

## Structure and Expression of Murine Cytomegalovirus Immediate-Early Gene 2

MARTIN MESSERLE,<sup>1</sup> GÜNTHER M. KEIL,<sup>2</sup> AND ULRICH H. KOSZINOWSKI<sup>1\*</sup>

*Department of Virology, Institute for Microbiology, University of Ulm, 7900 Ulm,<sup>1</sup> and Federal Research Centre for Virus Diseases of Animals, 7400 Tübingen,<sup>2</sup> Federal Republic of Germany*

Received 8 August 1990/Accepted 26 November 1990

**The immediate-early gene *ie2* of murine cytomegalovirus was characterized. The 1.75-kb *ie2* transcript is spliced from three exons, of 78, 124, and 1,283 nucleotides, which are separated by introns of 1,245 and 364 nucleotides. An ATG codon located in the third exon leads into an open reading frame of 391 codons. Immediate-early expression of the predicted polypeptide was confirmed by immunoprecipitation of a 43-kDa protein by using an antiserum raised against a bacterial fusion protein. The predicted IE2 amino acid sequence has regions with similarity to amino acid sequences of members of the human cytomegalovirus US22 family.**

Herpesvirus genes are expressed in three sequential phases which are temporally controlled and regulated in a cascade fashion. Immediate-early (IE) genes are the first viral genes which are transcribed upon infection. They are operationally defined as those genes which are transcribed without the need for prior viral protein synthesis. Herpesviruses differ in their numbers of IE genes. In herpes simplex virus, for example, five IE polypeptides are encoded, two of which are essential for replication in cell culture (9, 31). Herpes simplex virus IE proteins modulate early and late gene transcription (6, 11, 26), play a role in establishment and reactivation from latency (21), and may be required for replication in certain cell types (24, 29).

For cytomegalovirus (CMV), the organization of IE genes and the functions of their products during the viral replication cycle are under scrutiny. Some common features of IE genes of CMV from different species have emerged. A complex regulatory sequence containing binding sites for various transcription factors is located upstream of the abundantly expressed IE genes (1, 3, 10, 12, 39). This enhancer sequence regulates the expression of two genes, referred to as IE1 and IE2 in human CMV (HCMV) (34-38) and simian CMV (20) and *ie1* and *ie3* in murine CMV (MCMV) (15, 16). The most thorough analysis of this region has been carried out with HCMV (13, 14, 22, 28, 35). The two genes form a complex transcriptional unit from which mRNAs derived from at least five exons arise by alternative splicing. Whereas IE1 and *ie1* gene products are translated from mRNAs which originate from the first four exons, the IE2 and *ie3* products do not use information encoded within exon 4 (14-16, 37).

A remarkable divergence from the organization of abundantly expressed CMV IE genes is represented by the MCMV *ie2* gene, which has no obvious counterpart in HCMV and simian CMV (4). This gene is located at the other side of the enhancer sequence and is transcribed in the direction opposite to that of *ie1* and *ie3* (16). No role in transcriptional regulation could be assigned to the *ie2* gene product (17), nor is it essential for virus growth in tissues (23). There is, however, evidence for its involvement in the reactivation of latent MCMV in spleen cells (25). Here, we describe the structural organization of the *ie2* gene, the

sequence of the expressed IE2 polypeptide product, and the pattern of *ie2* gene expression.

In MCMV, abundant IE transcription originates from a region of approximately 11 kbp located between map units 0.769 and 0.818 (Fig. 1). Previous data indicated that transcription of the *ie2* gene begins close to map unit 0.802 and ceases between map units 0.814 and 0.818 (16). The DNA sequence between map units 0.802 and 0.818 was determined in order to identify potential open reading frames. Figure 2 provides the nucleotide sequence of the *ie2* gene, starting 18 nucleotides (nt) upstream of the TATA box and ending 38 nt downstream of the potential polyadenylation signal. The exon-intron structure of the *ie2* gene was analyzed by nuclease protection experiments. The experimental procedures used for the structural mapping have been reported previously (15). Briefly, in vitro-transcribed cRNAs or 5'-end-labeled DNA fragments were hybridized to IE RNA isolated from cells which were infected in the presence of cycloheximide (CH). After digestion with either RNases A and T<sub>1</sub> or nuclease S1, the nuclease-resistant fragments were size fractionated on denaturing polyacrylamide gels. The results of the nuclease protection experiments are summarized in Fig. 3.

