# Analysis of Transcripts Expressed from the UL47 Gene of Human Cytomegalovirus

**Jong-Jun Hyun, Hyo-Soon Park, Ki-Ho Kim and Hong-Jin Kim**

*College of Pharmacy, Chung Ang University, Seoul, 156-756, Korea*

*(Received Ocotober 15, 1999)*

The UL47 gene (b 60390-b 60338) located in the unique long region of the human cytomegalovirus (HCMV) AD169 strain genome was analyzed by RNA mapping. Northern blot analysis showed that the UL47 gene was expressed at late times after infection (72 h postinfection). The 9.7-kb transcript was expressed in the infected cells but not in phosphonoformate-treated cells at 72 hpi, indicating that the UL47 gene was only expressed at late times after infection. To map the 5'-end and 3'-end of UL47 transcripts, primer extension and RNase protection analysis were performed. Primer extension analysis revealed that the transcription initiation site of UL47 was located in 27 bp downstream (b 60323) of the TATA box motif. The sizes of UL47 ORF (approximately 2.9-kb) and UL48 ORF (approximately 6.7-kb) deduced from computer sequence analysis suggest that the expressed 9.7-kb transcript of UL47 uses the 3'-end polyadenylation signal of UL48. The result of RNase protection determined that the 3'-end of UL47 RNA utilized the 3'-end polyadenylation signal of UL48, which is located in HCMV genome b 70082.

**Key words:** Human cytomegalovirus, UL47, Northern blot, Primer extension, RNase protection

## INTRODUCTION

Human cytomegalovirus (HCMV) causes an asymptomatic infection in normal, healthy adults and then remains in a latent state after primary infection. In immunocompromised hosts, including transplant recipients and AIDS patients, HCMV is often reactivated from the latent state and causes a variety of diseases, such as adrenalitis, retinitis, encephalitis, colitis, hepatitis, pneumonitis, and viremia (Britt and Alford, 1996). Moreover, HCMV causes mental abnormalities (i.e., mental retardation) in some newborns with congenital viral infection. Cytomegaloviruses are highly species-specific agents, therefore HCMV can be propagated *in vitro* only in fibroblasts of human origin (Roizman *et al.*, 1981).

HCMV, a member of the betaherpesvirus group, has a linear double-stranded DNA genome (Chee *et al.*, 1990), which has more than 200 significant open reading frames (ORFs). The gene expression of the HCMV genome after infection of permissive host cells occurs in a temporally regulated pattern and has been divided into immediate-early (IE), early, and late classes (DeMarchi, 1981; Geballe *et al.*, 1986; Jones and Muzithras, 1991; Wathen and Stinski, 1982; Wathen *et al.*, 1981).

Expression of the immediate-early genes requires no de novo protein synthesis (McDonough and Spector, 1983; Stinski *et al.*, 1978; Wathen and Stinski, 1982; Wathen *et al.*, 1981). The protein products of IE genes are required for the regulation of early and late gene expression (Chang *et al.*, 1989; Davis and Huang, 1985; DeMarchi, 1981; Kouzarides *et al.*, 1988; Stenberg *et al.*, 1990; Stenberg *et al.*, 1985; Tenny and Colberg-Poley, 1991; Weston, 1988). Among these, IE1 and IE2 proteins are encoded by IE genes located in the long unique segment of the genome. A single strong promoter activates the transcription of these genes, and is expressed via differential splicing. A 1.95-kb mRNA is transcribed from the IE1 region and two RNAs of 2.25- and 1.7-kb are transcribed from the IE2 region (Kouzarides *et al.*, 1987a; Kouzarides *et al.*, 1987b; Stenberg *et al.*, 1990; Stenberg *et al.*, 1985).

Early genes are expressed before viral DNA replication and regulated by viral IE gene products (DeMarchi *et al.*, 1980; Kim, 1996; McDonough and Spector, 1983; Spector *et al.*, 1990; Staprans and Spector, 1986; Wathen and Stinski, 1982; Wathen *et al.*, 1981). However, the

*Correspondence to*: Hong-Jin Kim, College of Pharmacy, Chung Ang University, 221, Huksuk-Dong, Dongjak-Ku, Seoul, 156-756, Korea
E-mail: hongjink@cau.ac.kr

ways in which the expression of HCMV E genes are regulated by viral IE gene products have not been well understood. Examples of HCMV early genes are the viral DNA polymerase gene and the US11 gene, whose transcription is regulated by the viral IE1 (72 kDa) and IE2 (86 kDa) proteins (Stenberg *et al.*, 1990; Kim, 1996).

