JOURNAL OF VIROLOGY, Jan. 1991, p. 206–212
0022-538X/91/010206-07$02.00/0

Vol. 65, No. 1

The Promoter, Transcriptional Unit, and Coding Sequence of Herpes Simplex Virus 1 Family 35 Proteins Are Contained within and in Frame with the U_L26 Open Reading Frame

FENYONG LIU AND BERNARD ROIZMAN*

The Marjorie B. Kovler Viral Oncology Laboratories, The University of Chicago, Chicago, Illinois 60637

Received 31 July 1990/Accepted 2 October 1990

The herpes simplex virus 1 (HSV-1) genome specifies an abundant capsid protein which in denaturing gels forms multiple bands designated family 35 proteins (D. K. Braun, B. Roizman, and L. Pereira, J. Virol. 49:142–153, 1984). Nucleotide-sequencing studies have assigned the coding sequences of these proteins to the open reading frame U_L26 (D. J. McGeoch, M. A. Dalrymple, A. J. Davidson, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor, J. Gen. Virol. 69:1531–1574, 1988). In studies reported here, a series of plasmid constructs containing deletions or insertions of an α4 promoter or of a sequence encoding a cytomegalovirus epitope reacting with a mouse monoclonal antibody revealed the following: the open reading frame previously designated U_L26 encodes two proteins which share amino acid sequences, and each coding domain is contained in its own transcriptional unit that terminates at a common, unique poly(A) site. On the basis of the transcription initiation site (+1), it was predicted that the U_L26 open reading frame encodes a protein of 635 amino acids, and a protein with an apparent molecular weight of approximately 75,000 has been identified. The second transcriptional unit, designated $U_L26.5$, predicted to specify a protein of 329 amino acids, encodes the family 35 proteins; it is transcribed by an mRNA which initiates at approximately nucleotide +1000 of the U_L26 transcription initiation site and is translated from the methionine initiation codon located at position +1099 of the U_L26 transcriptional unit. The DNA fragment comprising the sequences downstream of the *Hpa*I cleavage site (+832 of U_L26) contains both the promoter and the coding sequence of family 35 proteins and is both competent and efficient in expressing the proteins in transfected cells superinfected with HSV-1 or HSV-2.

Earlier, this laboratory reported a set of herpes simplex virus 1 (HSV-1) capsid proteins of particular interest. Thus, Gibson and Roizman (6, 7) described the virion protein VP22A which was present in capsids containing DNA and which had not been enveloped. This protein was replaced by VP22 in virions; however, capsids obtained by stripping the envelope from virions did not contain VP22. VP22A and VP22 shared properties, suggesting that they might be related as precursor and product, respectively (6, 7). In later reports, Braun et al. (2, 3) described an abundant set of proteins which reacted with several monoclonal antibodies. These proteins, designated infected-cell protein 35 (ICP35), formed as least four major and several minor bands in one-dimensional denaturing polyacrylamide gels and numerous spots in two-dimensional gels (2, 3). The polypeptides were phosphorylated, and their electrophoretic mobilities after a pulse and after a chase suggested that posttranslational processing included cleavage (3). Analyses of capsids extracted from infected cells indicated that ICP35 is present on the surfaces of full but not empty capsids (3). Although it has been suggested that VP22 and ICP35 are related, experiments designed to determine whether this is the case have not been done. VP22 and ICP35 are of particular interest with respect to their roles in the assembly and maturation of HSV capsids inasmuch as the processing of the proteins and the variability in capsid content suggests that these proteins may play an important role in capsid maturation.

One approach to the determination of the function of these proteins is through specific mutagenesis of their genes, but the gene specifying ICP35 has not been mapped precisely. Braun et al. (2) mapped the gene specifying ICP35 on the basis of analyses of HSV-1 × HSV-2 recombinants to a sequence located between the genes specifying thymidine kinase (U_L23) and glycoprotein B (U_L27). On the basis of the mapping of a *ts* mutation which precludes processing of a set of proteins analogous to ICP35 at the nonpermissive temperature (16), McGeoch et al. (13) concluded that the coding sequences of ICP35 are located in open reading frame U_L26. The predicted product of U_L26, however, is considerably larger than ICP35. In this report we show that the domain of U_L26 contains two transcriptional units yielding proteins which share amino acid sequence. ICP35 is encoded by the 3′ domain of U_L26, which we have designated $U_L26.5$.