Following hybridization of a cRNA transcribed from map units 0.818 (*Hind*III) to 0.797 (*Dra*I) to IE RNA and subsequent digestion with RNases A and T<sub>1</sub>, three protected fragments of 78, 112, and about 1,280 nt were detected (Fig. 3, panel IIA, lane 1), indicating that the *ie2* gene comprises three exons and two introns. A cRNA transcribed from map units 0.818 (*Hind*III) to 0.802 (*Hpa*I) (panel IIA, lane 2) protected fragments of 32, 112, and 1,280 nt which located the start point of the *ie2* mRNA close to the *Hpa*I site and identified the 78-nt fragment as the first exon. Sequence analysis had revealed at map unit 0.802 two *Hpa*I sites separated by 78 nt. S1 analysis with the 5'-end-labeled 78-nt *Hpa*I fragment yielded a protected fragment of 46 nt (Fig. 3, panel IIB, lane 1) and mapped the 5' cap site of the *ie2* mRNA 46 nt upstream of the second *Hpa*I site (nt 47 in Fig. 2). The position of the 3' end of the first exon was corroborated with cRNAs transcribed from map units 0.806 (*Pst*I) to 0.797 (*Dra*I) and 0.802 (*Hpa*I). Fragments 78 and 32 nt in length were protected (Fig. 3, panel IIA, lanes 3 and 4). A consensus splice donor site, GTAAGG, was detected at positions 125 to 130 of the nucleic acid sequence (Fig. 2).

A cRNA transcribed from map units 0.818 (*Hind*III) to

\* Corresponding author.

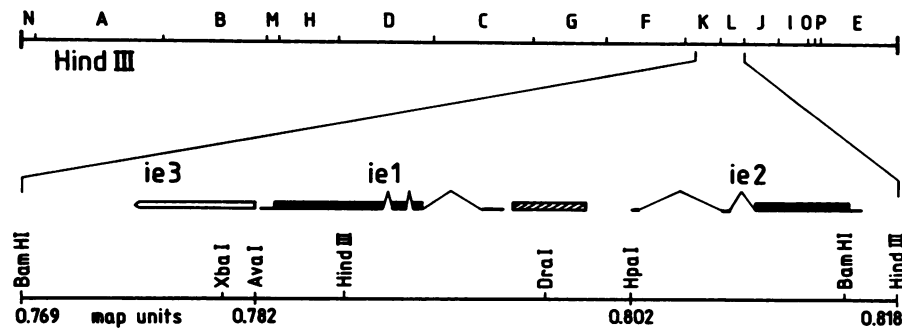


FIG. 1. Physical map of the MCMV IE region. The *Hind*III cleavage map of the MCMV genome is shown. The enlarged segment shows the gene organization of the IE region between map units 0.769 and 0.818. The hatched bar represents the enhancer region. The splicing patterns of the *ie1* and *ie2* genes and the orientation of transcription in the *ie3* region are detailed.

0.813 (*Pst*I) (Fig. 3, panel IIA, lane 5) protected a fragment of 617 nt and placed the largest exon at the 3' end of the transcription unit. The 3' end of the *ie2* mRNA was located 617 nt downstream of the *Pst*I site at position 3129. A consensus polyadenylation signal, AATAAA, was found at positions 3107 to 3112 (Fig. 2). The 5' end of the third exon was located 164 nt upstream of an *Eco*RV site at map unit 0.811 by using a 5'-end-labeled 703-nt *Eco*RV-*Xho*I fragment for S1 analysis (Fig. 3, panel IIB, lane 3). A consensus splice acceptor site was found at positions 1843 to 1859 (Fig. 2).

The second exon, which consisted of about 112 nt, had to be located between nt 125 and 1859 of the sequence. Potential splice acceptor and donor sites were detected at positions 1352 to 1370 and 1495 to 1500, respectively. S1 analysis with a 5'-end-labeled 125-nt *Ava*I fragment revealed a fragment of 79 nt and confirmed the predicted 5' end of the second exon at position 1371 (Fig. 3, panel IIB, lane 2). The location of splice sites predicted the second exon to be of 124 nt. However, the RNase-resistant fragments in Fig. 3, panel IIA, lanes 1 and 2, migrated at a position corresponding to a size of 112 nt. To clarify this discrepancy, a cDNA was prepared and amplified by the polymerase chain reaction. The sequence corresponding to positions 1967 to 1948, located on the noncoding strand in the third exon, and the sequence corresponding to positions 70 to 90, located in exon 1, served as 3' primer (primer 1) and 5' primer, respectively. Primer 1 was hybridized to *ie2* mRNA and elongated with reverse transcriptase. The resulting cDNA was amplified by the polymerase chain reaction in 35 cycles, by established procedures (32). A DNA product of about 300 nt was identified by agarose gel electrophoresis. This DNA fragment was isolated from the gel and directly sequenced by the chain termination method with primer 1 and polymerase T7. This analysis confirmed the location of the expected splice donor and acceptor sites (data not shown). The protected 124-nt cRNA fragment may have migrated with a higher mobility because of its propensity to form a stem loop.