Late genes are expressed only after the initiation of viral DNA replication (Chang *et al.*, 1989; Chee *et al.*, 1990; Davis and Huang, 1985; Depto and Stenberg, 1992; Klucher *et al.*, 1989; Lahijani *et al.*, 1991; Leach and Mocarski, 1989). The proteins encoded by late genes are generally structural proteins or enzymes involved in the maturation process of the virus. One example of an HCMV late gene is UL99, which is referred to as the 28-kDa virion phosphoprotein or pp28 (Depto and Stenberg, 1992; Martinez *et al.*, 1989; Martinez and St. Jeor, 1986; Meyer *et al.*, 1988; Pande *et al.*, 1988). However, many parts of these late genes have not yet been well studied.

The UL37 gene of Herpes simplex virus type 1 (HSV-1) is expressed at late times after infection, and is known to encode a virion protein. The HCMV UL47 gene shares sequence homology with the HSV-1 UL37 gene, suggesting that UL47 might also be expressed at late times after infection (Roizman and Sears, 1996; Chee *et al.* 1990). The HCMV UL47 gene is located in the *Hind*III M region of the AD169 strain genome, and the UL47 ORF consists of 982 amino acids (Fig. 1). The expression, regulation and function of the UL47 gene are unknown.

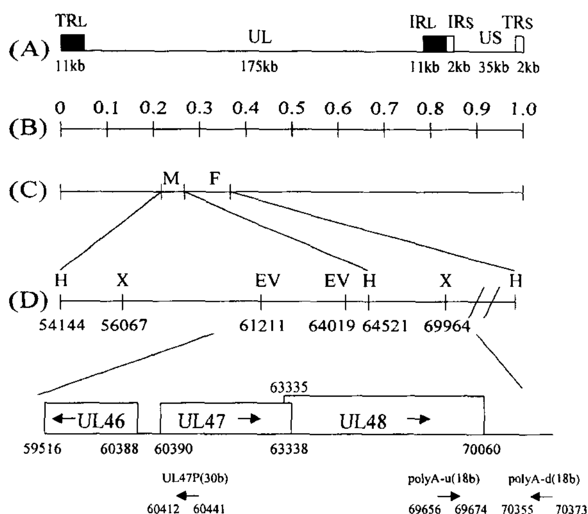



**Fig. 1.** Schematic representation of the HCMV (AD169 strain) genome in standard configuration showing the location of the UL47 open reading frame. (A) Diagram of the HCMV genome, including the unique long (UL) and unique short (US) components. (B) The map units for the arrangement of the viral genome are designated. (C) *Hind*III map of the HCMV (AD169 strain) genome was designated. (D) The expansion region of *Hind*III M and *Hind*III F map was designated (H, *Hind*III; EV, *Eco*RV; X, *Xba*I). Arrangement of the UL46, UL47, UL48 ORFs and oligonucleotides used in this report are shown.

In this study, we analyzed the expression pattern and structure of UL47 transcript during the course of viral infection. Total RNA was harvested from HCMV (AD169 strain)-infected human foreskin fibroblast (HFF) cells at immediate - early, early and late times after infection, and the expression pattern was confirmed by northern blot analysis. In order to map the 5' and 3' ends of the UL47 transcript, primer extension and RNase protection analyses were performed.

## MATERIALS AND METHODS

### Cells and virus

Human foreskin fibroblast (HFF) cells were isolated in this laboratory and used below passage 20. They were maintained in Dulbecco's modified Eagle medium (DMEM ; Gibco) containing 10% fetal bovine serum (Gibco), 25 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid; Gibco), and 100 μg/ml penicillin/streptomycin. HCMV strain AD169 was obtained from the American Type Culture Collection. The virus was propagated according to standard protocols (Stinski *et al.*, 1978).

### DNA sequence

The numbering system proposed by Chee *et al.* (1990) for the HCMV strain AD169 DNA sequence (GenBank accession number X17403) was used in this study.

