ICP22 Homolog of Equine Herpesvirus 1: Expression from Early and Late Promoters

V. ROGER HOLDEN, RAMANA R. YALAMANCHILI,[†] RONALD N. HARTY,
AND DENNIS J. O'CALLAGHAN*

*Department of Microbiology and Immunology, Louisiana State University Medical Center,
1501 Kings Highway, Shreveport, Louisiana 71130-3932*

Received 19 August 1991/Accepted 17 October 1991

The complete nucleotide sequence of the short region, made up of a unique segment (Us; 6.5 kb) bracketed by a pair of inverted repeat sequences (IR; 12.8 kb each), of the equine herpesvirus 1 (EHV-1) genome has been determined recently in our laboratory. Analysis of the IR segment revealed a major open reading frame (ORF) designated IR4. The IR4 ORF exhibits significant homology to the immediate-early gene US1 (ICP22) of herpes simplex virus type 1 and to the ICP22 homologs of varicella-zoster virus (ORF63), pseudorabies virus (RSp40), and equine herpesvirus 4 (ORF4). The IR4 ORF is located entirely within each of the inverted repeat sequences (nucleotides [nt] 7918 to 9327) and has the potential to encode a polypeptide of 469 amino acids (49,890 Da). Within the IR4 ORF are two reiterated sequences: a 7-nt sequence tandemly repeated 17 times and a 25-nt sequence tandemly repeated 13 times. Nucleotide sequence analyses of IR4 also revealed several potential *cis*-regulatory sequences, two TATA sequences separated by 287 nt, an in-frame translation initiation codon following each TATA sequence, and a single polyadenylation site. To address the nature of the mRNA species encoded by IR4, we used Northern (RNA) blot and S1 nuclease analyses. RNA mapping data revealed that IR4 has two promoters that are regulated differentially during a lytic infection. A 1.4-kb mRNA appears initially at 2 h postinfection and is an early transcript since its synthesis is not affected by the presence of phosphonoacetic acid, an inhibitor of EHV-1 DNA replication. In contrast, a 1.7-kb mRNA appears at later times postinfection and is designated as a gamma-1 transcript, since its synthesis is significantly reduced by phosphonoacetic acid. These IR4-specific mRNAs are 3' coterminal, have unique 5' termini, and would code for in-frame, overlapping, carboxy-coterminal proteins of 293 and 469 amino acids, respectively. Interestingly, the site of homologous recombination to generate the genome of EHV-1 defective interfering particles that initiate persistent infection occurs between nt 3244 and 3251 of UL3 (ICP27 homolog) and nt 9027 and 9034 of IR4 (ICP22 homolog). Thus, this recombination event would generate a unique ORF that would encode a potential protein whose amino end was derived from the N-terminal 193 amino acids of the ICP22 homolog and whose carboxyl end was derived from the C-terminal 68 amino acids of the ICP27 homolog.

Equine herpesvirus 1 (EHV-1), an alphaherpesvirus, is a major pathogen of equines, causing respiratory disease, neurological disorders, and spontaneous abortions in pregnant mares (2, 8, 44, 45). EHV-1 possesses a linear double-stranded DNA genome (34, 48, 55, 57, 64) of approximately 144 kbp with a G+C content of 57% (60). The genome of EHV-1 (Kentucky A strain) is composed of a long region (112 kbp) of unique sequences and a short region (32 kbp) which consists of identical inverted repeat sequences (IR; 12.8 kbp each) that bracket the short unique sequences. EHV-1 gene expression is regulated at the level of transcription (17, 18, 36), and the synthesis of EHV-1 gene products is coordinately regulated and sequentially ordered in an immediate-early, early, and late fashion (12, 25, 26). The immediate-early transcript is the first to be synthesized and is expressed in the absence of protein synthesis (25, 26). In contrast to herpes simplex virus type 1 (HSV-1), EHV-1 possesses only one immediate-early gene (25, 26), yet multiple immediate-early polypeptides are detected (11, 54). The early genes are expressed prior to viral DNA synthesis, and their expression is not prevented by inhibitors of viral DNA synthesis (12, 25, 26). Lastly, late genes are expressed

following the initiation of viral DNA synthesis, and their expression is either greatly reduced (gamma-1 genes) or completely blocked (gamma-2 genes) by inhibitors of viral DNA synthesis (12, 25, 26).

DNA sequencing and transcriptional analyses have identified six genes within the IR sequences of EHV-1 (IR1 through IR6) (7, 29-32, 35, 53, 67). Two of these genes (IR1 and IR2) have been characterized in detail at the DNA and RNA levels (29-32). The IR1 gene is an immediate-early gene that exhibits extensive homology to ICP4 of HSV-1 and ORF62 of varicella-zoster virus (VZV) (25, 26, 29-31). IR2, an early gene, is embedded within the immediate-early gene and encodes a 4.4-kb transcript that is 3' coterminal with the IR1 mRNA (32). IR3 is a gamma-1 gene that expresses a 0.9-kb mRNA that overlaps the 5' end of the IR1 mRNA and may encode a 10-kDa protein (35). IR5 is a gamma-1 gene that would encode a protein of 239 amino acids (69) and is homologous to US10 of HSV-1 and ORF63 of VZV, both of which encode a virion protein. IR6 maps within the IR sequence at the IRs/U_s junction and encodes a protein of approximately 30 kDa of unknown function (7).

In this report, we describe the IR4 gene sequence and show that IR4 is a homolog of the immediate-early gene US1 (ICP22) of HSV-1. Northern (RNA) blot hybridization and S1 nuclease analyses revealed that the IR4 gene is transcribed to yield a 1.4-kb early mRNA and a 1.7-kb late mRNA. These transcripts are 3' coterminal and encode

* Corresponding author.

[†] Present address: Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115.

polypeptides that are carboxy coterminal. These results will be the basis for future studies to localize the functional domain(s) of these putative regulatory proteins.

MATERIALS AND METHODS

Cells and virus. The Kentucky A strain of EHV-1 was passaged in L-M suspension cells at a multiplicity of infection of 0.05 PFU per cell to minimize the chance of generating defective interfering particles (3, 45, 47).