In order to analyze the expression kinetics of the *ie2* mRNA, whole-cell RNA was isolated at various times post-infection (p.i.) and Northern (RNA) blot hybridization was performed by methods detailed previously (16) (Fig. 4). The 1.75-kb size of the polyadenylated *ie2* mRNA is in good agreement with the structural analysis of the gene that should give rise to an RNA of about 1.5 kb prior to polyadenylation. The *ie2* mRNA became detectable 1 h after infection. In contrast to the *ie1/ie3* transcription unit, which is heterogeneous with regard to size and abundance of

individual transcripts, only one major mRNA originates from transcription unit *ie2* and there is no evidence for differentially spliced products. Expression of *ie2* mRNA was maximal at 2 h after infection (Fig. 4, lane 3) and declined rapidly thereafter. Between 6 and 12 h after infection, only small amounts of *ie2* mRNA were detectable (Fig. 4, lanes 5 and 6). At later times in the replication cycle (16 and 20 h p.i.), reexpression of *ie2* was seen (Fig. 4, lane 7). Reexpression required DNA replication, as it did not occur in the presence of the inhibitor phosphonoacetic acid (Fig. 4, lane 8). At IE times, the transcription of *ie2* mRNA was enhanced by the addition of CH (Fig. 4, lane 2). These expression kinetics are comparable to those of *ie1* mRNA, which is detectable 1 h p.i. and which has its maximum of expression at between 2 and 3 h p.i. (16). Also, *ie1* mRNA is reexpressed at late times in a phosphonoacetic acid-sensitive fashion (23, 30).

The open reading frame starting from the first ATG codon downstream of the 5' cap site stops after only 13 codons. The second ATG, which is located at position 1862 in the third exon, leads into an open reading frame of 391 codons. Translation of the *ie2* mRNA probably begins in exon 3 because this start codon fits the consensus sequence for initiation codons (A/G)NNATGG (19) better than the ATG in exon 2. The TGA stop codon of the reading frame in exon 3 is followed by the polyadenylation signal 69 nt downstream. The predicted protein of 391 amino acids has a calculated molecular mass of 43.8 kDa.

No IE2 protein of expected size could be identified following immunoprecipitation of infected cell lysates with murine antisera after MCMV infection (16). In order to confirm the expression of the predicted IE2 protein, a specific antiserum was produced. A 638-bp *Pvu*II-*Bam*HI DNA fragment (map units 0.810 to 0.814), representing 54% of the *ie2* open reading frame, was cloned into the expression vector pATH10 (8). A 63-kDa *trpE*-IE2 fusion protein was isolated from *Escherichia coli* C600 (33) and used to immunize BALB/c mice. Murine embryonic fibroblasts (MEF) were labeled with [<sup>35</sup>S]methionine at various times following infection with MCMV, and proteins were immunoprecipitated from cell lysates by using the antiserum to the fusion protein. A polypeptide with an apparent molecular mass of 43 kDa, which is in good agreement with the calculated molecular mass of the deduced amino acid sequence, was detected (Fig. 5). A second band of minor abundance which might represent a modification product of the 43-kDa IE2 protein was detected at 41 kDa. The IE2 proteins were produced predominantly between 0.5 and 3 h after infection