### Plasmids

In order to make riboprobe for RNase protection analysis, a 717-bp product containing polyadenylation signal (AATAAA) of HCMV UL48 ORF was generated by PCR amplification from pHCMVHindIIIF, using the oligonucleotide primers polyA-u (5'-AATATTCAGTTTCTG-CGT-3') and polyA-d (5'-ACAAGCGGCGACGAAAAAC-3') (Fig. 1D). This product was subcloned into pCR2.1 vector (Invitrogen) to generate plasmid pCR-UL47polyA (4617 bp) using a TA cloning kit (Invitrogen). To construct pGEMUL47polyA (Fig. 4), pCR-UL47polyA and pGEMEX-1 (Promega) were digested with *Eco*RI and ligated. All plasmid DNA manipulations were done according to standard protocols (Sambrook *et al.*, 1989).

### RNA extraction

Uninfected and HCMV-infected HFF cells were disrupted with a guanidine lysis buffer, and total RNA was extracted with acid-phenol (Frederick *et al.*, 1997). Immediate-early, early or late RNA was isolated from HFF cells at 24 h, 48 h or 72 h postinfection (hpi). RNA which is inhibited DNA replication was isolated from the

cells in the presence of 100 μg of phosphonoformate per ml at 72 hpi. RNA concentrations were measured by a spectrophotometer at 260 nm.

### Northern blot analysis

Ten micrograms of total RNA was electrophoresed through 1% formaldehyde agarose gels according to standard protocols (Sambrook *et al.*, 1989). Equal amounts of RNA loaded in each lane of the gel were visualized by ethidium bromide staining. The RNA was transferred overnight to Nytran (Schleicher & Schuell). Blots were prehybridized for at least 4 h at 42°C in a prehybridization buffer [50% formamide, 5 × SSC, 5 × Denhardt solution, 50 mM sodium phosphate pH 6.5, 1% glycine, 100 μg of salmon sperm DNA per ml, 0.1% sodium dodecyl sulfate (SDS)], and hybridized for 16 h at 42°C using a $^{32}$P-labeled DNA probe in a hybridization buffer [50% formamide, 5 × SSC, 1 × Denhardt solution, 20 mM sodium phosphate pH 6.5, 100 μg salmon sperm DNA per ml]. A double-stranded DNA probe was generated from pHCMVHindIIIM by *Bgl*II digestion and the resultant fragment (635 bp) was radiolabeled with [α-$^{32}$P]-dCTP using a random primer and klenow. Blots were washed twice for 5 min each time at room temperature in 2 × SSC-0.1% SDS, and three times at 50°C for 20 min in 0.05 × SSC-0.1% SDS. The washed blots were exposed to Hyperfilm (Amersham) with Hyperscreen™ (Amersham) at -70°C.

### Primer extension analysis

Primer extension was performed previously described (Sambrook *et al.*, 1989). First, a 30-bp oligonucleotide primer, UL47P (5'-GTGCCCGCAGCTGCTCTATCAACTT-TTTGA-3') (Fig 1D) was 5'-end labeled with T4 polynucleotide kinase (New England Biolabs) and [γ-$^{32}$P]-ATP. Twenty micrograms of uninfected RNA or 72 hpi RNA (late RNA) was hybridized with a 5'-end labeled probe for 12 h at 30°C in a 30 μl of hybridization buffer (40 mM PIPES pH 6.4, 1 mM EDTA pH 8.0, 0.4 M NaCl, 80% formamide). 20 μg of yeast tRNA was also used as a negative control. After hybridization, RNA hybrids were precipitated with ethanol and extended in 20 μl of reverse transcriptase buffer [50 mM Tris pH 8.3, 30 mM KCl, 8 mM $MgCl_2$, 10 mM dATP, 10 mM dGTP, 10 mM dTTP, 10 mM dCTP, 10 mM DTT, 20 U of RNasin (Promega), 100 U of Moloney murine leukemia virus reverse transcriptase (New England Biolabs)] for 90 min at 42°C. The reaction was stopped by the addition of RNase A to 50 ng/ml and EDTA to 20 mM, and incubated for 30 min at 37°C. The reaction was then subjected to phenol-chloroform extraction and ethanol precipitation. The primer extension products were analyzed on a 6% polyacrylamide-8 M urea gel along with appropriate dideoxynucleotide sequencing samples as size markers. DNA sequence analysis was done using the dideoxy chain termination method (Sanger *et al.*, 1977). pHCMVHindIII M was used as a template, and the DNA sequencing ladder was extended with the same primer (UL47p). After electrophoresis, the gel was dried and bands were detected by autoradiography.