Cloning and sequencing. The isolation and purification of EHV-1 clones have been described previously (3, 29, 68, 69). Clones pCS-1 and pCS-2 (map units 0.820 to 0.836 and 0.836 to 0.851, respectively) were inserted into pUC118, pUC119, M13mp18, and M13mp19 vectors by using standard methods (40). A series of overlapping subclones of the pCS-1 and pCS-2 region, generated by restriction endonuclease and exonuclease III digestion, were sequenced by using the dideoxy-chain termination method of Sanger et al. (58) with the Klenow fragment of *Escherichia coli* DNA polymerase (New England Nuclear, Boston, Mass.), reverse transcriptase (Promega, Madison, Wis.), or Sequenase (USB Corp., Cleveland, Ohio). To minimize the occurrence of compressions, deaza-dGTP was used in some reactions. Both DNA strands were sequenced for accuracy. DNA sequence analysis was carried out by using the computer facilities of Bionet, IBI/Pustell DNA/protein sequence analysis software, and PC/Gene software by IntelliGenetics (International Biotechnologies, Inc. New Haven, Conn.).

Isolation of poly(A)⁺ RNA. To isolate immediate-early mRNA, 1×10^8 L-M cells were pretreated with 100 μ g of cycloheximide (Sigma Chemical Co., St. Louis, Mo.) per ml for 1 h. The L-M cells were infected with EHV-1 at a multiplicity of infection of 15 PFU per cell, and the cells were harvested at 4 h postinfection. The cells were maintained in the presence of cycloheximide throughout the infection. To isolate early mRNA, 1×10^8 cells were treated with 100 μ g of phosphonoacetic acid (PAA; Abbot Laboratories, Chicago, Ill.) per ml at 1 h prior to infection and infected cells maintained in the presence of PAA were harvested at 6 h postinfection. These concentrations of cycloheximide and PAA have been shown to inhibit protein synthesis and DNA synthesis by 98 and 96%, respectively (12). To isolate late mRNA, we harvested infected cells at 8 h postinfection in the absence of metabolic inhibitors. Poly(A)⁺ RNA was isolated by using the Fast Track mRNA isolation kit (Invitrogen Corp., San Diego, Calif.) as described previously (32).

Northern blot analysis. Northern (RNA) blot analysis was performed by using 5.0 μ g of poly(A)⁺ RNA fractionated on a formaldehyde-1.2% agarose gel as described by Maniatis et al. (40). The RNA was blotted onto GeneScreen Plus filters (New England Nuclear) by the method of Southern (61). EHV-1 subclones were labeled with [α -³²P]deoxynucleoside triphosphates ([α -³²P]dNTP) as described by Rigby et al. (51) and hybridized to RNA on the filters as described previously (25, 26, 30, 32). The filters were washed extensively and exposed to X-OMAT film (Kodak, Rochester, N.Y.) with an intensifying screen at -70°C.

S1 nuclease analysis. S1 nuclease analysis was performed as described elsewhere (5, 30, 32, 62). Specific EHV-1 subclones were digested at the appropriate site and 5' end labeled with T4 polynucleotide kinase ([γ -³²P]-ATP) or 3' end labeled with the Klenow fragment ([α -³²P]dNTP). The probes were hybridized to 3.0 μ g of poly(A)-selected RNA in 25 μ l of S1 hybridization buffer [40 mM piperazine-N,N'-

bis(2-ethanesulfonic acid) (PIPES; pH 6.4), 400 mM NaCl, 1.0 mM EDTA, 80% formamide] at 65°C for 14 to 16 h. The DNA-RNA hybrids were digested in 300 μ l of S1 digestion buffer (250 mM NaCl, 30 mM sodium acetate [pH 4.6], 1.0 mM ZnCl₂, 60 U of S1 nuclease [Bethesda Research Laboratories, Inc.]) at 37°C for 30 min. The protected fragments were analyzed on a 6% polyacrylamide-urea sequencing gel.

Nucleotide sequence accession number. The DNA sequence has been submitted to GenBank under the accession number M76726.

RESULTS

DNA sequencing of the EHV-1 IR4 gene. To date, two genes in the inverted repeat sequence of EHV-1 have been characterized in some detail. IR1, the only immediate-early gene of EHV-1 (11, 12, 25, 26, 30, 31), encodes a spliced 6.0-kb transcript (Fig. 1) located entirely within the IR sequence (25, 26, 29, 30, 31). IR2 is an early gene encoding a 4.4-kb transcript that is 3' coterminal with IR1 mRNA (Fig. 1) (25, 26, 32). Recently, a second early gene (IR4) has been identified by DNA sequence analysis. The location of IR4, the subclones used to generate a family of overlapping M13 subclones, and the position of IR4 relative to that of the Ori_g are shown in Fig. 1. The IR4 gene lies within the central portion of each inverted repeat. The IR4 DNA sequence, two reiterated sequences (R1 and R2), an origin of replication (3), and a recombination site involved in the formation of the EHV-1 defective interfering particle genome (33, 69) are shown in Fig. 2. The sequence shown consists of 2,250 nucleotides with an overall G+C content of 65%. Analysis of the DNA sequence revealed a major open reading frame (ORF) of 469 amino acids, located between nucleotides 7918 and 9327 of each IR, that could encode a 49.9-kDa protein. The amino terminus of the 469-amino-acid ORF contains five in-frame ATGs located at nucleotides 7919 to 7921, 8054 to 8056, 8447 to 8449, 8492 to 8494, and 8648 to 8650. The third in-frame ATG codon is in the best context according to Kozak's consensus sequence (38). Initiation of translation at the third codon would result in a polypeptide of 293 amino acids with a predicted molecular mass of 32,815 Da. The IR4 ORF exhibits significant homology to the gene products of US1 (ICP22) of HSV-1 (41), ORF4 of EHV-4 (19), RSp40 of pseudorabies virus (PRV) (71), and ORF63 of VZV (22) (Fig. 3). For ICP22 of HSV-1, the potential IR4 protein of 293 amino acids exhibits an overall identity of 16%, whereas the identities with the RSp40 protein of PRV and the ORF63 protein of VZV are 34 and 24%, respectively. For EHV-4, only a partial sequence is known for its ICP22 homolog (19); however, the amino acid identity between this sequence and that of IR4 of EHV-1 exceeds 75%.