aacggggcga gttaaccaTA TAAttgctgt ccccatgcc attcagccc agarcagac gaccgcgc tcgatacag 80  
ctactctagc ttaacgagcc ggtgaacga gcccaccag cccggtaagg ttgattggg ctgctgtagt ggacggggat 160  
 ttggcccggt accttaatat gggtggggtc ttatgggtag ggggcttatt aatataacgc cccctacca tacatattgc 240  
 gtcactgaat tagaagcccc ctaccagccc attgcgccc gcccatacgc gtactttcgg ttctatacaa gatggctacc 320  
 gcggttttcg atcgctccc gggtctctat ttctatagct ggaaagcacc cgacccccat gatctgtgcc tttagaccca 400  
gggtactgac ctggataggt gggcagcccc tagcgctgce ttgtattttt atatcagaga tcgaaaatac cagccgcccg 480  
 taacagctgt cagggtata caagatggcg accgctactt tcgactgctt ttcttccctc gcatctgct tcgataaacg 560  
 aacccacccc cggggtatgt gccttagacc ctatagatgc cccccaggac gcaggaaaaa ttttagctcc ctctataatc 640  
 tttattcgag cgatgaaat ccgaggacc cccgggtttt tactttcgtt ttctcggtcg ccgctcgctt cggattccga 720  
 gcagaaaaag ggctagaaa ctgcctcac ctgcgaaac gcgtttgtcg gatcatccc ggaggccag cgtagtctcc 800  
 gttgctgggt tcggccgagg gatcagtcg agcggggcgc gggacgcgg gaacgagaga ggtacggag cggggccccc 880  
gggtgctcgg agcggccgg gagacgctct gcagagctcc agagctgctg gtccctgacac tgaacagttg gaaaaatac 960  
 tggatcttat aatgaccag ggaaggggt ccgagatcct cccagtgctc gagactgcca agagacactc ctgttgaaaa 1040  
 gcagctetaa aacgttttcc tgctcatctc tgccagaccc ccgccactg acagagctca attgatagca gatgtctctc 1120  
 ttggaaggac acacgcttcc acaaacacc ttcctcactg aggtgcatg ctcggggagg aagctcact ttctgttact 1200  
 tgtggacggt gtcggcccca ggctcggac gacctgtccc cgggttctct tcttctgat ccaggtagat tgactgtcgg 1280  
 acatcgcgaa cgaccgcagt cttcctgact tttgaatgac tcgagtacat gagcgacct ttcgagactg atctctcttg 1360  
 ttctccatag gtcagctacg cccctgctcc ggactcctca ctgagatacc ttttttcttc agactgacac cgctagagga 1440  
cttccctcgg agatgagga caccggaca ggtgctgctg acctgttgtt actagtgtg acagtggctg gctgacgtgt 1520  
 cccctgggat cttatacaat agtcttctgt tcctgacaga atagcattgt gggagattgt caagtgacc taggtctcac 1600  
 aggtgagata aagctcagag cagagctccc gagcacggag cagtaagac ggcgtgttc tctgcacgga tgctggtgac 1680  
 cggctatttc ttggtaagg gactcttctc ctcgcgcccc cgcggctgct ttcgtctgct cgtctgacct attctctctt 1760  
 ctgattcgag cacggagctc gggctgacg ctgtagttag agtctctcgg gtccggggcg gtagggagcg cgactgtcca 1840  
 cggttctgtt tgctgtaga tATGGAGCGT GTTCGGGGAG CTTCTCCTC AGACTCGGTG GGGACCCCTC ACGGCGATCA 1920  
CAACTCACG TGCATGGCGA HE ER VR GA SS SS DS SV AT LH GD HA 2000  
GGACCCCGC GACCGTGGC GATATCTTA TGCAGCGGT CTGTACAGAC ATCAAGCAGC TGCTCGGGCC GCTCGAATGT  
DPA TV D I F H Q R V C H D I N D V L R P V C G 2080  
CTCAGGACT TCCCGAATA CAGAAGAACC CTGAGGGTGG COGTGGAGGC CGCCACCGAA CTCCACGGGG TCTCGCTCT  
L T S F P E Y R R T L R V A V E A A T E L H G V S A L 2240  
GTCCGGCCAG CTGACCAAGT TCTTGGAGGA CCACGATGGA GGATGCTCC CCCTGGCCCTG GC0000GAAG AGGTACCTTC  
S C Q L T K F L E D H D G A H L P L A V F P N R Y L R 2320  
GCCTGACGTC GGCAGAAAG CTGGTTATC TGCAGCTGC GGAGAAGCTG CGCAAGGAAT ACCCCATGTG CAACGGTCCC  
L T S G R K L C Y L D V A E K L R K E Y F H C N G F 2400  
CTGAGGTTT TCGGGTGAT CCTCCGACCC GGACAGTCCC TCTACAGGGA CCAGACCCCC GTCTCTTGA TGGCCCTCAC  
L E V L G V I L G P Q S L Y S D Q H P V L H G L T 2480  
CGGACCATC TTCTCCACG CCAGAGGGAG GCCCGTCTGG AGTCCCGACT ACGATCCCGA CGGGAGGGCC GATCGACTCT  
G T I E L H A R G R F W S P D Y D P E R D A D R L F 2560  
TCTGGCCGC CGAGTCCCTG CAGTCTTCG GTAGGGAAGG ACTCTGCAGG TGCAGCAAGG TCTACACCGA GGACGGGGCC  
L E V L G V I L G P Q S L Y S D Q H P V L H G L T 2640  
GTCTCGTAC CTACTCCCGA GGACCCCGTG CTCAAGAACA TGTGTCTAC GGCCGATCTC GGCCGGAAGG CGCTCCACCA  
A P V A T P E D P V L K N I V F T P H L G C K A L D S 2720  
GCAGATCTGT AAGATCAAGG GACACACCTG GTACTGTGAC GGTGGCCCGC GCATGCTGAA GGACAGGGTG TTCGTCGCA  
Q I C K I K C H T Y L N G C F C H L K D E V E V A T 2800  
CGCCGGTACT CCCCCCTTTC GTCCGCGCAG TGAACCTGGA ACTCTTTTGA CATAGTITTC TACCCATCGG TAGGGTGAAC  
P D I F P F V P H V M L E L F G H R F L P I C N G T 2880  
CGTTC0000 AGGATCCCGA GTGTGAGATG TTTATCATGG TGAGCCCGG GGGGGCCATC TACGGGCACA TGTTAGATAC  
R S P E D P E C H E F I N V D A C G A I Y G H L D S 2960  
CGGAAAGTGT AGGAGGCTAG CTGATAACTT TCACCATGTC ATCGGATGG GCACCCGGAG GCTTATTICA ATTCCAGCT  
K V R R L A D N F D Q F M R H G T R R V L S I S R W 3040  
GTTAAGGGA GTAAGTCTCA ACACCGAGTA CGACGAGACC TGTTTGGGA CAAAGGACCA CGGGTCTTCT CTCTTgacca  
L R E V V S T P S T R R P R L G P R T T G S F T 3120  
agacctggt gacctcagg agaagattc arttagagga cacagatgt gtcatggttt ttaatcaata ataaagagg  
ttttattcat cggacagctg ttgtagcctg 3150