### RNase protection analysis

All RNase protection analysis was carried out according to standard protocols (Sambrook *et al.*, 1989). A single-stranded RNA probe was generated from linearized plasmid pGEMUL47polyA. A riboprobe was made by using reagents supplied with the Riboprobe kit (Promega) and [α-$^{32}$P]-UTP (10 mCi/ml ; Amersham). The labeled probe was purified by G-50 spin column chromatography (Boehringer Mannheim). It was then hybridized for 16 h at 45°C with 20 μg of total RNA extracted from mock-infected or 72 hpi infected cells in the hybridization buffer (40 mM PIPES pH 6.4, 1 mM EDTA pH 8.0, 0.4 M NaCl, 80% formamide). RNA hybrids were digested for 1 h at 30°C in 300 μl of RNase digestion mixture (300 mM NaCl, 10 mM Tris pH 7.4, 5 mM EDTA pH 7.5, 40 unit/ml RNase T1, and 40 μg/ml RNase A). After a 15 min digestion with proteinase K., phenol-chloroform extraction and ethanol precipitation were performed. The products were analyzed on a 6% polyacrylamide-8 M urea gel. Then, the gel was dried and exposed to X-ray film. The sizes of the protected fragments were estimated in comparison to a radiolabeled size marker. pBR322 cleaved with *Msp*I was 5'-end labeled with [γ-$^{32}$P]-ATP and was then used as a size marker. In this study, the mobilities of the RNA products could be approximately 5 to 15% different from those of the single-stranded DNA markers.

## RESULTS

### Kinetic analysis of the expression of transcripts from UL47

In order to determine the expression pattern and the size of UL47 transcripts, Northern blot analysis was performed with total RNA from HCMV (AD169 strain)-infected HFF cells at immediate-early (24 hpi), early (48 hpi) and late (72 hpi) times after infection (Fig. 2). Blots were hybridized with radiolabelled probes specific to UL47. The Northern blot results indicated that the 9.7-kb transcipt was only expressed at late times after infection (72 hpi). In order to confirm that the UL47 gene is a late gene whose expression requires viral DNA replication, the infected cells were treated with phosphonoformate to inhibit viral DNA replication. Then

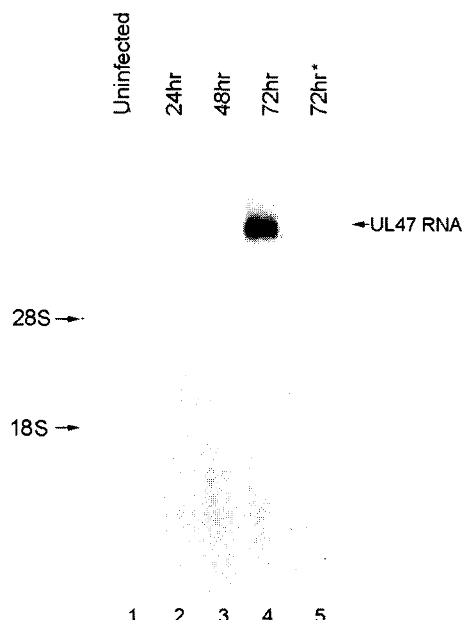



**Fig. 2.** Northern blot analysis of the UL47 transcript. Total RNA extracted from uninfected or infected cells at immediate-early (lane 24 hr), early (lane 48 hr), or late (lane 72 hr) times after infection was used. RNA blots were hybridized with a UL47-specific probe. 72 hr* RNA (lane 72 hr*) contains RNA isolated from cells in the presence of 100 μg of phosphonoformate per ml at late times (72 hr) postinfection. 18S and 28S represent 18S and 28S ribosomal RNA, respectively.

RNAs were harvested at 72 hpi, and Northern blot was performed. No transcript was detected in the phosphonoformate treated cells, indicating that the UL47 gene is indeed a late gene, which is only expressed at late times after infection.

### 5'-end mapping of UL47 transcripts by primer extension

Although the location of the start codon of UL47 gene is deduced to be in b 60390 of HCMV AD169 strain genome from sequence analysis by computer program, primer extension was performed to identify the exact RNA initiation site (5'-end) of UL47 RNA. Uninfected RNA, late RNA (72 hpi), and yeast tRNA were hybridized with a $^{32}$P-end labeled probe (UL47P), and then primer extension was performed (Fig 3). Late RNA hybridized to the probe to form a major RNA-hybrid product at b 60323 (TATA box downstream 27 bp). Therefore, the transcription initiation site of UL 47 was mapped to -67 bp upstream of the start codon of the UL47 ORF, approximately 27 bp downstream from the TATA box promoter motifs located in upstream of the UL47 ORF.

