. Transcripts that initiated at LSS1 contained three 5′-proximal ORFs, ORF45, ORF42, and ORF94, and transcripts that initiated at LSS2 contained ORF42 and ORF94.

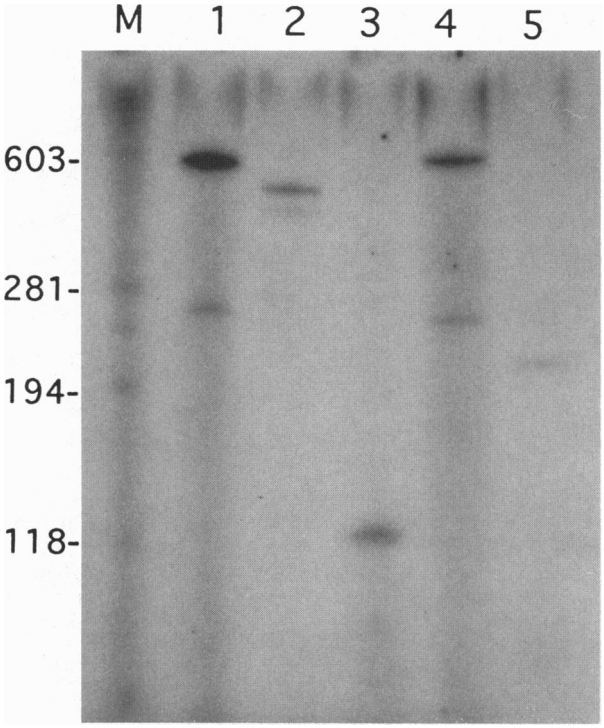

FIG. 3. RNase protection analysis of CLTs. RNA was extracted from 10^6 infected GM-Ps (lanes 2 and 5) or 10^3 infected HFs at 2 h pi (lane 3) as described in Fig. 1. RNase protection assay was performed, products were resolved following electrophoresis in an 8 M urea/5% polyacrylamide gel (23), and the gel was autoradiographed on Kodak XAR film. M, 5′-end labeled ϕoX174 DNA *Hae*III digest; lane 1, probe 1 alone; lane 2, RNA from infected GM-Ps hybridized with probe 1 and RNase digested; lane 3, RNA from 2 h pi HF hybridized with probe 1 and RNase treated; lane 4, probe 2 alone; lane 5, RNA from infected GM-Ps hybridized with probe 2 and RNase digested.

The amino terminus of the ORF94 represents a portion of UL126 (32). Three ORFs were identified in the antisense transcript, ORF59, ORF154, and ORF152 (where ORF152 corresponds to UL124; ref. 32). All of these ORFs were conserved in the Towne and AD169 genomes, although ORF94 (UL126) was not initially reported in Towne strain sequence due to a single missed base (27, 28). Although we have not yet evaluated expression of these ORFs in GM-Ps, their presence in sense transcripts would be expected to down-regulate expression the *ie*1 and *ie*2 productive gene products (24, 26, 33–35). Consistent with this expectation, we failed to detect expression of productive phase proteins $IE1_{491aa}$ or $IE2_{579aa}$ using immunofluorescence analysis (ref. 5; G. Hahn and E.S.M., unpublished results) with murine monoclonal antibody, CH160 (36).

Presence of CLTs in Naturally Infected Individuals. BM aspirates from 15 healthy adult donors at Stanford University Hospital were evaluated for CLTs using RT-PCR amplification. The CMV serology was unknown at the time of analysis. Sense transcripts were amplified from random-primed cDNA using latency-specific primers (IEP1K and IEP3D followed by IEP1G and IEP2D) and was detected in five of seven seropositive donors (Table 1). Antisense CLT was amplified from two of these five cDNA samples following PCR using a nested primer set (IEP2AII and IEP4J followed by IEP3C and IEP4BII; Table 1). CLTs were detected only in seropositive donors and not in any of eight seronegative donors, and, in reconstructions, this method was capable of detecting one cDNA copy in RNA from 10^8 cells. RT-PCR analysis for other viral α genes (UL36 and TRS1) was uniformly negative for all BM mononuclear cells RNA samples, similar to the pattern seen in GM-Ps (ref. 5; data not shown). We failed to recover virus from any BM sample despite cocultivation with HFs for 2 months (data not shown) using conditions that reproducibly led to reactivation in our experimental latent infection (5). These data show that sense CLTs are present in more than 70% of seropositive BM donors and suggest they may encode proteins that play important roles during latency.