Two reiterated sequences (Fig. 1 and 2), designated repeat 1 (R1) and repeat 2 (R2), are located at the 5' end of the IR4 ORF between the first and third in-frame ATGs at nucleotides 7930 to 8048 and 8120 to 8444, respectively. The R1 sequence consists of 7 nucleotides (GGTGGTC) reiterated 17 times, and this repeat sequence would code for 7 amino acids (V-V-G-G-R-W-S) that are reiterated five complete times and once partially. The R2 sequence consists of 25 nucleotides (CCCAGCTCCGGCGACCCCGGCCAG) reiterated 13 times, and R2 would code for 25 amino acids (P-S-S-G-D-P-G-P-A-Q-L-R-R-P-R-P-S-P-A-P-A-T-P-A-Q) that are reiterated four complete times and once partially.

The 5' region of IR4 contains several potential *cis*-acting elements (Table 1). There are two potential TATA boxes (4, 9, 16, 42) located at nucleotides 7795 to 7800 and 8087 to 8192

FIG. 1. Genomic structure of EHV-1 illustrating the location of genes IR1 (IE), IR2, and IR4. (A) The genome consists of unique long sequences (U_L), unique short sequences (U_S), and a pair of inverted repeats (IRs, TRs). TRs refers to the terminal inverted repeat sequence. (B) Map units 0.827 to 0.843 of IRs are expanded, showing relevant IR4 subclones within this region. The nucleotide numbers indicate the distances from the L-S junction, nucleotide 1 being the first nucleotide within the S region. (C) The locations of the IR4-specific transcripts of 1.4 and 1.7 kb are designated by arrows. The Ori_s is shown as a black box, and the repeat elements (R1 and R2) (3) are shown as hatched boxes.

(Fig. 2). The first TATA box is located 118 nucleotides upstream of the major IR4 ORF and conforms to the consensus TATA sequence. The second TATA box is located between the two repeat elements and contains a sequence similar to the TATA consensus sequence. A weight matrix analysis identified two potential cap sites (9) positioned 14 and 19 nucleotides downstream of TATA 1 and TATA 2, respectively (Fig. 2; Table 1). Other potential *cis*-acting elements identified by DNA sequencing include a CAAT box (4, 9, 16, 42) at nucleotides 7658 to 7665 and 7697 to 7704; SP1-binding sites (9, 23, 42) at nucleotides 7772 to 7777, 7554 to 7559, and 8085 to 8063; and an NF1-binding site (28) at nucleotides 7521 to 7534.

A polyadenylation consensus signal (AATAAA) (43) was identified at nucleotides 9509 to 9514 (Fig. 2). This poly(A) signal maps 180 nucleotides downstream of the IR4 termination codon. Sequences immediately downstream of the poly(A) signal conform to those found at consensus poly(A) signals (6, 63, 70). A G+T-rich region (G/T cluster) is found 25 to 54 nucleotides downstream (Fig. 2; Table 1) of the polyadenylation signal (6, 63, 70). Lastly, as reported previously (3), this region of the genome contains the EHV-1 origin of replication (Fig. 1 and 2). In addition, the 8-mer (GAGCGGCG) sequence which is the site of homologous recombination between the IR sequence and the UL3 ORF (ICP27 homolog) of the L region that generates the genome of EHV-1 defective interfering particles (33, 35, 69) is located within the IR4 gene at nucleotides 9027 to 9034.

Northern blot analysis of IR4. Northern blot analysis was performed to identify and characterize mRNAs mapping to the IR4 gene region. A series of subclones from the IR4 region (Fig. 1) was used to probe viral mRNA isolated under immediate-early, early, and late conditions. Clones p1-117, p1-118, p1-119, and p1-132 were 32 P labeled and used to probe Northern blots containing late poly(A)⁺ RNA. Clone p1-132 hybridized to a 1.7-kb transcript (Fig. 4, lane 2), whereas clones p1-118 and p1-119 hybridized to the 1.7-kb transcript and a 1.4 kb-transcript (Fig. 4, lanes 3 and 4).

Clone p1-117 failed to hybridize to either transcript (Fig. 4, lane 1). A subclone of p1-117, p1-128 (Fig. 1), also failed to hybridize to either transcript (data not shown). These results suggest that the 1.4- and 1.7-kb transcripts overlap and initiate downstream of the Ori_s . Both clones p1-132 and p1-117 span the Ori_s ; however, clone p1-117 terminates at the *Hind*III site (nucleotide 7842), whereas clone p1-132 extends to the *Nru*I site (nucleotide 8052) (Fig. 1). Thus, the 5' end of the 1.7-kb transcript appears to map upstream of R1. All clones used in Northern blot analyses failed to hybridize to poly(A)⁺ RNA isolated from mock-infected cells (data not shown). Use of single-stranded M13mp18 clones and synthetic oligonucleotides as probes in Northern blot analyses revealed that both the 1.7- and 1.4-kb transcripts are encoded by the same DNA strand of IR4 DNA and confirmed that both transcripts are transcribed in the same direction as the 469-amino-acid ORF (data not shown).

To determine the kinetic class to which these two transcripts belong, we performed Northern blot analysis by using poly(A)⁺ RNA performed under mock, immediate-early, early, and late conditions. Probes p1-132 and p1-118 were hybridized to mRNA isolated from EHV-1-infected L-M cells maintained in the absence or presence of the appropriate metabolic inhibitor (see Materials and Methods) (Fig. 5A and B, respectively). Clone p1-132 hybridized to the 1.7-kb transcript at early and late times (Fig. 5A, lanes 3 and 4); however, the signal obtained when using probe p1-132 with late mRNA was significantly stronger than that obtained with early mRNA. This observation suggests that the 1.7-kb transcript may belong to the gamma-1 kinetic class. Probe p1-132 failed to hybridize to immediate-early or mock mRNA (Fig. 5A, lanes 1 and 2). Probe p1-118 hybridized to both the 1.7- and 1.4-kb transcripts under both early and late conditions (Fig. 5B, lanes 3 and 4). The 1.4-kb transcript was detected more readily at early times than was the 1.7-kb transcript. Probe p1-118 failed to hybridize to immediate-early or mock mRNA (Fig. 5B, lanes 1 and 2). These data, as well as DNA sequence analyses, suggest that the IR4 ORF