### 3'-end mapping of UL47 transcripts by RNase protection

The results of Northern blot analysis indicated that a DNA probe detected a 9.7-kb transcript of UL47. However, from computer sequence analysis, the size of UL47 ORF is approximately 2.9-kb, and the size of UL48 ORF is approximately 6.7-kb (Fig. 1D). Therefore, we postulated that the UL47 transcript might use the UL48 polyadenylation signal. To test this possibility, we performed RNase protection analysis for the identification of the 3'-end of the UL47 transcript. For the experiment, we constructed pGEMUL47polyA which was inserted about the 700 bp fragment containing the polyadenylation signal (AATAAA sequence) of UL 48 ORF, in order to generate a $^{32}$P-labeled riboprobe complementary to the mRNA transcribed from UL48 and possibly UL47 (Fig. 4). Then, the labeled RNA probes were hybridized to RNA from mock-infected and 72 hpi infected cells, digested with RNase, and analyzed on sequencing gels as described in Materials and Methods. The results showed that the late RNA detected a 432-nt protection product (Fig. 5). As expected, the 432-nt protected RNA indicated that the UL47 transcript utilized the UL48 polyadenylation signal, and that the 3'-end of the UL47 transcript was the polyadenylation signal (AATAAA) region located in b 70082 of the HCMV genome.

## DISCUSSION

In this report, we presented the results of RNA mapping analyses of the UL47 gene (b 60390~63338), located in the unique long region of HCMV AD169 strain genome. Northern blot analysis with RNA harvested from cells at various times of infection showed that the UL47 gene had a 9.7-kb transcript and expressed at late times. Furthermore, we confirmed that the UL47 gene was a late gene, which required viral DNA replication, by showing that no transcript was detected when the infected cells were treated with phosphonoformate.

For the analysis of the structure of UL47 RNA transcripts, primer extension was used to map the 5'-end of UL47 RNA, and RNase protection assay was used to map the 3'-end of UL47 RNA. As a result of primer extension (Fig. 3), the 5'-end of UL47 transcript was mapped 27 bp downstream from a TATA box motif. However, as shown in Fig. 3, the result of dideoxynucleotide sequencing used as a size marker was unclear, and vague minor bands appeared with a distinct major band on the same horizontal lines. This was due to the high GC content of the region for primer extension, as high GC content lowers the efficiency of sequencing.

Northern blot analysis showed a 9.7-kb transcript of the UL47 RNA. However, from computer sequence

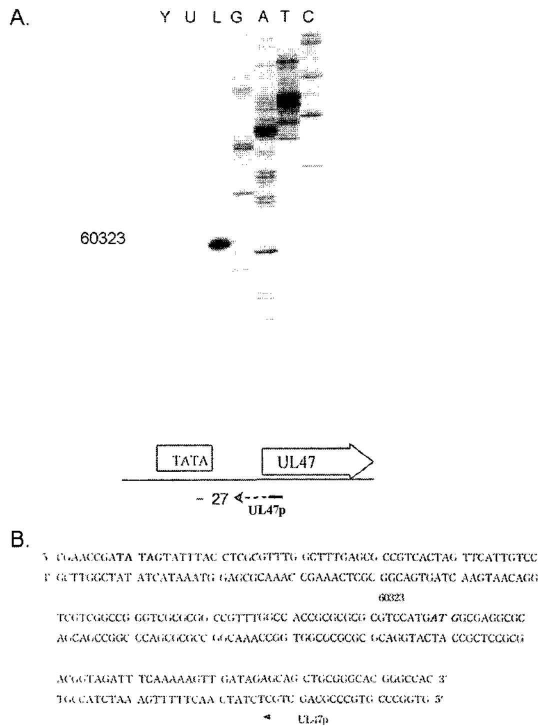



**Fig. 3.** The 5'-end mapping of the UL47 transcript. (A) Primer extension analysis of the UL47 transcript. Total RNA extracted from uninfected (lane U) or infected cells at late (lane L) times after infection was used, except yeast tRNA (lane Y). Analysis of UL47 transcript used oligomers 5'-GTGCCCCGCAGCTGCTCTATCAACTTTTTGA-3' as the primer. A dideoxy sequencing ladder using the respective oligomer is adjacent to the primer extension to determine the exact initiation site of the RNA. (B) Nucleotide sequence of the part of the HCMV AD169 strain genome. The TATA box, start codon, and 5'-end site of UL47 ORF are indicated in the figure. The site of the primer is also indicated in the figure.