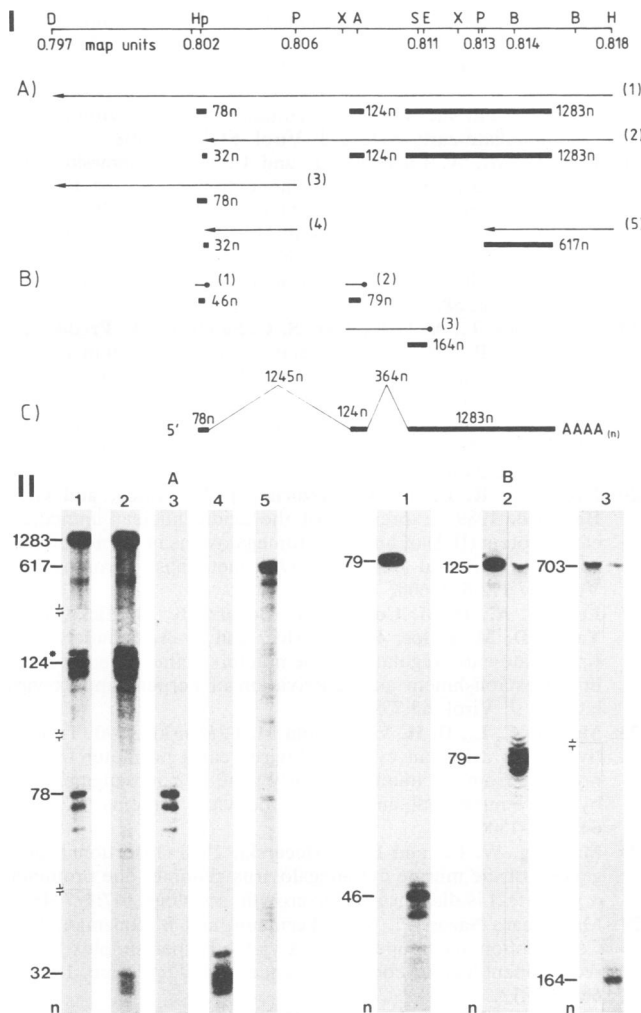
FIG. 2. Nucleotide sequence of the ie2 gene. The nucleotide sequence is shown from 18 nt upstream of the TATA box to 38 nt downstream of the polyadenylation signal. Exons are underlined, and the coding sequence, the TATA box, and the polyadenylation signal are in capital letters. The deduced amino acid sequence of the IE2 protein (one-letter code) is indicated below the open reading frame.

(Fig. 5, lane 3). During the early phase of infection (4 to 8 h p.i.), only traces of the proteins were detectable (data not shown). Restriction of expression to IE gene products was achieved by infection in the presence of CH, followed 3 h later by incubation with dactinomycin for a further 3 h. This resulted in an enhanced amount of IE2 proteins (Fig. 5, lane 2). Reexpression of the IE2 proteins at late times was not detected by using the antiserum. It is probably the low affinity and the low titer of the antiserum which prevented the detection. Precipitation following cell fractionation suggested that the IE2 protein is located in both the cytoplasm and the nucleus (data not shown).

The analysis of the amino acid sequence revealed no obvious indication for the function of the protein. It has no acidic or basic domains, no glycosylation sites, and no signals for a secretory pathway or for nuclear transport. The high overall content of hydrophobic amino acids is remarkable (39%).