FIG. 2. DNA sequence of the central portion of the IR and the predicted amino acid sequence encoded by the IR4 gene. The nucleotide numbers represent the distance from the L-S junction, nucleotide 1 being the first nucleotide in the S region. The amino acid sequence of IR4 is shown by the single-letter code beneath the DNA sequence. The numbering of the amino acids begins with the first methionine. The five in-frame methionines are underlined. The predicted start site of the 293-amino-acid ORF (third in-frame methionine) is indicated by an arrow. The repeat elements are indicated by dashed lines: repeat 1 (R1) is a 7-mer repeated 17 times, and repeat 2 (R2) is a 25-mer repeated 13 times. The Ori_s is underlined and is located upstream of the repeat elements (3). Potential *cis*-acting elements are shown by the following symbols: Tc_i, transcription initiation site (cap site); Tc_t, transcription termination site; ^, NFI-binding site; 0, GC box; #, CAAT box; X, polyadenylation site; *, G/C cluster. The TATA boxes are underlined for the 1.4-kb mRNA (TATA 2) and the 1.7-kb mRNA (TATA 1). The 8-mer (GAGCGGCG) which is the site of homologous recombination between IR4 and UL3 (ICP27) (69) to generate the genome of EHV-1 defective interfering particles lies at nucleotides 9027 to 9034 and is designated by a @ sign.

FIG. 5. Northern blot analysis of EHV-1 transcripts isolated at immediate-early (IE), early (E), and late (L) times and probed with IR4-specific clones. Immediate-early, early, and late poly(A)⁺ RNAs were isolated at 4, 6, and 8 h postinfection, respectively. The immediate-early mRNA was isolated from cells treated with cycloheximide, and the early mRNA was isolated from cells treated with PAA as described in Materials and Methods. The probes used include clone p1-132 (A) and p1-118 (B) (Fig. 1). The sizes of the transcripts are in kilobases.

[γ -³²P]ATP. This 5'-end-labeled probe hybridized to late mRNA (Fig. 6A, lane 7), as depicted in the schematic at the bottom of Fig. 6. After digestion with S1 nuclease, the protected fragments were fractionated on a 6% polyacrylamide-urea gel. Two major fragments of 239 and 233 nucleotides were detected (Fig. 6A, lane 7). These results mapped the 5' initiation site of the 1.7-kb transcript approximately 14

and 20 nucleotides downstream of the first TATA box (Fig. 2). These two protected fragments may represent two different transcription initiation sites. The 5' initiation site for the 1.7-kb transcript mapped to sequences that conform to a consensus cap site (9) (Fig. 2; Table 1). Probe pCS-1(*Nco*I) failed to hybridize to mock mRNA (Fig. 6A, lane 5). Although clone pCS-1(*Nco*I) failed to hybridize readily to early poly(A)⁺ RNA (Fig. 6A, lane 6), upon overexposure of the gel S1-protected fragments corresponding to those detected by using late mRNA were observed (data not shown). These findings that the synthesis of the 1.7-kb mRNA species is inhibited significantly, but not completely, by PAA suggest that the 1.7-kb mRNA is expressed as a gamma-1 transcript. To map the 5' terminus of the 1.4-kb transcript, clone pCS-1 was 5' end labeled at the *Bam*HI site (Fig. 1). This probe would overlap the 5' ends of both the 1.7- and 1.4-kb mRNAs. The probe was hybridized to both early (isolated from PAA-treated cells) and late mRNA (Fig. 6B, lanes 7 and 8, respectively). Digestion with S1 nuclease resulted in the generation of two major fragments. The largest fragment was approximately 677 nucleotides (Fig. 6B, lane 8) and mapped in close proximity to the first TATA box (Fig. 2). This finding is in agreement with the results obtained by using the pCS-1(*Nco*I) probe. The 677-nucleotide band was detected by using early mRNA upon overexposure of the gel (data not shown). These findings confirm the start site of the 1.7-kb mRNA and are consistent with the conclusion that it is synthesized as a gamma-1 transcript. The second fragment was approximately 370 to 372 nucleotides (Fig. 6B, lanes 7

FIG. 6. S1 nuclease analyses to map the 5' termini of the IR4-specific mRNAs. (A) Clone pCS-1 (Fig. 1) was 5' labeled at its *Nco*I site (*) with T4 kinase and [γ -³²P]ATP. The 5'-end-labeled probe was hybridized to early and late poly(A)⁺ RNAs (lanes 6 and 7, respectively). Early mRNA was isolated from infected cells maintained in the presence of PAA as described in Materials and Methods. The S1-resistant fragments were fractionated on a 6% polyacrylamide-urea sequencing gel. Lanes 1, 2, 3, and 4 contain the DNA sequence of M13mp18 as molecular weight markers. Lanes 5, 6, and 7 represent mock, early, and late poly(A)⁺ RNA, respectively. Lane 8 contains *Hae*III-digested ϕ X174-RF DNA as additional molecular size markers. (B) Clone pCS-1 was 5' end labeled at the *Bam*HI site (Fig. 1) and hybridized to the 1.4- and 1.7-kb IR4 transcripts as depicted below the gel (not to scale). Lanes 6, 7, and 8 represent the mock, early, and late poly(A)⁺ RNAs, respectively. Lane 5 contains the molecular size markers. Molecular sizes are in nucleotides.

FIG. 7. S1 nuclease analysis to map the 3' termini of the IR4-specific mRNAs. Probe pCS-2 was 3' labeled at the *S*tyI site (Fig. 1) with the Klenow fragment and [α - 32 P]dNTP. The 3'-end-labeled clone was hybridized to early (data not shown) and late (lane 6) poly(A)⁺ RNA. The S1 resistant fragment was analyzed on a 6% polyacrylamide-urea sequencing gel. Lanes 1, 2, 3, and 4 contain the DNA sequence of M13mp18, representing molecular size markers. Lane 7 represents a *Hae*III digest of ϕ X174-RF DNA used as additional molecular weight markers. The molecular size markers and the protected fragment are given in base pairs. The schematic at the bottom of the figure depicts the appropriate position of the probe relative to the 1.4- and 1.7-kb transcripts (not to scale).

and 8) and mapped the 5' initiation site of the 1.4-kb transcript to approximately 18 to 20 nucleotides downstream of the second TATA box (Fig. 2). The 5' end of the early 1.4-kb transcript also mapped to sequences representing a consensus cap site (9) (Fig. 2). The observations that these protected fragments were generated in approximately equal amounts from the 1.4-kb mRNA isolated from both PAA-treated and uninhibited infected cells suggest that this transcript is expressed as an early mRNA. The 1.4-kb mRNA was readily detected by S1 nuclease analysis of mRNA isolated at 2 h postinfection, whereas the 1.7-kb mRNA was not readily detected until 4 h postinfection (data not shown). Overall, these data reveal that the IR4 gene encodes two transcripts using two different promoters, such that the mRNAs are differentially expressed under early and late conditions.