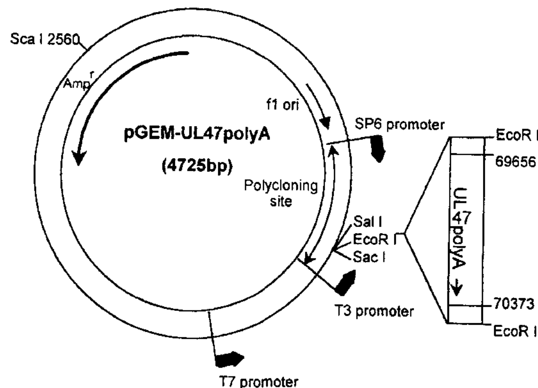
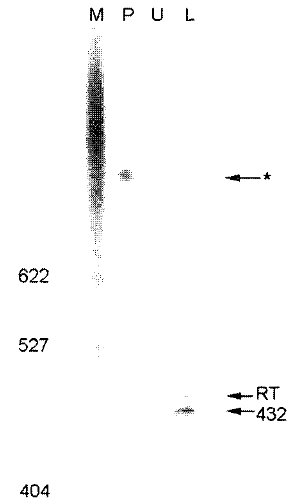



**Fig. 4.** Map of pGEMUL47polyA construct. Approximately 717 bp fragment containing the polyadenylation signal (AATAAA) of HCMV UL48 ORF is inserted into the *Eco*RI site of pG-EMEX-1 (Promega). Detail course to construct pGEMUL47-polyA was described in the Materials and Methods section. To generate polyA riboprobe (about 750 bases) with T3 RNA polymerase for RNase protection, pGEMUL47polyA was linearized with *Sal*I.

analysis, the size of UL47 ORF was deduced to be approximately 2.9-kb and the size of UL48 ORF was deduced to be approximately 6.7-kb. In addition, UL47 ORF overlaps four bases with UL48 ORF at the 3'-terminal end in the same direction (Fig 1D). HCMV ORFs often have 3'-end coterminals. For example, 3'-end mapping of US10 and US11 ORFs identified that the US10 and US11 transcripts both used the 3'-end polyadenylation signal of US10 (Jones and Muzithras, 1991). Moreover, UL47 ORF has no polyadenylation signal. The discrepancy in size deduction and the lack of a polyadenylation signal of UL47 led us to postulate that UL47 transcript uses the polyadenylation signal of UL48. To determine that, the 3'-terminal end of the UL47 transcript was identified by RNase pro-



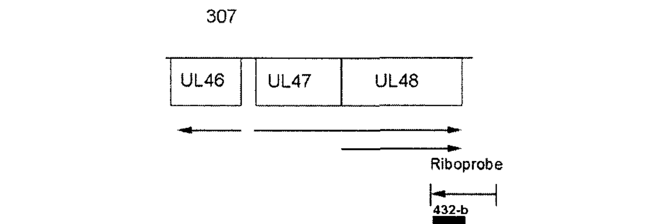




**Fig. 5.** RNase protection mapping of the UL47 transcript. Total RNA extracted from uninfected or infected (lane U) cells at late (lane L; 72 h) times after infection was used. Riboprobe produced from pGEMUL47polyA (Fig. 4) was used in the RNase protection assay to map the 3'-end of UL47. The location of the probe (lane P; *) and the major protected fragment (in lane L; 432 bp) are indicated to the right of the panel by arrows. A protected fragment corresponding to a readthrough transcript (RT) is also indicated in the figure. The molecular weight markers are end-labeled *Msp*I-digested pBR 322 (lane M) DNA fragments.

tection. The result was analyzed by using a riboprobe containing the 3′-end of UL48, and it showed a 432 bp protected fragment, as shown in Fig. 5. This RNase protection result demonstrates that UL47 transcript utilizes the polyadenylation signal of UL48. Therefore, the 3'-end of the UL47 transcript is a polyadenylation signal (AATAAAA) of the HCMV genome b 70082.

