JOURNAL OF VIROLOGY, July 1991, p. 3915–3918
0022-538X/91/073915-04$02.00/0

Vol. 65, No. 7

Cytomegalovirus Homologs of Cellular G Protein-Coupled Receptor Genes Are Transcribed

ANTHONY R. WELCH, LISA M. McGREGOR, AND WADE GIBSON*

Virology Laboratories, Department of Pharmacology and Molecular Sciences, The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205

Received 11 March 1991/Accepted 8 April 1991

The genome of human cytomegalovirus (HCMV) contains three genes with homology to cellular G protein-coupled receptors (GCRs). Evidence is presented here that all three HCMV GCR genes are transcribed during infection, that they are transcribed as two sets of 3′-coterminal mRNAs, and that their transcription is restricted to the late phase of infection.

A family of three human cytomegalovirus (HCMV; strain AD169) genes showing homology to cellular G protein-coupled receptors (GCRs) was recently identified by computer analysis of the genome (3, 4). Interestingly, some of the earliest metabolic changes described in cells infected with HCMV are typical of G protein-mediated signal transduction responses, such as increased levels of intracellular calcium, cyclic adenosine monophosphate, and diacylglycerol (1, 2, 9). The finding that CMV has homologs of cellular GCR genes raised the possibility that products of one or more of these viral genes may cause the observed G protein-like metabolic changes. This speculation assumes that the CMV GCR genes are in fact expressed. The experiments reported here were done to test this assumption by determining whether these viral genes are transcribed.

The genomic locations of the three HCMV GCR genes and their general organization are diagrammed in Fig. 1A and B. Two of the genes are contiguous in the unique short (US) region of the genome and contain the open reading frames (ORFs) US27 and US28 (3, 4). Although both have potential TATA promoter elements and 3′ stop translation codons, only US28 has an adjacent 3′ polyadenylation signal. It has been suggested that the upstream US27 gene may utilize this same polyadenylation signal and be expressed as a longer 3′-coterminal transcript (4, 11). The third HCMV GCR homolog is located in the unique long (UL) region of the genome and contains the UL33 ORF which, like US27, has a potential TATA promoter element and a 3′ stop translation codon, but lacks a nearby polyadenylation signal (4). The presence of a polyadenylation signal immediately following the downstream UL33 ORF, however, suggested that UL33 and UL34 may also be expressed as overlapping 3′-coterminal transcripts. Further inspection of the UL34 ORF suggested that the coding sequence of the gene begins at the ATG codon 368 nucleotides downstream of the UL33 stop translation codon and that a potential TATA promoter element (i.e., TATTA) for this gene, referred to as UL34a, is present in the resulting intergenic region.

The following experiments were done to determine whether and when the three HCMV (strain AD169) GCR genes are transcribed and to investigate their structures in relation to their genetic organization. The experimental approach was to identify and characterize the transcripts by using single-strand, ^{32}P-labeled, antisense oligodeoxynucleotides as probes in Northern (RNA) assays of infected-cell mRNAs. Five probes were synthesized, each 29 or 30 nucleotides long and representing sequences unique to US27, US28, UL33, UL34, and UL34a, respectively, and not present in other reported GCR sequences (4, 11; EMBL data base accession no. X17403):

5′-GTCCTGGTAAGGTATCTCCTGCTTCACGC-3′
5′-CTCGGCCAGCAGACAGTGTAGTTCTTGCG-3′
5′-CGTAGCCACTGATTGTGAATTGTTGCCAGC-3′
5′-GCGAAATCTGTGTACCGTGATGACAGATCA-3′
5′-CTGGCTAGAGGTCCGTCTTCTTCGTCGGCG-3′

These probes are represented by rectangles in Fig. 1C. An antisense riboprobe to β-actin was also prepared (10) and used to compare relative amounts of mRNAs in different samples.