DNA sequencing identified a polyadenylation sequence (AATAAA) located 180 nucleotides downstream of the IR4 termination codon. From results of Northern blot experiments, this polyadenylation site was predicted to be the functional polyadenylation site of the IR4 mRNAs. S1 nuclease analysis was used to map the 3' termini of both the 1.4- and 1.7-kb mRNAs. Clone pCS-2 (Fig. 1) was 3' end labeled at the *S*tyI site. The 3'-end-labeled pCS-2 probe hybridized to late mRNA (Fig. 7, lane 6), as shown in the schematic at the bottom of Fig. 7. A single S1 nuclease-resistant fragment of 239 nucleotides was detected after S1 nuclease digestion (Fig. 7, lane 6). The same-sized fragment was observed when using early mRNA isolated from cells treated with PAA (data not shown). A single protected

fragment was obtained in all S1 nuclease analyses when using mRNA isolated at either 6 or 8 h postinfection from both PAA-treated as well as untreated infected cells. All results mapped the transcription termination site to lie approximately 15 to 19 nucleotides downstream of the poly(A) signal (9509 to 9514). The data suggest that IR4 is differentially expressed at early and late times during a lytic infection. Both transcripts appear to possess unique 5' termini, a common 3' terminus, and overlapping ORFs such that the 1.7-kb transcript would encode a protein of 469 amino acids and the 1.4-kb transcript would encode a truncated form of this protein (293 amino acids).

DISCUSSION

In this paper, we present the DNA sequence and transcriptional mapping of the IR4 gene of EHV-1 and demonstrate that it is an ICP22 homolog. Most interestingly, the IR4 gene is transcribed to yield two mRNAs which differ in the regulation of their expression. A 1.4-kb mRNA appears at early times, and its synthesis is not reduced by inhibition of viral DNA replication, indicating that this promoter is regulated as an early promoter. A 1.7-kb mRNA appears later in infection, and its synthesis is reduced greatly in the absence of viral DNA replication, suggesting that it belongs to the gamma-1 class. The 1.4- and 1.7-kb mRNAs may encode overlapping, carboxy-coterminal proteins of 293 and 469 amino acids, respectively. Thus, it will be of interest to determine whether the putative 469-amino-acid polypeptide may have functions different from those of the smaller polypeptide, such as playing a role in the transactivation of specific late genes or possibly in the autoregulation of the IR4 promoter.

Homologs of the US1 (ICP22) gene of HSV-1 have now been identified in VZV (ORF63) (22), PRV (RSp40) (71), EHV-4 (ORF4) (19), and EHV-1 (IR4). These ICP22 homologs are colinear with the US1 gene of HSV-1; however, only the HSV-1 gene lies within the U_s segment and therefore is present in one copy per genome. Also, only the US1 gene of HSV is regulated as an immediate-early gene, although the specific function(s) of the ICP22 phosphoprotein (66) in lytic infection remains to be elucidated in any detail. Sears et al. demonstrated that the ICP22 protein does affect the expression of some late HSV-1 genes (59). ICP22 mutants can be propagated in some cell lines but not in others, suggesting that permissive cells may complement the ICP22 protein by utilizing a host cell factor (1, 50, 59).

A comparison of the five homologous proteins (Fig. 3) revealed that the amino-terminal portion of the proteins contains the majority of conserved residues, whereas the carboxy-terminal portion consists mainly of nonconserved residues. Interestingly, the nonconserved region is not without structural similarities. In each protein, the nonconserved carboxyl region contains an acidic domain that is flanked by regions rich in basic residues, indicating a greater degree of structural conservation than that concluded only from the extent of identical residues.

A common feature of alphaherpesviruses is the presence of G+C-rich reiterated sequences located within the intergenic region of the IR sequences (19, 22, 65, 71). An exception to this feature is the reiterated sequences located within an intron at the 5' end of the US1 ORF of HSV-1 and HSV-2 (65). Unlike the reiterated sequences of US1, the IR4 reiterated sequences map within the IR4 major ORF and, as shown by S1 nuclease analysis, are not removed by splicing of the 1.4- or 1.7-kb mRNAs. Thus, these reiterated se-

quences may be translated to yield a repetitive amino acid sequence within the putative 469-amino-acid protein. The precise function of these reiterated sequences is not known; however, it has been proposed that repetitive sequences within the IR segment are involved with its expansion and contraction by a nonhomologous crossover event (21, 65).

Significant features of the IR4 sequence include the presence of numerous potential *cis*-acting elements. The potential promoter element for the 1.7-kb transcript which is expressed as a gamma-1 transcript includes a TATA box, two GC boxes, and two CAAT boxes that map upstream of the major IR4 ORF which would encode the 469-amino-acid polypeptide. The promoter that controls expression of the early 1.4-kb transcript consists of a TATA box and a potential GC box embedded within the 469-amino-acid ORF. Downstream of each promoter is an in-frame translation initiation codon, and a single polyadenylation signal lies downstream of the 469-amino-acid ORF. Transcriptional mapping data revealed two mRNA species, and the placement of these transcripts and the location of potential ORFs within the translated IR4 DNA sequence suggest that the IR4 gene may encode two in-frame, overlapping proteins of 293 and 469 amino acids. Indeed, the S1 mapping results confirmed that the early 1.4-kb and the late 1.7-kb mRNAs have unique 5' termini and are 3' coterminal. S1 nuclease analysis with the pCS-1(*Nco*I) probe mapped the 5' transcription initiation site of the 1.7-kb mRNA to lie 14 to 20 nucleotides downstream of TATA 1. The presence of two S1-resistant fragments (Fig. 6A) suggests that the 1.7-kb mRNA may utilize two transcriptional cap sites. Alternatively, this may be the result of aberrant cleavage by the S1 nuclease. Use of the probe pCS-1(*Bam*HI), which overlaps both TATA sequences, yielded S1-resistant fragments of 370 to 372 and 677 bp in S1 analyses (Fig. 6B), confirming that the 1.7- and 1.4-kb mRNAs have unique 5' termini. Finally, S1 mapping of these mRNAs identified a single transcription termination site downstream of a polyadenylation signal sequence, indicating that the 1.4- and 1.7-kb mRNAs are 3' coterminal. Thus, the data suggest that IR4 expresses two mRNAs that have distinct 5' ends, share a common 3' terminus, and may encode overlapping proteins. The generation of mRNA families with common 3' and distinct 5' termini in a single transcriptional unit is a phenomenon observed for various members of the herpesvirus family (13, 32, 39, 52).