As discussed above, we examined the expression of the UL47 gene to study the expression of the HCMV late gene, since the UL47 gene shares sequence homology with the UL37 gene of HSV-1 which is known to be expressed at late times after infection and encodes a virion protein (Roizman and Sears, 1996; Chee *et al.*, 1990). As a result of this experiment, it was confirmed that the UL47 gene is expressed at late times after HCMV infection. Considering the sequence homology and the expression time of the HCMV UL47 gene, this gene resembles the HSV UL37 gene and is highly expected to encode virion protein (Mocarski, 1996). A study of the biological function of UL47 has not been performed as of yet. Whether the UL47 gene encodes a viral structural protein remains unknown.

In our experiment, the fact that UL47 transcript used the 3'-end of UL48 ORF downstream suggests the need for future studies on whether only UL47 ORF containing UL48 ORF is expressed, or whether both UL47 ORF containing UL48 ORF and individual UL48 ORF are expressed, when the two RNAs have 3′-coterminals. The structure and expression of UL48 transcripts are currently under investigation.

## ACKNOWLEDGEMENTS

We would like to thank Dr. Im-Soon Lee (Seoul National University, College of Medicine) for the critical reading of this manuscript. "This research was supported by a Chung Ang University research grant in 1998.

## REFERENCES CITED

Britt, W. J. and Alford, C. A., Cytomegalovirus; *In Fields Virology.* Fields, B. N. (Ed.). Lippincott-Raven, Philadelphia, 3rd Ed., Vol. 2, pp. 2493-2523, 1996.

Chang, C. P., Malone, C. L. and Stinski, M. F., A human cytomegalovirus early gene has three inducible promoters that are regulated differentially at various times after infection. *J. Virol.,* 63, 281-290 (1989).

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A., Kouzarides, T., Martignetti, J. A., Preddie, E., Satchwell, S. C., Tomlinson, P., Weston, K. and Barrell, B. G., Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.,* 154, 125-169 (1990).

Davis, M. G. and Huang, E. S., Nucleotide sequence of a human cytomegalovirus DNA fragment encoding a 67-kilodalton phosphorylated viral protein. *J. Virol.,* 56, 7-11 (1985).

DeMarchi, J. M., Human cytomegalovirus DNA: restriction enzyme cleavage maps and map locations of immediate-early, early, and late RNAs. *Virology,* 114, 23-38 (1981).

DeMarchi, J. M., Schmidt, C. A. and Kaplan, A. S., Patterns of transcription of human cytomegalovirus in permissively infected cells. *J. Virol.,* 35, 227-286 (1980).

Depto, A. S. and Stenberg, R. M., Functional analysis of the true late human cytomegalovirus pp28 upstream promoter: *cis*-acting elements and viral *trans*-acting proteins necessary for promoter activation. *J. Virol.,* 66, 3241-3246 (1992).

Frederick, M. A., Roger, B., Robert, E. K., David, D. M., Seidmon, J. G., John, A. S. and Kevin, S., *Current protocols in molecular biology.* John Wiley & Sons, Inc., 1997.

Geballe, A. P., Spaete, R. R. and Mocarski, E. S., A *cis*-acting element within the 5' leader of a cytomegalovirus beta transcript determines kinetic class. *Cell,* 46, 865-872 (1986).

Jones, T. R. and Muzithras, V. P., Fine mapping of transcripts expressed from the US6 gene family of human cytomegalovirus strain AD169. *J. Virol.,* 65, 2024-2036 (1991).

Kim, H. J., Regulation of the human cytomegalovirus US11 gene expression. *Mol. Cells,* 6, 61-66 (1996).

Klucher, K. M., Rabert, D. K. and Spector, D. H., Sequences in the human cytomegalovirus 2.7 kilobase RNA promoter which mediate its relation as an early gene. *J. Virol.,* 63, 5334-5343 (1989).

Kouzarides, T., Bankier, A. T., Satchwell, S. C., Preddy, E. and Barrell, B. G., An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology,* 165, 151-164 (1988).

Kouzarides, T., Bankier, A. T., Satchwell, S. C., Weston, K., Tomlinson, P. and Barrell, B. G., Large scale rearrangement of homologous regions in the genomes of HCMV and EBV. *Virology,* 157, 397-413 (1987a).

Kouzarides, T., Bankier, A. T., Satchwell, S. C., Weston, K., Tomlinson, P. and Barrell, B. G., Sequence and transcription analysis of the human cytomegalovirus DNA polymerase gene. *J. Virol.,* 61, 125-133 (1987b).