Table 1. CLTs in healthy adult bone marrow donors

Donor (SPN)	Sense CLT	Antisense CLT	ELISA	Gender	Age
841	+	+	+	M	34
854	+	−	+	M	39
858	+	+	+	M	39
865	−	−	−	F	29
872	−	−	+	M	45
878	−	−	−	M	28
900	−	−	−	M	40
904	−	−	+	F	28
907	−	−	−	F	26
935	+	−	+	F	47
936	−	−	−	M	36
957	−	−	−	M	39
972	−	−	−	F	49
987	+	−	+	F	38
991	−	−	−	F	37

SPN, Stanford patient number.

Serological Response to Latent Proteins. To investigate which ORFs might be expressed in naturally infected seropositive individuals, serum antibody to ORF94, ORF154, and ORF152 was investigated using GST fusion proteins. The entire ORF94 and ORF152, as well as aa 6–120 of the ORF154 and aa 232–400 of $IE1_{491aa}$, were cloned into *E. coli* as carboxyl-terminal fusions with GST (21). Following affinity purification, these proteins were greater than 90% pure (data not shown). Affinity-purified fusion proteins were separated by SDS/PAGE and subjected to immunoblot analysis. Cell lysates prepared from CMV-infected HFs (168 h pi) and the $IE1_{491aa}$ fusion protein were included as positive controls, whereas cell lysates prepared from mock-infected HFs (data not shown) and GST alone served as negative controls. A total of 27 serum samples were analyzed, and the results from the 15 seropositive samples are shown in Table 2. Among the CMV seropositive samples, 47% (7 of 15) were reactive with $IE1_{232-400}$, 47% (7 of 15) were reactive with ORF94, and 20% (3 of 15) were reactive, some weakly, with ORF152. Seronegative sera failed to react either with CMV-infected cell lysates or with any of the GST fusion proteins. These results indicate that antibodies to the ORF94 protein are present in sera from a high proportion of individuals that carry CMV. Thus, expression of proteins encoded by CLTs occurs during natural infection.

DISCUSSION

We have shown that molecular markers of latent CMV infection from an experimental infection of cultured GM lineage cells are detectable in naturally infected individuals. Cultured GM-Ps can now be investigated for additional insights into CMV latency and reactivation. Starting with an analysis that suggested an atypical pattern of *ie*1 region expression in GM-Ps (5, 13), this work has shown the level to which gene expression during latency differs from productive phase expression in the *ie*1/*ie*2 region and has revealed how novel latent gene products may be related to natural infection.

Because altered growth conditions can lead to reactivation of virus replication in cultured GM-Ps (5), the cell type and differentiation state clearly influence the balance between latency and replication. The importance of the interplay

Table 2. Immunoblot detection of CLT-encoded antigens by CMV-seropositive human sera

Serum sample	Infected HF	Recombinant proteins				
		GST alone	GST ORF94	GST ORF154	GST ORF152	GST $IE1_{232-400}$
1254567	+	−	−	−	−	−
1264772	+	−	−	−	−	+
1245056	+	−	+	−	±*	+
1231546	+	−	+	−	±*	+
1257315	+	−	+	−	−	−
1268039	+	−	+	−	−	+
841	+	−	−	−	−	+
854	+	−	−	−	−	−
858	+	−	−	−	−	−
935	+	−	+	−	−	+
1292882	+	−	−	−	−	−
1280687	+	−	+	−	+	+
1279463	+	−	−	−	−	−
1263554	+	−	+	−	−	−
1275872	+	−	−	−	−	−

*Detection required 10-fold longer exposure.

between viral and host cell functions needs to be carefully assessed to more fully understand these processes. The proteins encoded by CLTs may play roles in the establishment, maintenance, or reactivation of latent infection. In naturally infected individuals, self-renewing immature BM-derived progenitors may be a reservoir of latent virus from which reactivation occurs during immunosuppression or immunodeficiency, a situation that is analogous to the latent infection of B lymphocytes by EBV (3). The functions of CLTs or CLT-encoded proteins may be to insure genome maintenance analogous to the role EBNA-1 plays in EBV, or, alternatively, may be to play regulatory roles such as repression of α gene expression.

We thank Harry Sices and Danushka Formankova for technical assistance, Jeanette Baker and Karl Blume from the Stanford University Hospital Bone Marrow Transplantation Program for donor samples, William Dworsky for liver samples, and Hideto Kaneshima for many helpful discussions. Gabriele Hahn first used a higher starting cell density in establishing GM-P cultures. This work was supported by U.S. Public Health Service Grants SCOR HL33811 and RO1 AI33852 and by a grant from SmithKline Beecham.

