

## Structural Organization, Expression, and Functional Characterization of the Murine Cytomegalovirus Immediate-Early Gene 3

MARTIN MESSERLE,<sup>1</sup> BRIGITTE BÜHLER,<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> Germany

Received 14 August 1991/Accepted 26 September 1991

We have previously defined *ie3* as a coding region located downstream of the *ie1* gene which gives rise to a 2.75-kb immediate-early (IE) transcript. Here we describe the structural organization of the *ie3* gene, the amino acid sequence of the gene product, and some of the functional properties of the protein. The 2.75-kb *ie3* mRNA is generated by splicing and is composed of four exons. The first three exons, of 300, 111, and 191 nucleotides (nt), are shared with the *ie1* mRNA and are spliced to exon 5, which is located downstream of the fourth exon used by the *ie1* mRNA. Exon 5 starts 28 nt downstream of the 3' end of the *ie1* mRNA and has a length of 1,701 nt. The IE3 protein contains 611 amino acids, the first 99 of which are shared with the *ie1* product pp89. The IE3 protein expressed at IE times has a relative mobility of 88 kDa in gels, and a mobility shift to 90 kDa during the early phase is indicative of posttranslational modification. Sequence comparison reveals significant homology of the exon 5-encoded amino acid sequence with the respective sequence of UL 122, a component of the IE1-IE2 complex of human cytomegalovirus (HCMV). This homology is also apparent at the functional level. The IE3 protein is a strong transcriptional activator of the murine cytomegalovirus (MCMV) *e1* promoter and shows an autoregulatory function by repression of the MCMV *ie1/ie3* promoter. The high degree of conservation between the MCMV *ie3* and HCMV IE2 genes and their products with regard to gene structure, amino acid sequence, and protein functions suggests that these genes play a comparable role in the transcriptional control of the two cytomegaloviruses.

Murine cytomegalovirus (MCMV) is a member of the highly diverse group of herpesviruses. As with other herpesviruses, the expression of MCMV genes occurs in a regulated fashion (21). The immediate-early (IE) genes are the first genes to be expressed after infection. Transcription of IE genes does not require de novo protein synthesis. IE proteins have regulatory functions that are required for the coordinated expression of early and late genes.

The genomic region of MCMV, which is abundantly transcribed at IE times, is located between map units 0.769 and 0.818 of the genome, and its structural organization is similar to that of the IE regions of human and simian cytomegaloviruses (HCMV and SCMV, respectively) (19, 22, 23, 45, 46). Transcription of the MCMV IE genes is under the control of a large and complex enhancer which controls promoters flanking this sequence on both sides (13). Three IE genes have been identified within the MCMV IE region (22, 23, 35). The most abundantly transcribed IE gene of MCMV, *ie1*, is, like the homologous genes of HCMV and SCMV, composed of a first noncoding exon, two small exons, and a large fourth exon (19, 22, 45). The abundantly expressed IE protein encoded by the MCMV *ie1* gene has a molecular mass of 89 kDa (22, 24, 50). Like its counterparts from HCMV and SCMV, it is a phosphorylated nuclear protein with a high percentage of acidic residues (19, 24, 45). Remarkably, however, there is very little homology between the nucleic acid sequences and the amino acid sequences. A unique feature of MCMV is the location of a second IE transcription unit, *ie2*, on the other side of the enhancer sequence (23, 35). The function of the 43-kDa IE2 protein is

unknown (25, 35), and the *ie2* gene seems to be nonessential for growth of the virus in cell culture (32).

Downstream of the abundantly expressed IE gene is another transcribed region which is referred to as IE2 in HCMV and SCMV and *ie3* in MCMV (18, 23, 38, 46). In HCMV and SCMV, the major IE2 transcript is spliced together from the first three exons of the IE1 gene and a fifth exon in the IE2 region (18, 43, 46). This HCMV IE2 transcript is translated into a 579-amino-acid polypeptide with an apparent molecular mass of 88 kDa (30, 43, 44).

The most thorough analysis of the regulatory functions of the IE proteins has been performed with HCMV (18, 30, 37, 38, 44). Transient transfection experiments using chloramphenicol acetyltransferase (CAT) indicator plasmids with various heterologous as well as homologous promoters revealed that the HCMV IE2 protein is a strong activator of early promoters. In addition, the HCMV IE2 protein represses the major IE promoter of HCMV, suggesting an autoregulative role for IE2 (6, 17, 28, 37). The HCMV IE1 protein seems to cooperate with the IE2 protein in the regulation of some promoters (4), but IE1 alone is not sufficient for the activation of early promoters (30, 44).

We have suggested previously that the MCMV *ie3* gene may encode a regulatory protein with similarities to the HCMV IE2 protein (23). In this communication, we report on the structural analysis of the *ie3* gene, the determination of the *ie3* open reading frame, and the identification of the IE3 protein. Partial homology between the amino acid sequences of the MCMV IE3 and the HCMV IE2 proteins suggests a correlation between protein structure and function. Functional assays using the MCMV early promoter *e1* (3) and the MCMV *ie1/ie3* promoter show that the MCMV

\* Corresponding author.

IE3 protein has regulatory properties comparable to those of the HCMV IE2 protein.

## MATERIALS AND METHODS

**Virus and cell culture.** MCMV (mouse salivary gland virus strain Smith; ATCC VR-194) was propagated on BALB/c mouse embryonal fibroblasts (MEF) as described previously (21). Recombinant vaccinia viruses were produced by following established procedures (29) and using the recombination vector pCS43 (1), vaccinia virus DNA of strain Copenhagen, and the temperature-sensitive mutant *ts7* (14). For recombination of the *ie3* open reading frame into the vaccinia virus genome, the continuous *ie3* coding sequence was constructed. A 717-bp amplified DNA fragment (see Fig. 3C, lane E3), which contained 133 nucleotides (nt) of exon 3 sequences fused to 584 nt of exon 5 sequences (see Fig. 1), was cloned into the *HincII* site of pEMBL19. The resulting plasmid was digested with *BamHI* and *MscI*, and a 420-bp *BglIII-MscI* fragment isolated from plasmid p62/3 (10) was inserted. The 420-bp fragment contains 149 nt of the first intron, an *EcoRI* site in front of the first ATG, exon 2, and 173 nt of exon 3. The insertion of the 420-bp fragment restored the beginning of the *ie3* open reading frame. To provide the complete sequences of exon 5, the construct was digested with *XbaI* and *HindIII*, and exon 5 sequences were inserted as a 2.7-kbp *XbaI-HindIII* fragment comprising map units 0.780 to 0.769. The continuous *ie3* open reading frame was then isolated as a 2.9-kbp *EcoRI* fragment and inserted into the *EcoRI* site of the recombination vector pCS43.