The 1.4-kb mRNA could encode a polypeptide of 293 amino acids that is in frame and carboxy coterminal with the larger potential protein of 469 amino acids encoded by the 1.7-kb late transcript. Interestingly, the region of the potential IR4 protein of 469 amino acids that is most closely identical in amino acid sequence to ICP22 and to the ICP22 homologs of VZV, PRV, and EHV-4 is contained entirely within the smaller potential protein of 293 amino acids. This 293-amino-acid ORF contains three potential translation start codons, the first having the most favorable context for initiation of translation according to Kozak's scanning model for translation (38). Initiation of protein synthesis from the first ATG would produce a polypeptide of 293 amino acids with a predicted molecular mass of 32,815 Da, a size comparable to the products of RSp40 of PRV (71), ORF63 of VZV (22), and ORF4 of EHV-4 (19).

In contrast, the 1.7-kb mRNA has the potential to encode one of three proteins. The first is a 469-amino-acid protein (predicted molecular weight of 49,890), which would have repetitive amino acid sequences at the amino-terminal end. These repetitive sequences would not be present in the putative 32.8-kDa protein and could allow the 49.9-kDa

protein to have functions different from those of the 32.8-kDa protein. Other examples of herpesvirus proteins with repetitive amino acid sequences have been reported (15, 52). The second protein is the 32.8-kDa protein, which could be produced if the ribosome were to scan down the 1.7-kb mRNA and initiate translation at the third in-frame ATG. The third protein is a protein of 160 amino acids, which is present in the small ORF which lies at the 5' end of the translated IR4 sequence and is out of frame with the major IR4 ORF. This potential protein would not be produced by translation of the 1.4-kb transcript and would consist of amino acids with repetitive sequences. Further work is necessary to characterize the product(s) of the 1.7-kb mRNA.

The late 1.7-kb mRNA has a cap site 287 nucleotides upstream of the cap site of the 1.4-kb transcript and initiates in close proximity to the Ori_L. Since the late mRNA is present in greatly reduced amounts when viral DNA replication is inhibited, the temporal shift appears to be dependent on viral DNA replication. Similar shifts in transcription initiation sites have been reported for the early mRNAs of several DNA viruses such as human cytomegalovirus (13), JC virus (37), simian virus 40 (10, 24), and adenovirus (49). The factors controlling these shifts remain unclear. It has been proposed that changes in chromosome structure as a result of viral DNA replication and the coordinate synthesis of new regulatory proteins may result in the expression of these new transcription initiation sites (10, 14).

Finally, we have determined the precise location of homologous recombination that occurs between EHV-1 sequences near the L terminus and those in the IR sequence to generate the genome of defective interfering particles of EHV-1 (68). An 8-bp sequence located near the L terminus within the UL3 ORF (ICP27 homolog) and within the IR4 ORF is involved in a homologous recombination event that results in the generation of a novel ORF within the defective interfering particle genome (68). Interestingly, the majority of the IR4 ORF is conserved in this novel ORF, including the region which is homologous to ICP22. Thus, an understanding of IR4 function might provide insight into the possible importance of this novel ORF located in the genome of EHV-1 defective interfering particles that mediate persistent infection (20, 27, 45, 46, 56).

ACKNOWLEDGMENTS

We are grateful to Suzanne Zavecz, Bridget Higginbotham, and Scarlett Flowers for excellent technical assistance.

This investigation was supported by Public Health Service research grant AI 22001 from the National Institutes of Health, a Grayson-Jockey Club Research Foundation Inc. research grant, and grant 89-37266-4735 from the U.S. Department of Agriculture Animal Molecular Biology Program.

REFERENCES

1. Ackermann, M., M. Sarmiento, and B. Roizman. 1985. Application of antibody to synthetic peptides for characterization of the intact and truncated $\alpha 22$ -protein specified by herpes simplex virus 1 and the R325 alpha-deletion mutant. *J. Virol.* **56**:207-215.
2. Allen, G. P., and J. T. Bryans. 1986. Molecular epizootiology, pathogenesis, and prophylaxis of equine herpes virus type 1 infections. *Prog. Vet. Microbiol. Immunol.* **2**:78-144.
3. Baumann, R. P., R. R. Yalamanchili, and D. J. O'Callaghan. 1989. Functional mapping and DNA sequence of an equine herpesvirus type 1 origin of replication. *J. Virol.* **63**:1275-1283.
4. Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon. 1980. The ovalbumin gene-sequence of putative control regions. Nu-