MATERIALS AND METHODS

Virus and cells. The properties of HSV-1(F) and HSV-2(G), the prototype HSV-1 and HSV-2 strains used in this laboratory, were described elsewhere (5, 17). The maintenance and propagation of thymidine kinase-minus baby hamster kidney (designated BHK) and Vero cells have been described previously (1, 17). The BHK cells were used for analyses of polypeptides specified by transfected DNA fragments, whereas the Vero cells were used for preparation of RNA.

Monoclonal antibodies. Monoclonal antibodies used in these studies were obtained from Lenore Pereira. The monoclonal antibody H725 reacts with ICP35 of HSV-1 but not with HSV-2 proteins (2, 3). CH28-2, a gift of L. Pereira, is a monoclonal antibody directed against human cytomegalovirus (CMV) glycoprotein A. The epitope of this antibody has

* Corresponding author.

been mapped to a 20-amino-acid peptide, N-KGQKPNLL DRLRHRKNGYRH-C, by assaying the reactivity of a series of overlapping peptides synthesized according to the predicted nucleotide sequence of the protein (19).

Construction of plasmids. pRB4026 was constructed by insertion of the HSV-1 *Kpn*I S fragment into the *Kpn*I site of pUC18. pRB4057 contains the entire encoding sequence of the U_L26 open reading frame, extending 3′ from nucleotide −900 relative to the translation initiation site of the gene to approximately 650 nucleotides downstream from its polyadenylation site. pRB4060 was constructed by replacing the viral DNA sequence 23 bp upstream from the translation initiation site of the U_L26 open reading frame in pRB4057 with the HSV-1 DNA *Bam*HI Z fragment. The *Bam*HI Z fragment contains at one terminus portions of the 5′ transcribed noncoding sequences of the α4 gene starting with nucleotide +33 and the upstream untranscribed domains of this gene. The *Bam*HI Z fragment was inserted in the orientation that would juxtapose the α4 promoter in the proper transcriptional orientation to the intact and truncated domains of the U_L26 open reading frame. pRB4056, pRB4058, pRB4087, and pRB4093 were derived from pRB4057, and pRB4088 and pRB4089 were derived from pRB4060 by generating deletions by the subcloning techniques described by Sambrook et al. (18). Two pairs of oligonucleotides, i.e., oligonucleotide A (5′-AAGGGACAG AAGCCCAACCTGCTAGACCGACTGCGACACCGCA AAAACGGGTACCGACAC-3′) with its complement and oligonucleotide B (5′-AAAGGGACAGAAGCCCAACCTGC TAGACCGACTGCGACACCGCAAAAACGGGTACCGA CACGA-3′) with its complement, were synthesized in Applied Biosystems DNA Synthesizer 380A (Foster City, Calif.). Each oligonucleotide and its complement encode the epitope of the CH28-2 monoclonal antibody and contain a *Kpn*I site at the 3′ end for convenient screening of the oligonucleotide insertion in plasmids. pRB4079 and pRB4080 were derived by inserting the oligonucleotide A sequence into the unique *Hpa*I site of pRB4057 and pRB4060, respectively. pRB4092 was derived by inserting the oligonucleotide B sequence into the unique *Mst*II site of pRB4060. pRB4094, pRB4095, pRB4096, and pRB4102 were derived from pRB4092 by generating deletions by means of common subcloning techniques (18). All insertion sites of these plasmids were sequenced to verify that the CMV epitope was inserted in the same frame as the U_L26 open reading frame.