**Production of specific antisera.** The expression vector pATH11, which contains a part of the *Escherichia coli* TrpE reading frame downstream of the inducible TrpE promoter (12), was chosen to construct a plasmid encoding a TrpE-IE3 fusion protein. Plasmid pATH11 was digested with *XbaI* and *PstI*, and the 870-bp *XbaI-PstI* fragment of the MCMV *ie3* region (map units 0.780 to 0.776) was inserted. The expression plasmid pATH11-IE3 contains an open reading frame encoding a polypeptide which consists of 334 amino acids of the TrpE protein followed by 290 amino acids of the putative IE3 protein. The production of the fusion protein was induced in *E. coli* C600 with indolacrylic acid (10 µg/ml) in the absence of tryptophan as described previously (40). Bacteria were harvested 5 h later, and following suspension in lysis buffer (0.8 mg of lysozyme per ml in 6.7% sucrose–80 mM EDTA–30 mM Tris [pH 8.0]) and sonication, the proteins were isolated by successive extraction with 1 and 7 M urea. The 66-kDa TrpE-IE3 fusion protein was mainly contained in the 7 M urea fraction. The isolated protein was dialyzed against 0.1 M ammonium carbamate. Rabbits were injected subcutaneously with 50 µg of protein three times with 14-day intervals and bled after 8 weeks. Production of the monoclonal antibody (MAb) 6/20/1 and the antipeptide serum b5/1 has been described previously (10, 24).

**Plasmids for the analysis of the regulatory functions of the IE proteins.** Recombinant plasmids were constructed according to established procedures (31). For the construction of the indicator plasmid pE1CAT, the plasmid pBB5.5 was used. Plasmid pBB5.5 contains the sequences from map units 0.698 to 0.721 of the MCMV genome as a 5.5-kbp *PstI* fragment. This 5.5-kbp *PstI* fragment includes 2.6 kbp of the *el* promoter and 2.9 kbp of the *el* coding sequences (3). pBB5.5 was digested with *NcoI* and *HindIII*, thereby removing all coding sequences of the *el* gene. The *NcoI* site is located 118 nt downstream of the 5' cap site of the *el* gene directly at the translation start codon of the *el* open reading

frame. The ATG at the *NcoI* site was removed by nuclease S1 digestion, and a *HindIII* linker was added. For isolation of the CAT gene, plasmid pSV2CAT (16) was linearized with *BamHI*, the *BamHI* site was filled in by Klenow enzyme, and a *HindIII* linker was added. The CAT gene was then isolated by digestion with *HindIII* as a 1.6-kbp *HindIII* fragment and inserted into the modified pBB5.5 plasmid 118 nt downstream of the 5' cap site of the *el* gene.

For the construction of the indicator plasmid pE1CAT, the MCMV enhancer sequences, the *ie1/ie3* promoter, and the first 49 nt of exon 1 of the *ie1* gene were isolated from the plasmid pAMB33 (23). pAMB33 contains a 2.25-kbp *PstI* fragment from map units 0.806 to 0.796. pAMB33 was digested with *SmaI* in the polylinker sequences, and a *HindIII* linker was added. Then, a 2.25-kbp *HindIII* fragment was isolated and ligated into the *HindIII* site of pSV0CAT (16). In the resulting plasmid, pE1CAT, the CAT gene is fused to the leader sequence of the *ie1* gene at position +49.

To construct a plasmid selectively encoding the IE3 protein, the plasmid pE111 (23) was cut with *MscI* and *XbaI*, thereby removing all sequences downstream of the *MscI* site in the third exon of the *ie1* gene. A 552-bp *MscI-XbaI* fragment was isolated from the 717-bp amplification product (see Fig. 3C, lane E3). This fragment, which contained 18 nt from the end of exon 3 linked to the first 534 nt of exon 5, was fused at the *MscI* site in exon 3. The remaining part of the *ie3* region was added at the *XbaI* site corresponding to map unit 0.780 as a 2.7-kbp *XbaI-HindIII* fragment (map units 0.780 to 0.769).

**Synthetic oligonucleotides.** Synthetic oligonucleotides for polymerase chain reactions and sequencing were synthesized by an Applied Biosystems 381A DNA synthesizer and purified by urea-polyacrylamide gel electrophoresis. The 5'-to 3'-sequences of the oligonucleotides are as follows: E1, CTGCCCAAGAAAGAGCGACCTTCCCC; E2, CCGTTGCAACATGATCATGATCGC; E3, GACCTGACTCTGGAGGACATGTTGGG; E5, GGTACCTGAACGCCACCATCCTTTC; IE3.1, GAAGTCTGGGGTGTGGTCTTCTGATTG; IE3.END, TCTGCATTTGAGTCAGG; (dT)17-R1-R0, AAGGATCCGTCGACATCGATAATACGACTCACTATAGGGATTTTTTTTTTTTTTTTTTTT; and R0, AAGGATCGTCGACATC (see also Fig. 3A).

**Isolation of IE RNA.** MEF were infected for 6 h with MCMV at a multiplicity of infection of 20 PFU per cell in the presence of cycloheximide (50 µg/ml) by using the technique of centrifugal enhancement of infectivity at 800 × *g* for 30 min. Total cell RNA was prepared by following published procedures (7).