Lahijani, R. S., Otteson, E. W., Adlish, J. D. and St. Jeor, S. C., Characterization of a human cytomegalovirus 1.6 kilobase late mRNA and identification of its putative protein product. *J. Virol.,* 65, 373-381 (1991).

Leach, F. S. and Mocarski, E. S., Regulation of cytomegalovirus late-gene expression: differential use of three start sites in the transcriptional activation of ICP36 gene expression. *J. Virol.,* 63, 1783-1791 (1989).

Martinez, J., Lahijani, R. S. and St. Jeor, S. C., Analysis of a region of the human cytomegalovirus (AD169)

genome coding for a 25-kilodalton virion protein. *J. Virol.,* 63, 233-241 (1989).
Martinez, J. and St. Jeor, S. C., Molecular cloning and analysis of three cDNA clones homologous to human cytomegalovirus RNAs present during late infection. *J. Virol.,* 60, 531-538 (1986).
McDonough, S. H. and Spector, D. H., Transcription in human fibroblasts permissively infected by human cytomegalovirus strain AD169. *Virology,* 125, 31-46 (1983).
Meyer, H., Bankier, A. T., Landini, M. P., Brown, C. M., Barrell, B. G., Ruger, B. and Mach, M., Identification and procaryotic expression of the gene coding for the highly immunogenic 28-kilodalton structural phosphoprotein (pp28) of human cytomegalovirus. *J. Virol.,* 62, 2243-2250 (1988).
Mocarski, JR. E. S., Cytomegaloviruses and Their Replication; *In Fields Virology.* Fields, B. N. (Ed.). Lippincott-Raven, Philadelphia, 3rd Ed., Vol. 2, pp. 2447-2492, 1996.
Pande, H., Campo, K., Churchill, M. A., Clark, B. R. and Zaia, J. A., Genomic localization of the gene encoding a 32-kDa capsid protein of human cytomegalovirus. *Virology,* 167, 306-310 (1988).
Roizman, B., Carmichael, L. E., Deinhardt, F., de-The, G., Nahmias, A. J., Plowright, W., Rapp, F., Sheldrick, P., Takahashi, M. and Wolf, K., Herpesviridae. Definition, provisional nomenclature, and taxonomy. The Herpesvirus Study Group, the International Committee on Taxonomy of Viruses. *Intervirology,* 16, 201-217 (1981).
Roizman, B. and Sears, A. E., Herpes Simplex Viruses and Their Replication; *In Fields Virology.* Fields., B. N. (Ed.). Lippincott-Raven, Philadelphia, 3rd Ed., Vol. 2, pp. 2231-2295, 1996.
Sambrook, J., Fritisch, E. F. and Maniatis, T., *Molecular Cloning : A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.
Sanger, F., Nicklen, S. and Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA,* 74, 5463-5467 (1977).
Spector, D. H., Klucher, K. M., Rabert, D. K. and Wright, D. A., Human cytomegalovirus early gene expression. *Curr. Top. Microbiol. Immunol.,* 154, 21-45 (1990).
Staprans, S. I. and Spector, D. H., 2.2-kilobase class of early transcripts encoded by cell-related sequences in human cytomegalovirus strain AD169. *J. Virol.,* 57, 591-602 (1986).
Stenberg, R. M., Fortney, J., Barlow, S. W., Magrane, B. P., Nelson, J. A. and Ghazal, P., Promoter-specific trans activation and repression by human cytomegalovirus immediate-early proteins involves common and unique protein domains. *J. Virol.,* 64, 1556-1565 (1990).
Stenberg, R. M., Witte, P. R. and Stinski, M. F., Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate-early region 1. *J. Virol.,* 56, 665-675 (1985).
Stinski, M. F., Sequence of protein synthesis in cells infected by human cytomegalovirus: early and late virus-induced polypeptides. *J. Virol.,* 26, 686-701 (1978).
Tenney, D. J. and Colberg-Poley, A. M., Expression of the human cytomegalovirus UL36-38 immediate-early region during permissive infection. *Virology,* 182, 199-210 (1991).
Wathen, M. W. and Stinski, M. F., 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 (1982).
Wathen, M. W., Thomsen, D. R. and Stinski, M. F., Temporal regulation of human cytomegalovirus transcription at immediate early and early times after infection. *J. Virol.,* 38, 446-459 (1981).
Weston, K., An enhancer element in the short unique region of human cytomegalovirus regulates the production of a group of abundant immediate early transcripts. *Virology,* 162, 406-416 (1988).