Transfections and superinfection of cells transfected with plasmid DNAs. All experiments were done in six-well Costar (Cambridge, Mass.) dish cultures of BHK cells (approximately 10^6 cells per well). Transfections were done as described previously (11) except that each well was transfected with 10 μg of plasmid DNA.

In most experiments, the transfected cells were exposed for 18 to 20 h posttransfection to 10 PFU of HSV-1(F) or HSV-2(G) per cell as stated below in Results. After 2 h of adsorption, the inoculum was replaced with Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. At 20 h postinfection, the cells were labeled for 2 h with 50 μCi of [^{35}S]methionine (Dupont, NEN Research Products) in 1 ml of the medium consisting of mixture 199 lacking methionine but supplemented with 1% calf serum. The cells were then harvested, washed once with phosphate-buffered saline, pelleted by centrifugation in a Sorvall RC_2B GSA rotor at 4,000 rpm for 5 min, suspended in disruption buffer (0.05 M Tris [pH 7.0], 8.5% [vol/vol] sucrose, 5% [vol/vol] β-mercaptoethanol, 2% [vol/vol] sodium dodecyl sulfate), sonicated for 20 s in ice, and boiled for 1 min before electrophoretic separation in denaturing gels.

Electrophoretic separation and staining of infected-cell proteins with monoclonal antibody. The infected-cell polypeptides were denatured, separated on 9.5% (vol/vol) polyacrylamide gels containing sodium dodecyl sulfate, cross-linked with *N,N′*-diallyltartardiamide, transferred electrically to nitrocellulose membranes, and reacted with the monoclonal antibody H725 against HSV-1 ICP35 protein or CH28-2 against the CMV epitope, as described previously by Braun et al. (2).

Isolation and S1 analysis of cytoplasmic RNA. Cytoplasmic RNA was purified as described previously by Jenkins and Howett (10) from Vero cells mock infected or infected with 20 PFU of HSV-1(F) per cell and maintained for 12 h. HSV-1 DNA probe (0.02 pmol) described in Fig. 1 was 5′ end labeled with [γ-^{32}P]ATP (Dupont, NEN Research Products), hybridized to 50 μg of total cytoplasmic RNA, digested with S1 nuclease, and separated on 7% polyacrylamide gels in the presence of 8 M urea (10).

RESULTS

Experimental design and construction of plasmids. The major objective of these studies was to map the nucleotide sequences that encode ICP35. This was done in two series of experiments. The first series mapped the 5′ termini of cytoplasmic RNA homologous to the open reading frame U_L26. These studies, detailed below, revealed the presence of two overlapping transcriptional units with different transcription initiation sites and methionine initiation codons. The second set of experiments dealt with identification of the protein products and verification of the reading frame for each transcriptional unit. This was done by analysis of the proteins specified by a series of plasmids carrying the wild-type and derivative forms of the *Eco*NI-*Pst*I DNA fragment located between 0.325 and 0.355 map units in the HSV-1(F) genome in the prototype orientation (Fig. 1, lines 1 to 3). Three series of plasmid constructs were made.

The first series included plasmid constructs designated A through I in Fig. 1. Collectively, the HSV-1 sequences carried in these plasmids contain deletions which encompass the entire domain of the U_L26 open reading frame. In some instances, e.g., plasmid constructs B and D, we juxtaposed the promoter of the α4 gene in the form of the *Bam*HI Z fragment to force a higher level of transcription.

The second series included plasmid constructs designated J through N in Fig. 1. The HSV DNAs in the plasmids of this group carry a sequence encoding a CMV epitope inserted into the unique *Mst*II site of these fragments. In all but one plasmid construct we also inserted the *Bam*HI Z fragment to augment transcription by the α4 promoter contained in this fragment. The exception was plasmid construct N.

The last series comprised plasmid constructs O and P. In this group the CMV epitope was inserted into the unique *Hpa*I site and only the P plasmid contains the α4 promoter in the form of the *Bam*HI Z fragment.