**Preparation of cDNA and polymerase chain reaction.** A 10-µl volume containing 5 µg of IE RNA and 20 pM 3' primer in H<sub>2</sub>O were heated at 65°C for 5 min and then quenched on ice. Reverse transcription of the RNA was performed in a 20-µl reaction mixture containing 100 mM Tris (pH 8.3), 150 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 400 µM deoxynucleoside triphosphates (dNTPs), 5 U of human placental RNase inhibitor, and 10 U of murine reverse transcriptase (Pharmacia, Uppsala, Sweden) for 1 h at 42°C. The synthesized cDNA was subjected to amplification with 2.5 U of *Taq* polymerase (Perkin Elmer Cetus, Norwalk, Conn.) in a 100-µl reaction mixture which contained 1 µM 3' and 5' primers, 200 µM dNTPs, 10 mM Tris (pH 8.0), 50 mM KCl, and 1.5 mM MgCl<sub>2</sub>. Thirty-five cycles were carried out in a thermocycler (Biomed, Obertheres, Germany) under the following conditions: denaturation at 94°C for 2 min, primer annealing at 55°C for 4 min, and

elongation at 72°C for 4 to 7 min depending on the expected length of the amplified product. The determination of the 3' end of the *ie3* mRNA was done by the protocol of Frohman et al. (15).

**DNA sequencing.** DNA sequence analysis was carried out by using the method of Maxam and Gilbert (33) and a modified version of the enzymatic chain termination method (48). Direct sequencing of amplified DNA fragments was performed as follows. The DNA fragment was isolated from agarose gels by using the standard protocol for GeneClean DNA purification (GeneClean; BIO 101, Inc., La Jolla, Calif.). For the annealing reaction, the isolated DNA (500 ng) was suspended in a reaction mixture containing 1.5 pM primer, 100 mM Tris (pH 8.3), 150 mM KCl, 10 mM MgCl<sub>2</sub>, and 10 mM dithiothreitol, and was then incubated for 10 min at 100°C and for 10 min at room temperature. Labeling and extension reactions with polymerase T7 and [<sup>35</sup>S]dATP (Amersham, Braunschweig, Germany) were performed as recommended by the supplier of the sequencing kit (Pharmacia).

**Western blotting (immunoblotting).** MEF were infected with MCMV at 20 PFU per cell by using the technique of centrifugal enhancement of infectivity at 800 × *g* for 30 min. For selective expression of IE proteins, the cells were incubated from 0 to 3 h postinfection (p.i.) in the presence of cycloheximide (50 µg/ml), which was replaced from 3 to 7 h p.i. with actinomycin D (5 µg/ml). Infections with vaccinia viruses were performed at a multiplicity of infection of 20. Samples (10<sup>6</sup> cells) were lysed in protein sample buffer (3.4% sodium dodecyl sulfate [SDS], 20 mM dithiothreitol, 70 mM Tris [pH 6.8]), sonicated three times for 10 s, and boiled for 5 min. The polypeptides of the cell lysates were separated by 7.5 to 15% gradient SDS-polyacrylamide gel electrophoresis and transferred electrophoretically to nitrocellulose filters. Incubation of the filters with antisera and immunostaining were performed as described previously (3, 24).

**DNA transfection and CAT assay.** All plasmids were purified twice by cesium chloride gradients. Quantification of plasmids was performed spectrophotometrically. Quality of plasmid DNAs was examined by restriction enzyme digestion and gel electrophoresis. For transient gene expression assays, second-passage MEF were seeded into 60-mm-diameter plastic dishes (Greiner, Nürtingen, Germany) at 16 h prior to transfection. Effector and indicator plasmids (1 pmol) were transfected in duplicate by the calcium phosphate technique. All transfection experiments were repeated at least three times with different plasmid preparations. The experiments discussed here were carried out without carrier DNA, since the addition of carrier DNA to equalize the amount of DNA used for transfection had no effect on the resulting enzyme activity. Cells were harvested 40 h after transfection, and cell extracts were prepared according to the procedure of Gorman et al. (16). The 1-h standard assays and thin-layer chromatography were performed as described previously (25). Radioactive chloramphenicol was visualized by overnight exposure to Kodak X-Omat S films. For quantification, the chloramphenicol spots were excised from the silica plates and counted in a liquid scintillation counter.

**Nucleotide sequence accession number.** The GenBank accession number for the nucleotide sequence reported in this article is M77846.

## RESULTS

**Nucleotide sequence and structural analysis of the *ie3* gene.** Our previous experiments have shown that two 2.75-kb

mRNAs originate from the abundantly transcribed IE region of MCMV (Fig. 1). As DNA probes from the *ie1* and *ie3* region hybridize with a 2.75-kb mRNA as well as with a 5.1-kb RNA, it was concluded that the two 2.75-kb mRNAs probably represent alternative splice products of a primary 5.1-kb transcript. Hybrid selection and in vitro translation experiments suggested that both transcripts originate at least partially from the same regions and that they are spliced from common exons (23). Furthermore, we have observed that sequences downstream of the *ie1* gene are transcribed at IE times and that the endpoint of an mRNA with moderate abundance is located between map units 0.776 (*Pst*I) and 0.773 (*Xho*I) (Fig. 1) (23). From these previous findings and by analogy to the structure of the HCMV IE region, it was predicted that the MCMV *ie3* gene would share its first three exons with the MCMV *ie1* gene and that alternative splicing from the third exon to a fifth exon downstream of exon 4 of the MCMV *ie1* gene would occur.

To identify the putative open reading frame of exon 5, the nucleic acid sequence of the region downstream of the *ie1* gene was analyzed. Figure 2 shows the nucleotide sequence from the polyadenylation signal of the *ie1* mRNA to 50 nt downstream of the putative polyadenylation signal of the *ie3* transcript. The restriction enzyme sites *Ava*I (map unit 0.783), *Xba*I (map unit 0.780), and *Pst*I (map unit 0.776) are indicated for orientation. An open reading frame encoding 528 amino acids was identified within this region.

To study whether this coding sequence is spliced in total or in part to exon sequences of the *ie1* gene, reverse transcription of the *ie3* mRNA and amplification of the resulting *ie3* cDNA were performed. Primers E1, E2, and E3 were chosen from the upper strands of exon 1, exon 2, and exon 3 of the *ie1* gene, respectively (Fig. 3A and reference 22). Primer E5 corresponds to positions 560 to 584 in the lower strand shown in Fig. 2 and is therefore located far downstream of the expected 5' end of exon 5.