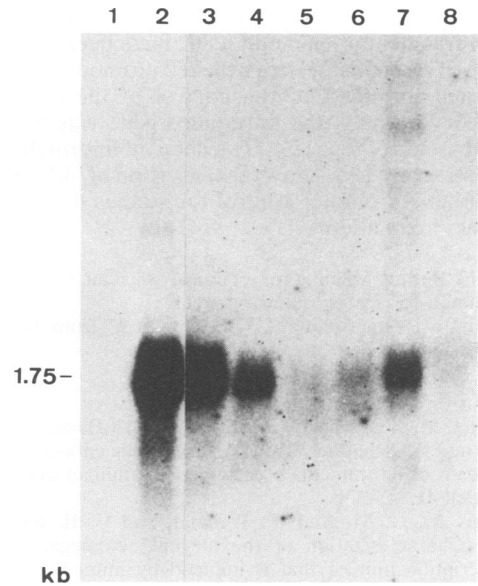
Comparison of the IE2 amino acid sequence and the

NBRF protein data base (release 23, December 1989) with the sequence analysis program FastA (7, 27) revealed a low similarity to the HCMV US22 family protein HCMVUS24 (FastA score, 90). The 95 carboxy-terminal amino acids of the IE2 protein show 24.2% identity to amino acids 235 to 330 of HCMVUS24. A detailed comparison with the HCMV family 4 proteins (40), which form part of the US22 family (4), detected a low similarity of the IE2 protein to other members of this family as well. For example, 20.4% of the amino acids at the amino terminus are shared between the IE2 protein and HCMVUS26. However, the degree of homology between members of family 4 is much higher. One member of the US22 family represents a putative IE membrane glycoprotein, and three sequence motifs are shared between this putative IE glycoprotein and the proteins of family 4 (18). Whereas motifs 2 and 3 show 30% identity to amino acids 198 to 230 and 325 to 355 of the IE2 protein, no counterpart to motif 1 was found in the IE2 protein, which makes a firm association of ie2 with members of this gene



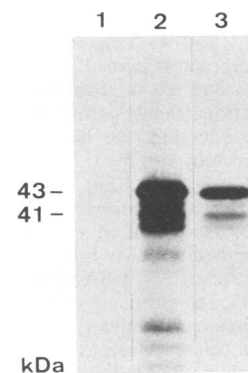
**FIG. 3.** Structural analysis of transcription unit *ie2*. Nuclease protection experiments were performed to analyze the exon-intron structure of the gene. (I) Design of the experiments; (II) experimental data. (I) The positions of some restriction enzyme sites are indicated. Abbreviations: A, *Ava*I; B, *Bam*HI; D, *Dra*I; E, *Eco*RV; H, *Hind*III; Hp, *Hpa*I; P, *Pst*I; S, *Sma*I; X, *Xho*I. (A and B) Experimental protocol. Evenly labeled cRNAs (←) and 5'-end-labeled DNA fragments (→) were hybridized to IE RNA and then were digested with RNases A and T<sub>1</sub> and nuclease S1, respectively. Solid bars and numbers indicate the locations and lengths, respectively, of the protected fragments. (C) Summary of the results. (II) Autoradiographs of protected fragments, which were separated by electrophoresis in denaturing polyacrylamide gels. To present all relevant information in one figure, the autoradiographs were cropped and refitted as indicated. (±) Size markers were obtained from commercial suppliers or prepared by cleaving pBR322 with *Hpa*II and labeling the fragments at the 5' ends. (A) Protected fragments from the cRNAs after digestion with RNases A and T<sub>1</sub>. \*, Migration of protected fragments according to a length of 112 nt. (B) In each pair of lanes, the left lane shows the 5'-end-labeled fragments prior to digestion with nuclease S1 and the right lane reveals the nuclease-resistant fragments.

family questionable. The low degree of similarity between *ie2* and HCMV genes is not unexpected. The homology between IE1 of HCMV and *ie1* of MCMV is apparent from the number and the size of exons and introns, the organization of coding and noncoding exons, the regulation of



**FIG. 4.** Expression kinetics of the *ie2* mRNA. Northern blot analysis with an *ie2*-specific probe was performed to determine the kinetics of the synthesis of the *ie2* mRNA. MEF were mock infected (lane 1) or were infected in the presence of CH (lane 2) or phosphonoacetic acid (lane 8) or without inhibitors (lanes 3 through 7). Whole-cell RNA was isolated at 2 (lanes 2 and 3), 3 (lane 4), 8 (lane 5), 12 (lane 6), and 24 (lanes 7 and 8) h p.i. Exposure time of lane 2 was three times shorter than for the other lanes.

transcription, and the functional homology but not, however, from the nucleic acid and amino acid sequences. On the other hand, this does not mean that there is no sufficient relationship in general between amino acid sequences of HCMV and MCMV gene products to identify homologs. The *gB* gene of MCMV could be clearly identified by its similarity to its HCMV counterpart (5, 14a), and an HCMV homolog to the MCMV *e1* gene has also been identified (2). On the basis of our preliminary sequence information,