Nucleotide sequences of U_L26 are contained in two transcriptional units. The nucleotide sequence of the region surrounding the U_L26 open reading frame shows a single poly(A) site (13). To map the transcripts of U_L26, we prepared two probes. Probe 1, designed to identify the 5′ terminus of U_L26 mRNA, consisted of the *Eco*NI-*Bam*HI fragment labeled at the *Bam*HI site, whereas probe 2 consisted of the *Xcm*I-*Bst*EII fragment labeled at the *Bst*EII site

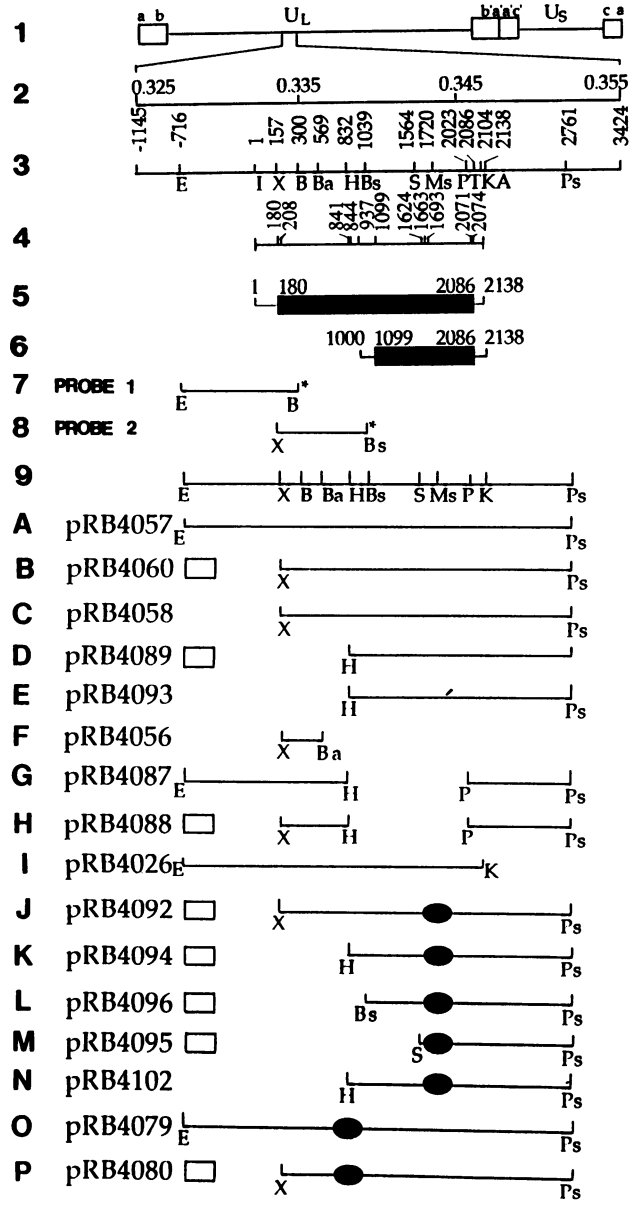

FIG. 1. Sequence arrangement of HSV-1 genome showing the positions of U_L26 and $U_L26.5$ open reading frames and their transcripts and of the DNA probes and plasmids constructed for these studies. Line 1, Schematic representation of the sequence arrangement of HSV-1 genome. U_L and U_S refer to the long and short unique sequences flanked by the terminal inverted repeats ab b'a' and a'c' ca, respectively. Line 2, Genome map position. Line 3, Nucleotide numbers relative to the approximate transcription initiation site of U_L26 (I, nucleotide +1) and restriction endonuclease sites of the HSV-1 *Eco*NI-*Pst*I DNA fragment. The line also shows the position of the translational termination codon (T) and of the single poly(A) signal (A) which serves both the U_L26 and $U_L26.5$ RNAs. Line 4, Locations of the methionine codons contained in the U_L26 open reading frame. Lines 5 and 6, Nucleotide numbers for transcription initiation, translation initiation and termination, and the poly(A) signal for U_L26 and $U_L26.5$, respectively. All numbers shown are relative to the transcription initiation site of the U_L26 mRNA. The thick bar represents the coding sequences. Lines 7 and 8, Locations of the DNA probes used to map the 5' termini of the mRNAs of U_L26 and $U_L26.5$ open reading frames, respectively. The asterisk identifies the terminus of the probes that was labeled with ^{32}P. Line 9, Restriction endonuclease map drawn to scale with reference to lines A through P, which are schematic representations of the HSV-1 sequences contained in the plasmid constructs used in the studies described in this report. The construction of the plasmids is described in Materials and Methods. The vector for all plasmids was derived form pUC18. The *Bam*HI Z fragment used as the source of the α4 gene promoter and inserted in the proper transcriptional orientation relative to that of the U_L26 open reading frame is shown as an open rectangle. The CMV epitope is shown as a filled oval. The restriction endonuclease sites were as follows: B, *Bam*HI; Ba, *Bal*I; Bs, *Bst*EII; E, *Eco*NI; H, *Hpa*I; K, *Kpn*I; Ms, *Mst*II; P, *Pml*I; Ps, *Pst*I; S, *Sal*I; X, *Xcm*I.

(Fig. 1, lines 7 and 8). The results of the S1 analyses shown in Fig. 2 indicate the following.