The 3' primer E5 was hybridized to IE RNA and elongated with reverse transcriptase. The synthesized cDNA was amplified by using primer pairs E1-E5, E2-E5, and E3-E5. The amplification resulted in DNA fragments of 1,100, 850, and 700 bp, thus confirming the prediction that sequences of the *ie3* region are spliced to the third exon of the *ie1* gene (Fig. 3C). Some smaller fragments of an amount that varied between experiments were interpreted as single-stranded templates because of an unbalanced amount of the two primers and were not further analyzed.

To locate the 5' end of exon 5 and to confirm that exon sequences of the *ie1* gene are included in the *ie3* transcript, sequence analysis of the amplified fragments was performed with primers E2, E3, and IE3.1. Primer IE3.1 corresponds to positions 79 to 107 from the expected start of exon 5 in the lower strand shown in Fig. 2. The 700- and 850-bp DNA fragments were eluted from the agarose gel and sequenced directly. The sequence analysis located the 5' end of the fifth exon 538 nt upstream of the *Xba*I site (map unit 0.780). A splice acceptor consensus sequence, CTTTATGGTTCAC AG, was identified upstream of the 5' end of exon 5 (positions -15 to -1 in Fig. 2). Therefore, the 5' end of exon 5 is located 54 and 28 nt downstream of the polyadenylation signal and the 3' end of the 2.75-kb *ie1* transcript, respectively (22).

The sequence analysis of the amplified DNA fragments confirmed that the *ie3* transcript is composed of the first three exons of the *ie1* gene and exon 5 in the *ie3* region. The open reading frame of the *ie3* transcript starts at the beginning of exon 2, continues through the exon 3 sequences, and

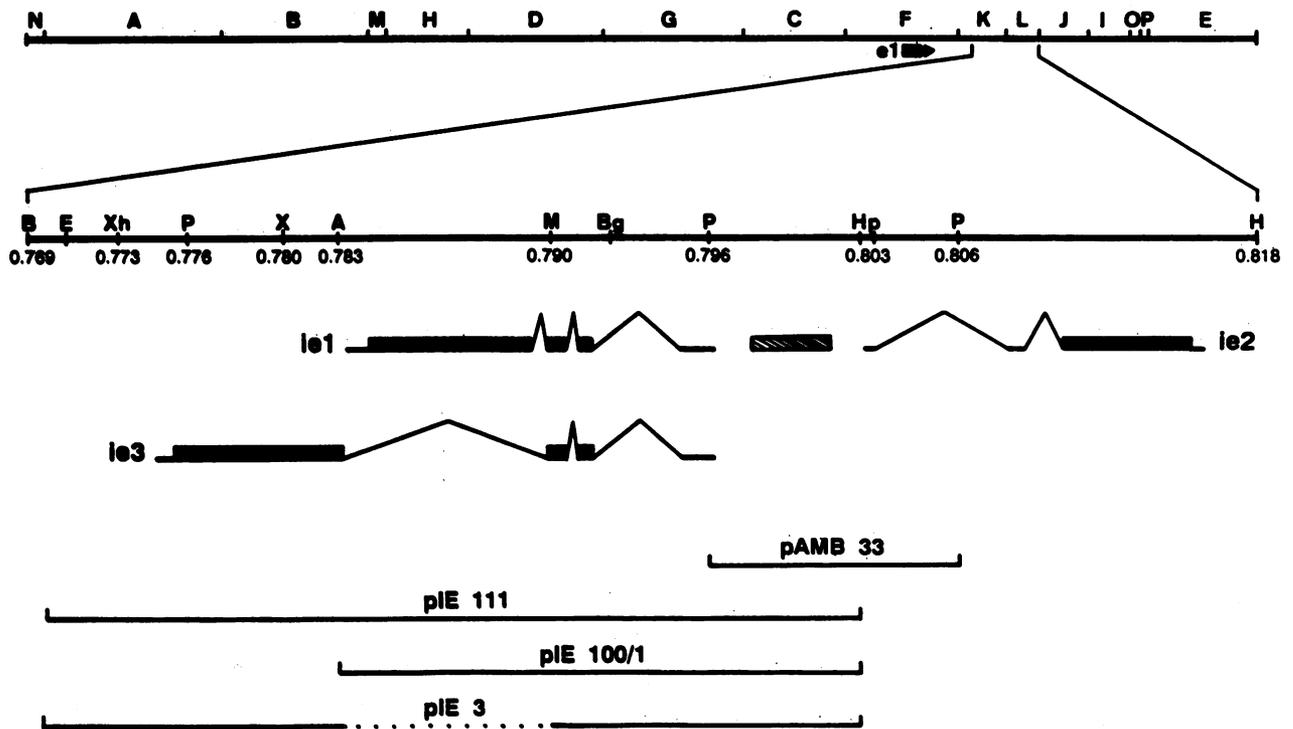


FIG. 1. (Top) Physical map of the MCMV genome with *Hind*III cleavage sites and location of the transcription unit *e1*. (Middle) Organization of the abundantly transcribed IE region between map units 0.769 and 0.818. Abbreviations of restriction enzyme sites used for *ie3* gene characterization and cloning purposes: A, *Ava*I; B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; Hp, *Hpa*I; M, *Msc*I; P, *Pst*I; X, *Xba*I; Xh, *Xho*I. The splicing patterns of the *ie1*, *ie2*, and *ie3* genes are drawn to scale. Coding sequences are shown as black bars. The hatched bar represents the enhancer sequences. (Bottom) Effector plasmids used for analysis of the regulatory functions of the IE proteins. The dotted line indicates deleted sequences.

uses a further 511 of the 528 codons of the open reading frame in the *ie3* region (Fig. 2). A polyadenylation consensus sequence, AATAAA, is located 137 nt downstream of the stop codon TGA (positions 1676 to 1682 in Fig. 2). The sequence GCACTTGTTCTTGT, 20 nt downstream of the putative polyadenylation signal (positions 1703 to 1717 in Fig. 2), shows homology to the typical 3' end consensus sequences CAC/TTG and YGTGTTY (2).