- cleic Acids Res. 8:127-142.
5. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
 6. Birnstiel, M. L., M. Busslinger, and K. Strub. 1985. Transcription termination and 3' processing: the end is in site. *Cell* 41:349-359.
 7. Breeden, C. A., and D. J. O'Callaghan. Unpublished data.
 8. Bryans, J. T., and G. P. Allen. 1986. Equine viral rhinopneumonitis. *Rev. Sci. Tech. O.I.E.* 5:837-847.
 9. Bucher, P. 1990. Weight matrix description of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* 212:563-578.
 10. Buchman, A. R., M. Fromm, and P. Berg. 1984. Complex regulation of simian virus 40 early-region transcription from different overlapping promoters. *Mol. Cell. Biol.* 4:1900-1914.
 11. Caughman, G. B., A. T. Robertson, W. L. Gray, D. C. Sullivan, and D. J. O'Callaghan. 1988. Characterization of equine herpesvirus type 1 immediate early proteins. *Virology* 163:563-571.
 12. Caughman, G. B., J. Staczek, and D. J. O'Callaghan. 1985. Equine herpesvirus type 1 infected cell polypeptides: evidence for immediate-early, early, and late regulation of viral gene expression. *Virology* 145:49-61.
 13. Chang, C., C. 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.
 14. Chen, W., and K. Struhl. 1988. Saturation mutagenesis of yeast his3 "TATA element": genetic evidence for a specific TATA-binding protein. *Proc. Natl. Acad. Sci. USA* 85:2691-2695.
 15. Chou, J., and B. Roizman. 1990. The herpes simplex virus 1 gene for ICP34.5, which maps in inverted repeats, is conserved in several limited-passage isolates but not in strain 17syn+. *J. Virol.* 64:1014-1020.
 16. Coen, D. M., S. P. Weinheimer, and S. L. McKnight. 1986. A genetic approach to promoter regulation during trans induction of viral gene expression. *Science* 234:53-58.
 17. Cohen, J. C., M. L. Perdue, C. C. Randall, and D. J. O'Callaghan. 1977. Herpesvirus transcription: altered regulation induced by FUDR. *Virology* 76:621-633.
 18. Cohen, J. C., C. C. Randall, and D. J. O'Callaghan. 1975. Transcription of equine herpesvirus type 1: evidence for classes of transcripts differing in abundance. *Virology* 68:561-565.
 19. Cullinane, A. A., F. J. Rixon, and A. J. Davison. 1988. Characterization of the genome of equine herpesvirus 1 subtype 2. *J. Gen. Virol.* 69:1575-1590.
 20. Dauenhauer, S. A., R. A. Robinson, and D. J. O'Callaghan. 1982. Chronic production of defective-interfering particles by hamster embryo cultures of herpesvirus persistently infected and oncogenically transformed cells. *J. Gen. Virol.* 60:1-14.
 21. Davison, A. J., and D. J. McGeoch. 1986. Evolutionary comparisons of the S segments in the genomes of herpes simplex type 1 and varicella-zoster virus. *J. Gen. Virol.* 67:597-611.
 22. Davison, A. J., and J. E. Scott. 1986. The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* 67:1759-1816.
 23. Dynan, W. S., and R. Tjian. 1983. The promoter-specific transcriptional factor SP1 binds to upstream sequences in the SV40 early promoter. *Cell* 35:79-87.
 24. Ghosh, P. K., and P. Lebowitz. 1981. Simian virus 40 early mRNAs contain multiple 5' termini upstream and downstream from a Hogness-Goldberg sequence: a shift in 5' termini during the lytic cycle is mediated by large T antigen. *J. Virol.* 40:224-240.
 25. Gray, W. L., R. P. Baumann, A. T. Robertson, G. B. Caughman, D. J. O'Callaghan, and J. Staczek. 1987. Regulation of equine herpesvirus type 1 gene expression: characterization of immediate early, early, and late transcription. *Virology* 158:79-87.
 26. Gray, W. L., R. P. Baumann, A. T. Robertson, D. J. O'Callaghan, and J. S. Staczek. 1987. Characterization and mapping of equine herpesvirus type 1 immediate early, early, and late transcripts. *Virus Res.* 8:233-244.
 27. Gray, W. L., R. Yalamanchili, B. Reangsakulrach, R. B. Baumann, J. Staczek, and D. J. O'Callaghan. 1989. Viral transcripts in cells infected with defective interfering particles of equine herpesvirus type 1. *Virology* 172:1-10.
 28. Gronostajski, R. M., S. Adhya, K. Nagata, R. A. Guggenheimer, and J. Hurwitz. 1985. Site-specific DNA binding of nuclear factor I: analysis of cellular binding sites. *Mol. Cell. Biol.* 5:964-971.
 29. Grundy, F. J., R. P. Baumann, and D. J. O'Callaghan. 1989. DNA sequence and comparative analysis of the equine herpesvirus type 1 immediate-early gene. *Virology* 172:223-236.
 30. Harty, R. N., C. F. Colle, F. J. Grundy, and D. J. O'Callaghan. 1989. Mapping the termini and intron of the spliced immediate-early transcript of equine herpesvirus 1. *J. Virol.* 63:5101-5110.
 31. Harty, R. N., C. F. Colle, and D. J. O'Callaghan. 1991. Equine herpesvirus type 1 gene regulation: characterization of transcription from the IE gene region in a productive infection, p. 319-337. *In* E. K. Wagner (ed.), *Herpesvirus transcription and its regulation*. CRC Press, Inc., Boca Raton, Fla.
 32. Harty, R. N., and D. J. O'Callaghan. 1991. An early gene maps within and is 3' coterminal with the immediate-early gene of equine herpesvirus 1. *J. Virol.* 65:3829-3838.
 33. Harty, R. N., R. R. Yalamanchili, and D. J. O'Callaghan. 1991. Transcriptional analysis of the UL1 gene of equine herpesvirus 1: a gene conserved in the genome of defective-interfering particles. *Virology* 183:830-833.
 34. Henry, B. E., R. A. Robinson, S. A. Dauenhauer, S. S. Atherton, G. S. Hayward, and D. J. O'Callaghan. 1981. Structure of the genome of equine herpesvirus type 1. *Virology* 115:97-114.
 35. Holden, V. R., R. N. Harty, R. R. Yalamanchili, and D. J. O'Callaghan. Unpublished data.
 36. Huang, H. L., J. M. Szabocsik, C. C. Randall, and G. A. Gentry. 1971. Equine abortion (herpes) virus-specific RNA. *Virology* 45:381-389.
 37. Khalili, K., L. Feigenbaum, and G. Khoury. 1987. Evidence for a shift in 5'-termini of early viral RNA during the lytic cycle of JC virus. *Virology* 158:469-472.
 38. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229-241.
 39. Liu, F., and B. Roizman. 1991. The promoter, transcriptional unit, and coding sequence of herpes simplex virus 1 family 35 proteins are contained within and in frame with the UL26 open reading frame. *J. Virol.* 65:206-212.
 40. Maniatis, T., E. R. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 41. McGeoch, D. J., A. Dolan, S. Donald, and F. J. Rixon. 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. *J. Mol. Biol.* 181:1-13.
 42. McKnight, S. L., R. C. Kingsbury, A. Spence, and M. Smith. 1984. The distal transcription signals of the herpesvirus tk gene share a common hexanucleotide control sequence. *Cell* 37:253-262.
 43. Nordstorm, J. L., S. L. Hall, and M. M. Kessler. 1985. Polyadenylation of sea urchin histone RNA sequences in transfected COS cells. *Proc. Natl. Acad. Sci. USA* 82:1094-1098.
 44. O'Callaghan, D. J., G. P. Allen, and C. C. Randall. 1978. Structure and replication of equine herpesvirus, p. 1-32. *In* J. T. Bryans and H. Gerber (ed.), *Equine infectious diseases IV*. Veterinary Publishers, Princeton, N.J.
 45. O'Callaghan, D. J., G. A. Gentry, and C. C. Randall. 1983. The equine herpesviruses, p. 215-318. *In* B. Roizman (ed.), *The herpesviruses*, vol. 2. Comprehensive virology, series 2. Plenum Publishing Corp., New York.
 46. O'Callaghan, D. J., B. E. Henry, J. H. Wharton, S. A. Dauenhauer, R. B. Vance, J. Staczek, and R. A. Robinson. 1981. Equine herpesviruses: biochemical studies on genomic structure, DI particles, oncogenic transformation and persistent infection, p. 387-418. *In* Y. Becker (ed.), *Developments in molecular virology*, vol. 1. Herpesvirus DNA. Martinus Nijhoff, The Hague, The Netherlands.
 47. O'Callaghan, D. J., J. M. Hyde, G. A. Gentry, and C. C. Randall. 1968. Kinetics of viral deoxyribonucleic acid, protein, and infectious particle production and alterations in host mac-