In the first experiment, mRNAs were prepared from subconfluent cultures of noninfected human foreskin fibroblasts (Fig. 2, lanes Mock) and from fibroblasts 7 days after infection with HCMV; separated by agarose gel electrophoresis and transferred to nitrocellulose (6); and sequentially probed with each of the five oligonucleotides. All five regions of the genome were transcribed as indicated by the specific hybridization patterns obtained (Fig. 2). The US27 probe detected a 2.9-kb transcript, consistent in size with an RNA initiating upstream of US27, reading through both the US27 and US28 sequences, and terminating at the first polyadenylation signal—3′ of the US28 ORF (Fig. 1B and C). When the blot was stripped (8) and rehybridized with a probe specific for the US28 ORF, a 2.9-kb transcript was again detected, along with a 1.3-kb species. The size of the 1.3-kb RNA approximates that predicted for a transcript from the US28 gene. These results are most simply explained by the US27 and US28 genes giving rise to a pair of overlapping 3′-coterminal transcripts, 2.9 and 1.3 kb, respectively. The higher-M_r band (i.e., between the 2.9-kb band and the top of the gel) seen in these and the UL33, UL34, and UL34a panels is attributed to nonspecific hybridization, primarily because it was also detected by nonrelated probes representing different regions of the genome (Fig. 2 and additional data not presented).

Similar results were obtained with the third HCMV GCR homolog gene containing the UL33 ORF. The probe for UL33 detected a 3.3-kb transcript, and the probe for UL34a detected the same 3.3-kb RNA along with a 1.4-kb species. These results parallel those for the US27 and US28 genes and are consistent with the 3.3-kb UL33 mRNA overlapping

* Corresponding author.