The 3' end of the *ie3* transcript was determined by reverse transcription of the *ie3* mRNA and amplification of the cDNA end according to the method of Frohman et al. (15). In brief, the primer (dT)17-R1-R0 was annealed to IE RNA, and reverse transcription was performed. Amplification of the 3' end of the *ie3* cDNA was achieved by utilizing the *ie3* gene-specific primer IE3.END (Fig. 3A) and primer R0. A 350-bp fragment was identified after amplification and subcloned into pUC19. Sequence analysis located the 3' end of the *ie3* mRNA to position 1701 (Fig. 2 and 3D). S1 analysis with different 5' and 3' end-labeled probes of the *ie3* region confirmed the 5' and 3' ends of exon 5 we had determined and revealed no indication for additional splicing within the *ie3* region (data not shown).

Thus, the *ie3* mRNA contains four exons of 300, 111, 191, and 1,701 nt comprising 2.3 kb (Fig. 3B). The sizes of the introns are 825, 111, 191, and 1,856 nt. The observed size of 2.75 kb for the *ie3* mRNA in Northern (RNA) blot analysis is in accord with the calculated size if the increase in size of the mRNA by polyadenylation is considered. Furthermore, the calculated sizes of 2,305 nt for the *ie1* transcript (22) and 2,303 nt for the *ie3* transcript explain the comigration of the

*ie1* and *ie3* 2.75-kb mRNAs in agarose gels seen previously (23).

**Homology of MCMV IE3 and HCMV IE2 amino acid sequences.** The amino acid sequence of the IE3 protein was deduced from the nucleotide sequence of the *ie3* mRNA (Fig. 4). The predicted IE3 protein contains 611 amino acids and has a calculated molecular mass of 68 kDa. As the *ie1* mRNA and the *ie3* mRNA both use the information of exon 2 and 3, the first 99 amino acids at the amino-terminal ends of pp89 and the IE3 protein are identical. Some interesting regions within the unique amino acid sequence of the IE3 protein defined by exon 5 sequences were identified. Three clusters with acidic residues, i.e., amino acids 166 to 177, 180 to 188, and 231 to 239, were observed, although the overall content of acidic amino acids (92 residues) is balanced by a similar number of basic residues (110 residues), resulting in a calculated isoelectric point of 6.75. The sequence Lys-Lys-Ala-Lys-Lys-His-Lys between positions 141 and 147 shows similarity to the simian virus 40 large T antigen nuclear location signal Lys-Lys-Lys-Arg-Lys (20). Another series of basic amino acids, Arg-His-His-His-Lys-Arg-Lys between positions 279 and 285, may represent a second nuclear location signal. A cluster of five glutamine residues was found at positions 365 to 370 of the IE3 amino acid sequence. Serine residues, often clustered, are frequent (17%), especially in the middle of the IE3 amino acid sequence. The serine clusters may represent phosphorylation sites. According to Chou and Fasman (8), the high number of serine residues together with glycine and alanine





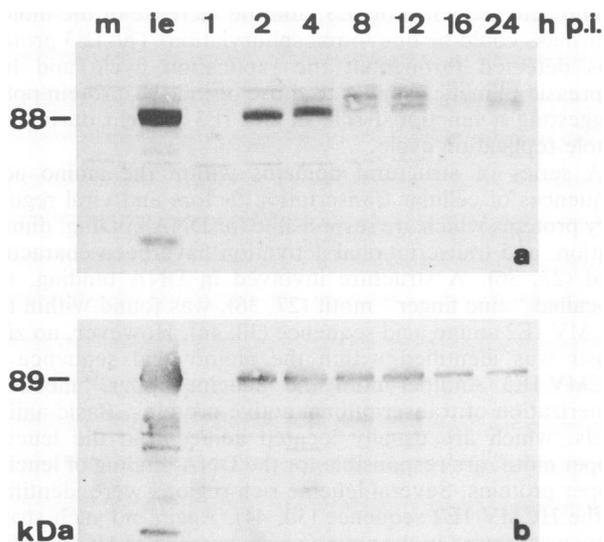


FIG. 6. Western blot analysis of IE3 protein expression kinetics. Lysates of mock-infected MEF (lane m), MCMV-infected MEF in the IE phase (lane ie), and MCMV-infected MEF at various hours p.i. (hours are given above each lane) were reacted with IE3-specific antiserum (a) and IE1-specific MAb 6/20/1 (b).

plasmid containing the indicator gene CAT under the control of the *el* promoter was constructed. Different effector plasmids were used for the expression of the IE1 and IE3 proteins (Fig. 1). The plasmid pIE111, which encompasses the MCMV IE region, encodes the complete *ie1/ie3* gene complex (13, 22, 23). The plasmid pIE100/1 encodes only IE1 proteins, as sequences downstream of the fourth exon at map unit 0.783 (*Ava*I) are not included. The plasmid pIE3 differs from pIE111 by the deletion of the third intron and the fourth exon and thus encodes only the IE3 protein.

MEF, which are permissive for MCMV infection, were used for transfection experiments. Transfection of the indicator plasmid pE1CAT alone resulted in a low basal level of CAT activity (Fig. 7, lanes 1). Following cotransfection of pE1CAT and pIE111, a 200-fold induction of CAT activity was observed (Fig. 7, lanes 2), demonstrating that IE proteins encoded by pIE111 activate the early promoter *el*. After cotransfection of pE1CAT and pIE100/1, only a low level of CAT activity was detected, indicating that the early promoter was not activated by the *ie1* gene product alone (Fig. 7, lanes 3). In contrast, cotransfection of pE1CAT and

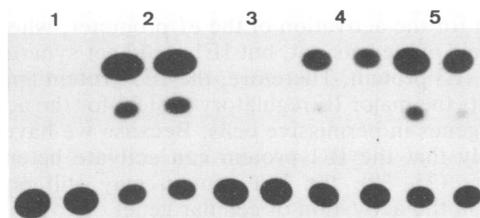


FIG. 7. Activation of the MCMV *el* promoter by IE proteins. Transfection experiments in MEF were performed in duplicate with the indicator plasmid pE1CAT alone (lanes 1) or in combination with the effector plasmids pIE111 (lanes 2), pIE100/1 (lanes 3), pIE3 (lanes 4), and pIE3 plus pIE100/1 (lanes 5). Cell extracts were prepared 40 h after transfection and assayed for CAT activity.