- romolecular synthesis in equine abortion (herpes) virus-infected cells. *J. Virol.* 2:793-804.
48. O'Callaghan, D. J., D. C. Sullivan, R. P. Baumann, G. B. Caughman, C. C. Flowers, A. T. Robertson, and J. Staczek. 1984. Genomes of the equine herpesviruses: molecular structure, regions of homology and DNA sequences associated with transformation. *UCLA Symp. Mol. Cell. Biol.* 21:507-525.
 49. Osborne, T. F., and A. J. Berk. 1983. Far upstream initiation sites for adenovirus early region 1A transcription are utilized after the onset of viral DNA replication. *J. Virol.* 45:594-599.
 50. 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 is not essential for growth. *Cell* 25:227-232.
 51. Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
 52. Rixon, F. J., and D. J. McGeoch. 1984. A 3' co-terminal family of mRNAs from the herpes simplex virus type 1 short region: two overlapping reading frames encode unrelated polypeptides one of which has a highly reiterated amino acid sequence. *Nucleic Acids Res.* 12:2473-2487.
 53. Robertson, A. T., R. P. Baumann, J. Staczek, and D. J. O'Callaghan. 1988. Molecular characterization of the gene products of the short region of the equid herpesvirus-1 genome, p. 132-139. *In* D. G. Powell (ed.), *Equine infectious diseases V. Proceedings of the Fifth International Conference*. University Press of Kentucky, Lexington.
 54. Robertson, A. T., G. B. Caughman, W. L. Gray, R. P. Baumann, J. Staczek, and D. J. O'Callaghan. 1988. Analysis of the in vitro translation products of the equine herpesvirus type 1 immediate early gene mRNA. *Virology* 166:451-462.
 55. Robinson, R. A., P. W. Tucker, S. A. Dauenhauer, and D. J. O'Callaghan. 1981. Molecular cloning of equine herpesvirus type 1 DNA: analysis of standard and defective viral genomes and viral sequences in oncogenically transformed cells. *Proc. Natl. Acad. Sci. USA* 78:6684-6688.
 56. Robinson, R. A., R. B. Vance, and D. J. O'Callaghan. 1980. Oncogenic transformation by equine herpesviruses. II. Coestablishment of persistent infection and oncogenic transformation of hamster embryo cells by equine herpesvirus type 1 preparations enriched for defective interfering particles. *J. Virol.* 36:204-219.
 57. Ruyechan, W. T., S. A. Dauenhauer, and D. J. O'Callaghan. 1982. Electron microscopic study of equine herpesvirus type 1 DNA. *J. Virol.* 42:297-300.
 58. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
 59. Sears, A. E., I. W. Halliburton, B. Meignier, S. Silver, and B. Roizman. 1985. Herpes simplex virus 1 mutant deleted in the α 22 gene: growth and gene expression in permissive and restrictive cells and establishment of latency in mice. *J. Virol.* 55:338-346.
 60. Soehner, R. L., G. A. Gentry, and C. C. Randall. 1965. Some physicochemical characteristics of equine abortion virus nucleic acid. *Virology* 26:394-405.
 61. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
 62. Weaver, R. F., and C. Weissman. 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: 5' termini of 15-S beta-globin messenger RNA precursor and mature 10-S beta globin messenger RNA have identical map coordinates. *Nucleic Acids Res.* 7:1175-1193.
 63. Weiss, E. A., G. M. Gilmartin, and J. R. Nevins. 1991. Poly(A) site efficiency reflects the stability of complex formation involving the downstream element. *EMBO J.* 10:215-219.
 64. Whalley, J. M., G. R. Robertson, and A. J. Davison. 1981. Analysis of the genome of equine herpesvirus type 1: arrangement of cleavage for restriction endonuclease *EcoRI*, *BglII*, and *BamHI*. *J. Gen. Virol.* 57:307-323.
 65. Whitton, L., and J. B. Clements. 1984. The junction between the repetitive and the short unique sequences of the herpes simplex virus genome are determined by the polypeptide-coding regions of two spliced immediate-early mRNA's. *J. Gen. Virol.* 65:451-466.
 66. Wilcox, K. W., A. Kohn, E. Sklyanskaya, and B. Roizman. 1980. Herpes simplex virus phosphoproteins. I. Phosphate cycles on and off some viral polypeptides and can alter their affinity for DNA. *J. Virol.* 33:167-182.
 67. Yalamanchili, R. R., V. R. Holden, R. N. Harty, and D. J. O'Callaghan. Unpublished data.
 68. Yalamanchili, R. R., and D. J. O'Callaghan. 1990. Sequence and organization of the genomic termini of equine herpesvirus type 1. *Virus Res.* 15:149-162.
 69. Yalamanchili, R. R., B. Raengsakulrach, R. P. Baumann, and D. J. O'Callaghan. 1990. Identification of the site of recombination in the generation of the genome of DI particles of equine herpesvirus type 1. *Virology* 175:448-455.
 70. Zarkower, D., and M. Wickens. 1987. Formation of mRNA 3' termini: stability and dissociation of a complex involving the AAUAAA sequence. *EMBO J.* 6:177-186.
 71. Zhang, G., and D. P. Laeder. 1990. The structure of the pseudorabies virus genome at the end of the inverted repeat sequences proximal to the junction with the short unique region. *J. Gen. Virol.* 71:2433-2441.