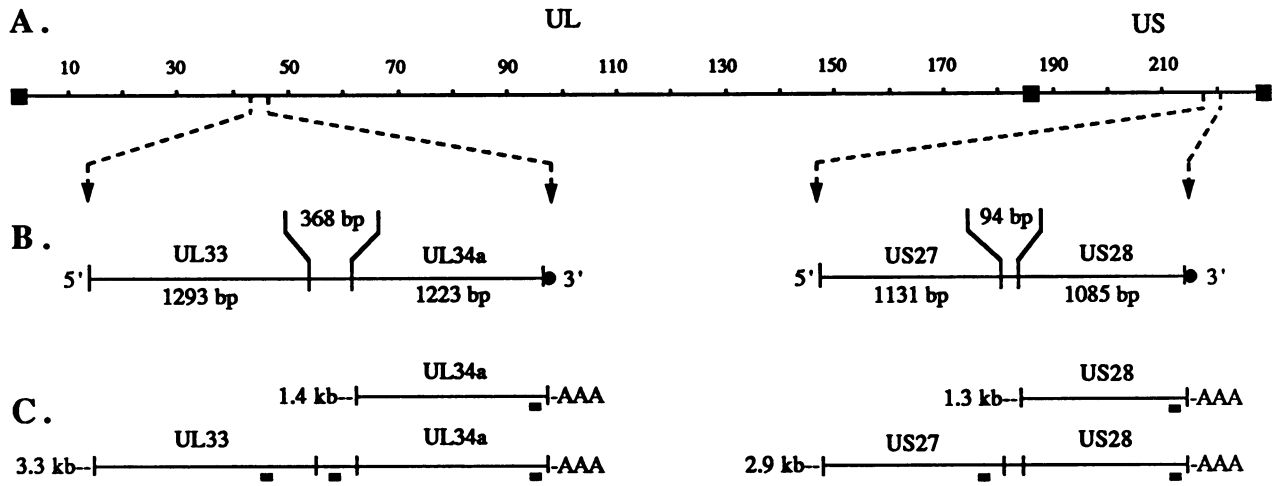

FIG. 1. General organization of HCMV GCR homolog genes, relative locations of specific oligonucleotide probes, and hypothetical structures of the HCMV GCR transcripts. (A) Locations of GCR homolog genes in the HCMV (AD169) genome (3, 4; EMBL data base accession no. X17403). The boundaries of the UL and US segments are indicated by filled squares, and units of measure are shown in 10-kb increments. (B) Genetic organization of the three HCMV GCR ORFs, US27, US28, and UL33, and of the UL34a portion of ORF UL34 (see text). The minimal lengths of the corresponding US27, US28, and UL33 genes are given as the number of nucleotides in their respective ORFs. The minimal length of the UL34a gene is given as the number of nucleotides between its proposed translational start ATG codon and the UL34 translational stop codon (see text). Intergenic, presumably nontranslated sequences are indicated between US27 and US28 (94 bp) and UL33 and UL34a (368 bp). (C) Proposed structures of the three HCMV GCR and UL34a transcripts and their experimentally determined sizes. 3′ polyadenosine tails are represented by -AAA, and putative 5′ untranslated sequences are indicated by the dotted portion of lines. Filled rectangles represent the antisense probes and the approximate locations of the sequences that they recognize.

and being 3′-coterminal with the 1.4-kb UL34a mRNA as indicated in Fig. 1C. The 3.3-kb RNA was also detected by the UL34 probe (Fig. 2), specific for sequences upstream of the putative promoter for UL34a (Fig. 1C), as would be expected if it is a read-through transcript containing this intergenic region downstream of the UL33 ORF (Fig. 1B and C). Failure of the UL34 probe to react with the 1.4-kb RNA indicates that this transcript does not contain the UL34 sequences upstream of the predicted UL34a promoter element. None of the five oligonucleotide probes reacted with RNA from mock-infected cells (Fig. 2) or from simian CMV strain Colburn-infected cells (data not shown), indicating that the transcripts detected are specific to HCMV-infected cells. A riboprobe to β-actin mRNA, a cellular transcript that does not change appreciably in amount following HCMV infection (5), showed comparable amounts of that transcript in the mock- and AD169-infected cell preparations (see arrow, Fig. 2), indicating that they represent similar numbers of cells.

Because the G protein-like metabolic changes observed in HCMV-infected cells are initiated very early after infection (1, 2, 9), it was of interest to determine whether the HCMV GCR genes are expressed as early (i.e., no preceding viral DNA synthesis is required) or late (i.e., viral DNA synthesis is required for maximal synthesis) transcripts. This was done by Northern assays of mRNAs prepared from normally infected cells and from cells infected and maintained in phosphonoformic acid (PFA; 200 μg/ml), an inhibitor of CMV DNA synthesis (7). mRNA was prepared 7 days after infection, when viral cytopathic effects were strong in the normally infected cells; separated by agarose gel electrophoresis and transferred to nitrocellulose (6); and detected by hybridization with the US28 and UL34a probes (Fig. 3). In agreement with the preceding results, the US28 probe detected a 2.9-kb and a 1.3-kb transcript, and the UL34a probe detected a 3.3-kb and a 1.4-kb transcript in normally infected cells (Fig. 3, lane AD169−) but not in mock-infected cells (lane Mock). Neither probe was reactive with RNAs in PFA-treated infected cells (lane AD169+), indicating that transcription of all three HCMV GCR genes, as well as that of the UL34a gene, is restricted to late times of infection. The US27 and UL33 probes were likewise nonreactive with RNAs in the PFA-treated infected cells (data not shown), establishing that they are not transcribed by alternate mechanisms (e.g., present as species other than the expected 2.9- and 3.3-kb RNAs, respectively) under these conditions. Results of this experiment were normalized as before by using the riboprobe to β-actin RNA (see arrow, Fig. 3), which showed that the preparations from nontreated and from PFA-treated infected cells contained essentially the same amount of β-actin RNA and therefore represented comparable numbers of cells. The comparatively greater amount of β-actin RNA in the preparation from mock-infected cells indicates that a greater number of cells were represented by that sample.