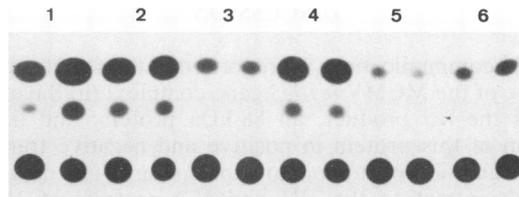


FIG. 8. Regulation of the MCMV *ie1/ie3* promoter by IE proteins. The indicator plasmid pE1CAT was transfected into MEF alone (lanes 1) or in combination with the effector plasmids pAMB33 (lanes 2), pIE111 (lanes 3), pIE100/1 (lanes 4), pIE3 (lanes 5), and pIE3 plus pIE100/1 (lanes 6). CAT activities were determined as described in Materials and Methods.

pIE3 resulted in a 150-fold induction of CAT activity (Fig. 7, lanes 4). Therefore, the IE3 protein is sufficient for *trans* activation of the *el* promoter, whereas the IE1 protein is not. To study a potential cooperation between the IE1 and IE3 proteins in the activation of the early promoter, pE1CAT and the two effector plasmids pIE100/1 and pIE3 were cotransfected. The observed CAT activity was 1.5 to 2 times higher than after transfection of pE1CAT and pIE3 (Fig. 7, lanes 5). This suggests that the IE1 protein has some additive effect on the *trans*-activating properties of the IE3 protein.

**Repression of the *ie1/ie3* promoter by the IE3 protein.** Transcription of the MCMV IE genes ceases during the early phase of infection (21, 24). A possible mechanism for this downregulation is an autoregulatory repression of the MCMV IE promoters by the IE proteins. To study the regulation of the *ie1/ie3* promoter, an indicator plasmid containing the CAT gene under the control of the *ie1/ie3* promoter and upstream enhancer sequences was constructed. Because the upstream sequences in all effector plasmids are identical to those in the indicator plasmid, the plasmid pAMB33 (23) was included as a control for competition for cellular transcription factors. pAMB33 contains a 2.25-kbp *Pst*I fragment (map units 0.806 to 0.796) encompassing the enhancer and the *ie1/ie3* promoter (Fig. 1).

Transfection of the indicator plasmid pE1CAT into MEF resulted in high CAT activity (Fig. 8, lanes 1) due to the MCMV enhancer activity (13, 25). Following transfection of pE1CAT and pAMB33, a reduction of CAT activity by 25% could be seen (Fig. 8, lanes 2), indicating a slight competition for transcription factors. Cotransfection of pE1CAT and pIE111, however, resulted in a fivefold repression of the *ie1/ie3* promoter (Fig. 8, lanes 3), whereas no suppression or even some activation was seen after cotransfection of pE1CAT and pIE100/1 (Fig. 8, lanes 4). This indicated that the IE1 protein alone has no major effect on the *ie1/ie3* promoter, whereas the expression of the complete *ie1/ie3* gene complex resulted in repression. To study whether the repression of the *ie1/ie3* promoter is a function of the IE3 protein alone, cotransfection of pE1CAT and pIE3 was carried out. The strong reduction of CAT expression (Fig. 8, lanes 5) confirmed that the *ie1/ie3* promoter is negatively regulated by the IE3 protein. Cotransfection of pE1CAT, pIE3, and pIE100/1 resulted in the same reduced CAT activity as that resulting from transfection of pE1CAT and pIE3 (Fig. 8, lanes 6), indicating that under the experimental conditions applied, the repressive properties of the IE3 protein prevented a potential activation of the *ie1/ie3* promoter by the IE1 protein.

## DISCUSSION

In this communication, we report on (i) the structure of the *ie3* gene of the MCMV *ie1/ie3* gene complex; (ii) the expression of the *ie3* product, an 88-kDa protein; and (iii) the function of this protein in positive and negative transcriptional regulation of homologous promoters. The data show that in contrast to the IE1 and IE2 proteins of MCMV described by us previously (22, 35), the IE3 protein shows a striking homology to its HCMV counterpart, the HCMV IE2 protein, in its amino acid sequence and a great similarity in regulatory function. Similar to the function of the HCMV IE2 protein in HCMV, the MCMV IE3 protein is probably the major transcriptional regulator for the initiation of early gene expression in MCMV. The availability of both genes will allow a direct and more-detailed comparison of their functions and thus aid the understanding of these important regulatory proteins of CMV.

The MCMV *ie1/ie3* gene complex comprises five exons. Differential splicing leads to the 2.75-kb *ie1* and *ie3* transcripts. Splicing from the first or second exon to the fifth exon was not observed. Therefore, two major IE transcripts originate from the MCMV *ie1/ie3* gene complex: the *ie1* mRNA, which is composed of exons 1, 2, 3, and 4, and the *ie3* mRNA, consisting of exons 1, 2, 3, and 5. These aspects of the structural organization of the MCMV IE region are identical to those of the HCMV IE region, which is also composed of five exons (18, 43, 45, 46). In HCMV, distinct minor mRNA species originating from exon 5 sequences have been identified (43, 46). It has been proven that a splicing event resulting in a transcript which encodes a 55-kDa HCMV IE2 protein can occur within the fifth exon. We were not able to detect distinct minor IE mRNA species in MCMV, although a smear of smaller mRNAs is indicative of their existence (23). S1 nuclease analysis with various end-labeled DNA fragments from the *ie3* region revealed no RNA species that are generated by a splicing event within exon 5. If potential shorter transcripts from the MCMV *ie3* region have escaped detection, they must be of low abundance. Further cDNA cloning experiments are needed to clarify the origin of the smaller IE transcripts.