1. Alford, C. A. & Britt, W. J. (1995) in *Fields Virology,* eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott–Raven, New York), pp. 2493–2534.
2. Roizman, B. & Sears, A. E. (1995) in *Fields Virology,* eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott–Raven, New York), pp. 2231–2296.
3. Kieff, E. (1995) in *Fields Virology,* eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott–Raven, New York), pp. 2343–2396.
4. Mocarski, E. S. (1995) in *Fields Virology,* eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott–Raven, New York), pp. 2447–2492.
5. Kondo, K., Kaneshima, H. & Mocarski, E. S. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11879–11883.
6. Minton, E. J., Tysoe, C., Sinclair, J. H. & Sissons, J. G. (1994) *J. Virol.* **68,** 4017–4021.
7. von Laer, D., Meyer-Koenig, U., Serr, A., Finke, J., Kanz, L., Fauser, A. A., Neumann-Haefelin, D., Brugger, W. & Hufert, F. T. (1995) *Blood* **86,** 4086–4090.
8. Bevan, I. S., Daw, R. A., Day, P. J., Ala, F. A. & Walker, M. R. (1991) *Br. J. Haematol.* **78,** 94–99.
9. Taylor-Wiedeman, J., Sissons, J. G., Borysiewicz, L. K. & Sinclair, J. H. (1991) *J. Gen. Virol.* **72,** 2059–2064.
10. Stanier, P., Kitchen, A. D., Taylor, D. L. & Tyms, A. S. (1992) *Mol. Cell. Probes* **6,** 51–58.
11. Schrier, R. D., Nelson, J. A. & Oldstone, M. B. (1985) *Science* **230,** 1048–1051.
12. Taylor-Wiedeman, J., Sissons, P. & Sinclair, J. (1994) *J. Virol.* **68,** 1597–1604.
13. Kondo, K. & Mocarski, E. S. (1995) *Scand. J. Infect. Dis. Suppl.* **99,** 63–67.
14. Baines, P., Masters, G., Booth, M. & Jacobs, A. (1987) *Exp. Hematol.* **15,** 809–813.
15. Spaete, R. R. & Mocarski, E. S. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7213–7217.
16. Spaete, R. R. & Mocarski, E. S. (1985) *J. Virol.* **54,** 817–824.
17. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
18. Ohara, O., Dorit, R. L. & Gilbert, W. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5673–5677.
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
20. Cherrington, J. M. & Mocarski, E. S. (1989) *J. Virol.* **63,** 1435–1440.
21. Smith, D. B. & Johnson, K. S. (1988) *Gene* **67,** 31–40.
22. Frangioni, J. V. & Neel, B. G. (1993) *Anal. Biochem.* **210,** 179–187.
23. Sambrook, J., Frisch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
24. Stenberg, R. M., Thomsen, D. R. & Stinski, M. F. (1984) *J. Virol.* **49,** 190–199.
25. Stenberg, R. M., Witte, P. R. & Stinski, M. F. (1985) *J. Virol.* **56,** 665–675.
26. Stenberg, R. M., Depto, A. S., Fortney, J. & Nelson, J. A. (1989) *J. Virol.* **63,** 2699–2708.
27. Boshart, M., Weber, F., Jahn, G., Dorsch-Hasler, K., Fleckenstein, B. & Schaffner, W. (1985) *Cell* **41,** 521–530.
28. Thomsen, D. R., Stenberg, R. M., Goins, W. F. & Stinski, M. F. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 659–663.
29. Kerry, J. A., Sehgal, A., Barlow, S. W., Cavanaugh, V. J., Fish, K., Nelson, J. A. & Stenberg, R. M. (1995) *J. Virol.* **69,** 3868–3872.
30. Sorscher, D. H., Yang, B., Bhaumik, D., Trangas, T., Philips, A. V., Chancellor, K. E. & Coleman, M. S. (1994) *Biochemistry* **33,** 11025–11032.
31. Bhaumik, D., Yang, B., Trangas, T., Bartlett, J. S., Coleman, M. S. & Sorscher, D. H. (1994) *J. Biol. Chem.* **269,** 15861–15867.
32. Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A. I., Kouzarides, T., Martignetti, J. A., Preddie, E., Satchwell, S. C., Tomlinson, P., Weston, K. M. & Barrell, B. G. (1990) *Curr. Top. Microbiol. Immunol.* **154,** 125–170.
33. Stinski, M. F. (1978) *J. Virol.* **26,** 686–701.
34. Stinski, M. F., Thomsen, D. R., Stenberg, R. M. & Goldstein, L. C. (1983) *J. Virol.* **46,** 1–14.
35. Wathen, M. W., Thomsen, D. R. & Stinski, M. F. (1981) *J. Virol.* **38,** 446–459.
36. Plachter, B., Britt, W., Vornhagen, R., Stamminger, T. & Jahn, G. (1993) *Virology* **193,** 642–652.