Taken together, these results demonstrate that all three HCMV GCR homolog genes and the UL34a gene are transcribed during lytic infection of human foreskin fibroblasts and are regulated temporally as late genes. The results also support the suggestion that transcripts from the US27 and UL33 gene utilize downstream polyadenylation signals since they lack their own (4, 11). Although the involvement of splicing or other RNA processing in generating the observed HCMV GCR-specific transcripts cannot be ruled out, the sizes (taking into account probable untranslated 5′ and 3′ sequences) and probe reactivities of the transcripts are most simply explained by the model diagrammed in Fig. 1C. It seems likely that the 1.3- and 1.4-kb RNAs are transcripts of the US28 and UL34a genes, respectively, and that the 2.9- and 3.3-kb RNAs are read-through transcripts containing the combined coding sequences of US27/US28 and UL33/UL34, respectively. No homology was found between the predicted

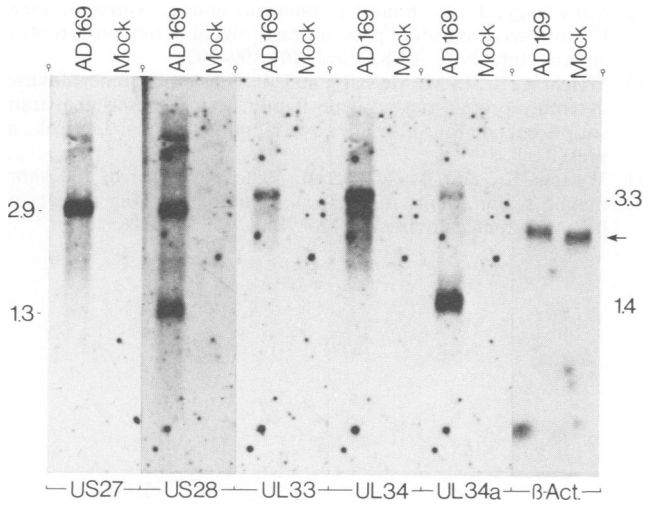

FIG. 2. Evidence that the HCMV GCR homolog genes and the UL34a gene are transcribed. Polyadenylated mRNA was recovered from cells 7 days after infection with HCMV strain AD169 (lanes AD169) or after mock infection (lanes Mock) by using the Fast Track kit (InVitrogen, San Diego, Calif.). The resulting RNA was separated in a denaturing (0.66 M formaldehyde) agarose gel, blotted onto nitrocellulose, and hybridized with the probes indicated at the bottom (6). Each lane contained approximately 1 μg of RNA as determined by A_{260}. The oligonucleotide probes (see text) were end labeled with [γ-^{32}P]ATP and incubated overnight (or for 6 h for the US28 probe) with the blot in a solution containing 6× SSPE, 5×Denhardt's solution, and 1% sodium dodecyl sulfate at 37°C with shaking. The blot was washed once in 6× SSPE at 37°C, once in 6× SSPE at 45°C, and once in 3× SSPE at 45°C; sealed in Saran Wrap to prevent drying; and visualized by fluorography or autoradiography. Blots hybridized with the UL33, UL34, and UL34a probes were additionally washed once in 1× SSPE at 55°C and twice in 0.5× SSPE at 55°C (UL33), or once with 1× SSPE at 55°C (UL34; UL34a). The same blot was probed sequentially with each of the five oligonucleotides and a ^{32}P-labeled riboprobe to β-actin mRNA (β-Act.) (10). Prior to reuse, the blot was stripped (8) and fluorographed overnight to verify removal of the preceding probe. Transcript sizes (in kilobases) were determined relative to RNA markers (Bethesda Research Laboratories, Gaithersburg, Md.) and are indicated in the left and right margins. Shown is a collage of the six resulting fluorograms, aligned by matching specific background spots common to all exposures except β-actin, which was aligned by positioning markers.

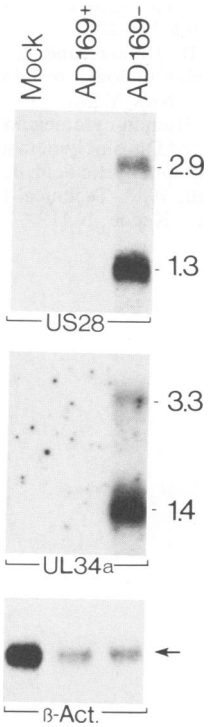

FIG. 3. Evidence that the HCMV GCR and UL34a genes are transcribed late. RNA prepared from noninfected cells (lane Mock) and from AD169-infected cells treated with PFA (200 μg/ml) (lane AD169+) or not treated (lane AD169−) was probed in sequential Northern assays, as described in the text and in the legend to Fig. 2, with ^{32}P-labeled oligonucleotides antisense to sequences within UL28 and UL34a and with the ^{32}P-labeled β-actin riboprobe (β-Act.). Shown are fluorographic images of the blot after reaction with each of the probes.