The open reading frame of the *ie3* transcript predicted a polypeptide of 611 amino acids with a calculated molecular mass of 68 kDa. An 88-kDa protein was identified as the major *ie3* gene product. Integration of the identified *ie3* open reading frame into vaccinia virus and expression of an 88-kDa protein in Vac-*ie3*-infected cells confirmed that the 88-kDa IE3 protein is encoded by the 2.75-kb *ie3* mRNA. The low-abundance 77-kDa protein probably represents a modification product of the 88-kDa protein, as the 77-kDa protein is also detected in lysates of Vac-*ie3*-infected cells. The IE3-specific antiserum reacted with a 54-kDa protein after enhanced expression of IE proteins by MCMV but not after infection with the splicing-incompetent vaccinia virus recombinant. The possibility that this protein is translated from a differentially spliced transcript is therefore not excluded, and further studies are required to clarify whether it represents the homolog of the HCMV 55-kDa IE2 protein.

The 88-kDa IE3 protein shows a slower mobility in SDS-polyacrylamide gels than expected from the calculated molecular mass of the deduced amino acid sequence, an observation also characteristic for the MCMV IE1 protein and the HCMV IE proteins (22, 24, 43–45). During the early phase of the MCMV replication cycle, the molecular mass of the IE3 protein shifted to 89 and 90 kDa, suggesting posttranslational modification. Large serine clusters are located within the

amino acid sequence of IE3, and the increase of the molecular mass could be due to phosphorylation. The IE3 protein was detected throughout the replication cycle and had expression kinetics similar to those of the IE1 protein pp89, suggesting a functional role of the IE3 protein during the whole replication cycle.

A series of structural domains within the amino acid sequences of cellular transcription factors and viral regulatory proteins which are responsible for DNA binding, dimerization, and transcriptional activation have been characterized (27, 36). A structure involved in DNA binding, the so-called "zinc finger" motif (27, 36), was found within the HCMV IE2 amino acid sequence (30, 44). However, no zinc finger was identified within the amino acid sequence of MCMV IE3. Another motif, the "leucine zipper," mediates dimerization of transcriptional active proteins. Basic amino acids, which are usually located adjacent to the leucine zipper motif, are responsible for the DNA binding of leucine zipper proteins. Several leucine-rich regions were identified in the HCMV IE2 sequence (30, 44). Again, no such structures were found in the amino acid sequence of MCMV IE3.

Transcriptional activator domains of yeast activator proteins and of the herpes simplex virus regulatory protein Vmw65 are characterized by acidic residues (47, 49). Furthermore, serine/threonine-, glutamine-, and proline-rich domains have been implicated in transcriptional activation by regulatory proteins (9, 34). IE3 has three short acidic clusters (amino acids 166 to 177, 180 to 188, and 231 to 239) and a series of five glutamine residues (amino acids 365 to 370). In addition, 10 of the 20 carboxy-terminal amino acids of the IE3 protein are acidic. The short acidic clusters have counterparts in the amino acid sequence of the HCMV IE2 protein, although the precise locations are not conserved (43). The 29 carboxy-terminal amino acids of the HCMV IE2 protein include 7 acidic residues, and they may fold into an amphipathic alpha helix (30). The precise positions of these negatively charged residues are not conserved in the amino acid sequence of MCMV IE3, and the propensity of the carboxy terminus of MCMV IE3 to form an amphipathic alpha helix is low. Altogether, the 200 carboxy-terminal amino acids of MCMV IE3 and HCMV IE2 represent the region of major homology, which suggests an important function for this domain of the CMV IE proteins.

Transfection experiments in MEF, which are permissive for MCMV infection, were performed with indicator genes under the control of the homologous early promoter *e1* or the *ie1/ie3* promoter and effector plasmids encoding the different MCMV IE proteins. Characterization of the MCMV early transcription unit *e1* (3) and structural analysis of the MCMV *ie1/ie3* gene complex allowed us to analyze in detail which IE proteins are required for the activation of the homologous early promoter *e1*. The IE3 protein alone was found to be sufficient for the activation of the *e1* promoter, whereas the IE1 protein alone was not, but IE1 could act synergistically with the IE3 protein. Therefore, the IE3 protein apparently represents the major IE regulatory protein for the activation of early genes in permissive cells. Because we have shown previously that the IE1 protein can activate heterologous promoters (25, 39), the IE1 protein may still possess a function in the activation of cellular genes.

In its requirement for IE proteins for *trans* activation, the MCMV *e1* promoter shows similarity to some HCMV early promoters. An HCMV E transcription unit corresponding to the MCMV transcription unit *e1* is located between map units 0.682 and 0.713 of the HCMV genome (3, 42). The promoter of this HCMV early gene can be activated by

HCMV IE1 and IE2 proteins (41). Similar to the induction of the MCMV *e1* promoter by the MCMV IE3 protein, the HCMV IE2 protein alone can also activate the HCMV E1.7 promoter (4). The activation of this HCMV promoter by HCMV IE2 is also augmented by the HCMV IE1 protein (30).

The autoregulative effect of the MCMV IE3 protein on the *ie1/ie3* promoter may explain the reduction of IE transcription at the beginning of the early phase (21, 24). Similar to the way the MCMV IE3 protein represses the MCMV *ie1/ie3* promoter, the HCMV IE2 protein downregulates the transcription of genes under the control of the HCMV major IE promoter (6, 17, 28, 37, 38, 44). The target sequences mediating repression of the HCMV major IE promoter by the HCMV IE2 protein have been defined and are conserved between HCMV and SCMV (6, 17, 28, 37) but do not exist in the MCMV *ie1/ie3* promoter and cannot be replaced by MCMV sequence motifs, since the HCMV IE2 protein cannot repress transcription from the MCMV *ie1/ie3* promoter (17). Thus, the MCMV IE3 protein mediates repression of the *ie1/ie3* promoter by different and as-yet-undefined target sequences.

Mutational analysis of domains which are conserved between the different CMV homologs will show which domains are important for the functions of the MCMV IE3 protein and will help us understand the control of transcriptional regulation in CMV.

#### ACKNOWLEDGMENTS

The technical assistance of D. Rettenmeier, K. Schneider, and A. Schroeder is gratefully acknowledged. We thank Karen Burke for helpful comments on the manuscript.

This study was supported by the Deutsche Forschungsgemeinschaft through SFB 322 grant A7.