**FIG. 5.** Synthesis of the IE2 protein. Lysates of infected cells labeled with [<sup>35</sup>S]methionine were immunoprecipitated with the antiserum directed against the fusion protein and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Lanes: 1, mock-infected MEF labeled with [<sup>35</sup>S]methionine for 3 h; 2, MCMV-infected MEF incubated from 0 to 3 h p.i. with CH and from 3 to 6 h p.i. with dactinomycin and labeled with [<sup>35</sup>S]methionine during the dactinomycin block; 3, MCMV-infected MEF incubated with [<sup>35</sup>S]methionine from 0.5 to 3 h p.i.

MCMV recombinants in which foreign sequences replace the *HpaI* fragment at map unit 0.802 have been constructed. This type of insertion destroys the *ie2* promoter and subjects the inserted promoter to the control of the enhancer sequence. Accordingly, the introduced gene was found to be regulated as an IE gene (23). The detailed information on the *ie2* gene structure now allows the insertion of indicator genes under variable promoter control for studies on cell tropism and latency regulation.

We thank Bettina Mächtig for technical assistance. We also thank Ingrid Bennett for secretarial assistance.

This work was supported by SFB 322 grant A7 from the Deutsche Forschungsgemeinschaft.

#### REFERENCES

- Boshart, M., F. Weber, G. Jahn, K. Dorsch-Häsler, B. Fleckenstein, and W. Schaffner. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* **41**:521-530.
- Bühler, B., G. M. Keil, F. Weiland, and U. H. Koszinowski. 1990. Characterization of the murine cytomegalovirus early transcription unit *e1* that is induced by immediate-early proteins. *J. Virol.* **64**:1907-1919.
- Chang, Y.-N., S. Crawford, J. Stall, D. R. Rawlins, K.-T. Jeang, and G. S. Hayward. 1990. The palindromic series I repeats in the simian cytomegalovirus major immediate-early promoter behave as both strong basal enhancers and cyclic AMP response elements. *J. Virol.* **64**:264-277.
- 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. Tomlinson, K. M. Weston, and B. G. Barrell. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD 169. *Curr. Top. Microbiol. Immunol.* **154**:125-169.
- Cranage, M. P., T. Kouzarides, A. T. Bankier, S. Satchwell, K. Weston, P. Tomlinson, B. Barrell, H. Hart, S. E. Bell, A. C. Minson, and G. L. Smith. 1986. Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. *EMBO J.* **5**:3057-3063.
- DeLuca, N. A., and P. A. Schaffer. 1985. Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. *Mol. Cell. Biol.* **5**:1997-2008.
- Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the Vax. *Nucleic Acids Res.* **12**:387-395.
- Dieckmann, C. L., and A. Tzagoloff. 1985. Assembly of the mitochondrial membrane system. *J. Biol. Chem.* **260**:1513-1520.
- Dixon, R. A. F., and P. A. Schaffer. 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. *J. Virol.* **36**:189-203.
- Dorsch-Häsler, K., G. M. Keil, F. Weber, M. Jasin, W. Schaffner, and U. H. Koszinowski. 1985. A long and complex enhancer activates transcription of the gene coding for the highly abundant immediate early mRNA in murine cytomegalovirus. *Proc. Natl. Acad. Sci. USA* **82**:8325-8329.
- Everett, R. D. 1986. The products of herpes simplex virus type 1 (HSV-1) immediate-early genes 1, 2 and 3 can activate HSV-1 gene expression in trans. *J. Gen. Virol.* **68**:2507-2513.
- Ghazal, P., H. Lubon, B. Fleckenstein, and L. Hennighausen. 1987. Binding of transcription factors and creation of a large nucleoprotein complex on the human cytomegalovirus enhancer. *Proc. Natl. Acad. Sci. USA* **84**:3658-3662.
- Hermiston, T. W., C. L. Malone, and M. F. Stinski. 1990. Human cytomegalovirus immediate-early two protein region involved in negative regulation of the major immediate-early promoter. *J. Virol.* **64**:3532-3536.
- Hermiston, T. W., C. L. Malone, P. R. Witte, and M. F. Stinski. 1987. Identification and characterization of the human cytomegalovirus immediate-early region 2 gene that stimulates gene expression from an inducible promoter. *J. Virol.* **61**:3214-3221.
- Keil, G. M., et al. Unpublished data.
- Keil, G. M., A. Ebeling-Keil, and U. H. Koszinowski. 1987. Sequence and structural organization of murine cytomegalovirus immediate-early gene 1. *J. Virol.* **61**:1901-1908.