UL34a protein and other data base sequences. The amino acid sequence encoded by UL34a suggests a slightly acidic, 45-kDa protein.

Although the late transcription of these genes does not correlate directly with the very early, G protein-like metabolic changes observed in infected cells (1, 2, 9), the two observations can be reconciled if the CMV GCR proteins are synthesized at late times after infection and become incorporated into the virion envelope. As constituents of the virion, they would be introduced into the cell membrane immediately upon infection, where they could interact with resident G proteins, leading to very early activation of signal transduction pathways.

We thank Peter Collins and Randy Reed for advice concerning Northern assays using oligonucleotide probes; Annie Colberg-Poley and Dan Tenney for the p18-2 plasmid (β-actin); Clark Riley for assistance with computer comparisons of sequences; and Tom Albrecht, Mark Chee, and Peter Devreotes for their helpful comments on the manuscript. The oligonucleotide probes used in this study were synthesized by Jodie Franklin and Ponniah Shenbagamurthi, Protein/Peptide/DNA Facility, School of Medicine.

This work was aided by research grant AI13718 from the National Institute for Allergy and Infectious Diseases. A.R.W. and L.M.M. are students in the School of Medicine and were supported by U.S. Public Health Service grants.

REFERENCES

1. **AbuBakar, S., I. Boldogh, and T. Albrecht.** 1990. Human cytomegalovirus stimulates arachidonic acid metabolism through pathways that are affected by inhibitors of phospholipase-A2 and protein kinase-C. Biochem. Biophys. Res. Commun. **166**:953–959.
2. **Albrecht, T., I. Boldogh, M. Fons, C. H. Lee, S. AbuBakar, J. M. Russell, and W. W. Au.** 1989. Cell-activation responses to cytomegalovirus infection: relationship to the phasing of CMV replication and to the induction of cellular damage. Subcell. Biochem. **15**:157–207.
3. **Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchison III, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. K. M. Weston, and B. G. Barrell.** 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr. Top. Microbiol. Immunol. **154**:123–169.
4. **Chee, M. S., S. C. Satchwell, E. Preddie, K. M. Weston, and B. G. Barrell.** 1990. Human cytomegalovirus encodes three G protein-coupled receptor homologues. Nature (London) **344**: 774–777.
5. **Colberg-Poley, A. M., and L. D. Santomenna.** 1988. Selective

induction of chromosomal gene expression by human cytomegalovirus. Virology **166**:217–228.
6. **Davis, L. G., M. D. Dibner, and J. F. Battey.** 1986. Basic methods in molecular biology, p. 143–146. Elsevier Science Publishing Co., Inc., New York.
7. **Huang, E.-S.** 1975. Human cytomegalovirus. IV. Specific inhibition of virus-induced DNA polymerase activity and viral DNA replication by phosphonoacetic acid. J. Virol. **16**:1560–1565.
8. **Schleicher & Schuell.** 1987. Technical bulletin no. 700. Schleicher & Schuell, Inc., Keene, N.H.
9. **Valyi-Nagy, T., Z. Bandi, I. Boldogh, and T. Albrecht.** 1988. Hydrolysis of inositol lipids: an early signal of human cytomegalovirus infection. Arch. Virol. **101**:199–207.
10. **Welch, A. R., L. M. McNally, and W. Gibson.** Cytomegalovirus assembly protein nested gene family: four 3′-coterminal transcripts encode four in-frame, overlapping proteins. J. Virol., in press.
11. **Weston, K., and B. G. Barrell.** 1986. Sequence of the short unique region, short repeats, and part of the long repeats of human cytomegalovirus. J. Mol. Biol. **192**:177–208.