#### REFERENCES

1. Arnold, B., M. Messerle, L. Jatsch, G. Kühlbeck, and U. H. Koszinowski. 1990. Transgenic mice expressing a soluble foreign H-2 class I antigen are tolerant to allogeneic fragments presented by self class I but not the whole membrane-bound alloantigen. *Proc. Natl. Acad. Sci. USA* **87**:1762-1766.
2. Birnstiel, M. L., M. Busslinger, and K. Strub. 1985. Transcription termination and 3' processing: the end is in site! *Cell* **41**:349-359.
3. 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.
4. Chang, C.-P., C. L. Malone, and M. F. Stinski. 1989. A human cytomegalovirus early gene has three inducible promoters that are regulated differentially at various times after infection. *J. Virol.* **63**:281-290.
5. 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.
6. Cherrington, J. M., E. L. Khoury, and E. S. Mocarski. 1991. Human cytomegalovirus *ie2* negatively regulates  $\alpha$  gene expression via a short target sequence near the transcription start site. *J. Virol.* **65**:887-896.
7. Chirgwin, J. M., A. E. Przybala, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**:5294-5299.
8. Chou, P. Y., and G. D. Fasman. 1974. Prediction of protein conformation. *Biochemistry* **13**:222-245.
9. Courey, A. J., and R. Tjian. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**:887-898.
10. Del Val, M., H. Volkmer, J. B. Rothbard, S. Jonjić, M. Messerle, J. Schickedanz, M. J. Reddehase, and U. H. Koszinowski. 1988. Molecular basis for cytolytic T-lymphocyte recognition of the murine cytomegalovirus immediate-early protein pp89. *J. Virol.* **62**:3965-3972.
11. Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the Vax. *Nucleic Acids Res.* **12**:387-395.
12. Dieckmann, C. L., and A. Tzagoloff. 1985. Assembly of the mitochondrial membrane system. *J. Biol. Chem.* **260**:1513-1520.
13. 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.
14. Drillien, R., and D. Spehner. 1983. Physical mapping of vaccinia virus temperature-sensitive mutations. *Virology* **131**:385-393.
15. Frohman, M. A., M. K. Dush, and G. R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**:8998-9002.
16. Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
17. 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.
18. 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.
19. Jeang, K.-T., M.-S. Cho, and G. S. Hayward. 1984. Abundant constitutive expression of the immediate-early 94K protein from cytomegalovirus (Colburn) in a DNA-transfected mouse cell line. *Mol. Cell. Biol.* **4**:2214-2223.
20. Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith. 1984. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature (London)* **311**:33-38.
21. Keil, G. M., A. Ebeling-Keil, and U. H. Koszinowski. 1984. Temporal regulation of murine cytomegalovirus transcription and mapping of viral RNA synthesized at immediate early times after infection. *J. Virol.* **50**:784-795.
22. 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.
23. 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.
24. Keil, G. M., M. R. Fibi, and U. H. Koszinowski. 1985. Characterization of the major immediate-early polypeptides encoded by murine cytomegalovirus. *J. Virol.* **54**:422-428.
25. 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.
26. Koszinowski, U. H., M. J. Reddehase, G. M. Keil, H. Volkmer, S. Jonjić, M. Messerle, M. DelVal, W. Mutter, K. Münch, and B. Bühler. 1987. Molecular analysis of herpesviral gene products recognized by protective cytolytic T lymphocytes. *Immunol. Lett.* **16**:185-192.
27. Latchman, D. S. 1990. Eukaryotic transcription factors. *Biochem. J.* **270**:281-289.
28. Liu, B., T. W. Hermiston, and M. F. Stinski. 1991. A *cis*-acting element in the major immediate-early (IE) promoter of human cytomegalovirus is required for negative regulation by IE2. *J. Virol.* **65**:897-903.
29. Mackett, M., G. L. Smith, and B. Moss. 1984. General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. *J. Virol.* **49**:857-864.
30. Malone, C. L., D. H. Vesole, and M. F. Stinski. 1990. Transac-

- tivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral protein. *J. Virol.* **64**:1498–1506.
31. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  32. 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.
  33. Maxam, A. M., and W. Gilbert. 1980. Sequencing with base-specific chemical cleavages. *Methods Enzymol.* **65**:499–560.
  34. Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian. 1989. The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell* **58**:741–753.
  35. Messerle, M., G. M. Keil, and U. H. Koszinowski. 1991. Structure and expression of the murine cytomegalovirus immediate-early gene 2. *J. Virol.* **65**:1638–1643.
  36. Mitchell, P. J., and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**:371–378.
  37. Pizzorno, M. C., and G. S. Hayward. 1990. The IE2 gene products of human cytomegalovirus specifically down-regulate expression from the major immediate-early promoter through a target sequence located near the cap site. *J. Virol.* **64**:6154–6165.
  38. 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.
  39. Schickedanz, J., L. Philipson, W. Ansoorge, R. Pepperkok, R. Klein, and U. H. Koszinowski. 1988. The 89,000-*M<sub>r</sub>* murine cytomegalovirus immediate-early protein stimulates *c-fos* expression and cellular DNA synthesis. *J. Virol.* **62**:3341–3347.
  40. 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.
  41. Staprans, S. I., D. K. Rabert, and D. H. Spector. 1988. Identification of sequence requirements and *trans*-acting functions necessary for regulated expression of a human cytomegalovirus early gene. *J. Virol.* **62**:3463–3473.
  42. Staprans, S. I., and D. H. Spector. 1986. 2.2-Kilobase class of early transcripts encoded by cell-related sequences in human cytomegalovirus strain AD169. *J. Virol.* **57**:591–602.
  43. 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.
  44. 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.
  45. 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.
  46. 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.
  47. Struhl, K. 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* **49**:295–297.
  48. Tabor, S., and C. C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84**:4767–4771.
  49. Triezenberg, S. J., R. C. Kingsbury, and S. L. McKnight. 1988. Functional dissection of VP16, the *trans*-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* **2**:718–729.
  50. Volkmer, H., C. Bertholet, S. Jonjic, R. Wittek, and U. H. Koszinowski. 1987. Cytolytic T lymphocyte recognition of the murine cytomegalovirus nonstructural immediate-early protein pp89 expressed by recombinant vaccinia virus. *J. Exp. Med.* **166**:668–677.