- Keil, G. M., A. Ebeling-Keil, and U. H. Koszinowski. 1987. Immediate-early genes of murine cytomegalovirus: location, transcripts, and translation products. *J. Virol.* **61**:526-533.
- Koszinowski, U. H., G. M. Keil, H. Volkmer, M. R. Fibi, A. Ebeling-Keil, and K. Münch. 1986. The 89,000-*M<sub>r</sub>* murine cytomegalovirus immediate-early protein activates gene transcription. *J. Virol.* **58**:59-66.
- Kouzarides, T., A. T. Bankier, S. C. Satchwell, E. Preddie, and B. G. Barrell. 1988. An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology* **165**:151-164.
- Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**:857-872.
- LaFemina, R. L., M. C. Pizzorno, J. D. Mosca, and G. S. Hayward. 1989. Expression of the acidic nuclear immediate-early protein (IE1) of human cytomegalovirus in stable cell lines and its preferential association with metaphase chromosomes. *Virology* **172**:584-600.
- Leib, D. A., D. M. Coen, C. L. Bogard, K. A. Hicks, D. R. Yager, D. M. Knipe, K. L. Tyler, and P. A. Schaffer. 1989. Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency. *J. Virol.* **63**:759-768.
- Malone, C. L., D. H. Vesole, and M. F. Stinski. 1990. Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral proteins. *J. Virol.* **64**:1498-1506.
- Manning, W. C., and E. S. Mocarski. 1988. Insertional mutagenesis of the murine cytomegalovirus genome: one prominent  $\alpha$  gene (*ie2*) is dispensable for growth. *Virology* **167**:477-484.
- Mavromara-Nazos, P., M. Ackermann, and B. Roizman. 1986. Construction and properties of a viable herpes simplex virus 1 recombinant lacking coding sequences of the  $\alpha 47$  gene. *J. Virol.* **60**:807-812.
- Mocarski, E. S., G. B. Abenes, W. C. Manning, L. C. Sambucetti, and J. M. Cherrington. 1990. Molecular genetic analysis of cytomegalovirus and gene regulation in growth, persistence and latency. *Curr. Top. Microbiol. Immunol.* **154**:47-74.
- O'Hare, P., and G. S. Hayward. 1985. Evidence for a direct role for both the 175,000- and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the transactivation of delayed-early promoters. *J. Virol.* **53**:751-760.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444-2448.
- Pizzorno, M. C., P. O'Hare, L. Sha, R. L. LaFemina, and G. S. Hayward. 1988. *trans*-Activation and autoregulation of gene expression by the immediate-early region 2 gene products of human cytomegalovirus. *J. Virol.* **62**:1167-1179.
- Post, L. E., and B. Roizman. 1981. A generalized technique for deletion of specific genes in large genomes:  $\alpha$  gene 22 of herpes simplex virus 1 is not essential for growth. *Cell* **25**:227-232.
- Reddehase, M. J., M. R. Fibi, G. M. Keil, and U. H. Koszinowski. 1986. Late-phase expression of a murine cytomegalovirus immediate-early antigen recognized by cytolytic T lymphocytes. *J. Virol.* **60**:1125-1129.
- Sacks, W. R., C. C. Greene, D. P. Aschman, and P. A. Schaffer. 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. *J. Virol.* **55**:796-805.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487-491.
- Spindler, K. R., D. S. E. Rosser, and A. J. Berk. 1984. Analysis of adenovirus transforming proteins from early regions 1A and

- 1B with antisera to inducible fusion antigens produced in *Escherichia coli*. *J. Virol.* **49**:132-141.
34. **Stenberg, R. M., A. S. Depto, J. Fortney, and J. A. Nelson.** 1989. Regulated expression of early and late RNAs and proteins from the human cytomegalovirus immediate-early gene region. *J. Virol.* **63**:2699-2708.
35. **Stenberg, R. M., J. Fortney, S. W. Barlow, B. P. Magrane, J. A. Nelson, and P. Ghazal.** 1990. Promoter-specific *trans* activation and repression by human cytomegalovirus immediate-early proteins involves common and unique protein domains. *J. Virol.* **64**:1556-1565.
36. **Stenberg, R. M., D. R. Thomsen, and M. F. Stinski.** 1984. Structural analysis of the major immediate early gene of human cytomegalovirus. *J. Virol.* **49**:190-199.
37. **Stenberg, R. M., P. R. Witte, and M. F. Stinski.** 1985. 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.
38. **Stinski, M. F., D. R. Thomsen, R. M. Stenberg, and L. C. Goldstein.** 1983. Organization and expression of the immediate early genes of human cytomegalovirus. *J. Virol.* **46**:1-14.
39. **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. USA* **81**:659-663.
40. **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.