(i) Cytoplasmic RNA hybridized to probe 1 protected a fragment approximately 300 nucleotides long (Fig. 2, lane 3). Although the precise transcription initiation site was not identified and could vary by as much as 5%, for convenience we designated nucleotide 300 upstream from the *Bam*HI site as nucleotide +1 of U_L26. It is noteworthy that sequences corresponding to a TATAA box are not present in the vicinity of the +1 nucleotide. The first methionine codon after the approximate transcription initiation site is located at position +180.

(ii) Cytoplasmic RNA hybridized to probe 2 yielded two sets of fragments protected from S1 digestion (Fig. 2, lane 9). The first fragment contained all of the HSV-1 DNA sequences (lane 9, band T), as would be expected, since the transcript detected with probe 1 would anneal to all of the HSV-1 sequences in probe 2. The second set of protected fragments formed several bands ranging from 35 to 40 nucleotides in length (lane 2, band $U_L26.5$). Thus the transcription initiation site of this transcript was approximately at nucleotide +1000 relative to nucleotide +1 of U_L26. A TATAA box was present at +976, i.e., 24 nucleotides upstream, and the first methionine codon downstream from the transcription initiation site of this RNA was at position +1099.

For convenience, we refer to the longer RNA as U_L26 RNA and to the RNA detected by probe 2 as $U_L26.5$ RNA.

Location of the gene specifying ICP35. In order to localize the position of the coding domain of the gene specifying ICP35, a series of deletions in open reading frame U_L26 were tested for their capacity to express ICP35. In this series of experiments, BHK cells were transfected with plasmids containing the construct A, B, C, D, E, F, G, H, or I (Fig. 1) and then infected with HSV-2. Analyses of the electrophoretically separated, electrically transferred polypeptides with the monoclonal antibody H725 specific to HSV-1 ICP35 showed (Fig. 3) the following.

(i) H725 reacts uniquely with HSV-1 ICP35, as evidenced by the absence of reactivity with electrophoretically separated polypeptides from HSV-2-infected cells in lane 11.

(ii) All plasmid constructs except F, H, G, and I specified a series of bands which comigrated with HSV-1 ICP35 and reacted with the H725 monoclonal antibody. The shortest fragment which yielded HSV-1 ICP35 was plasmid E (Fig. 3, lane 2). Since this plasmid construct was expressed from its own endogenous promoter, the results indicate that the sequences contained in the *Hpa*I-*Pst*I fragment contain both the coding sequences and the promoter of the gene encoding ICP35. The HSV-1 sequences in plasmid construct E contain all of the sequences of $U_L26.5$ RNA plus 168 nucleotides upstream from the transcription initiation site of this RNA.

(iii) The expression of HSV-1 ICP35 required all of the

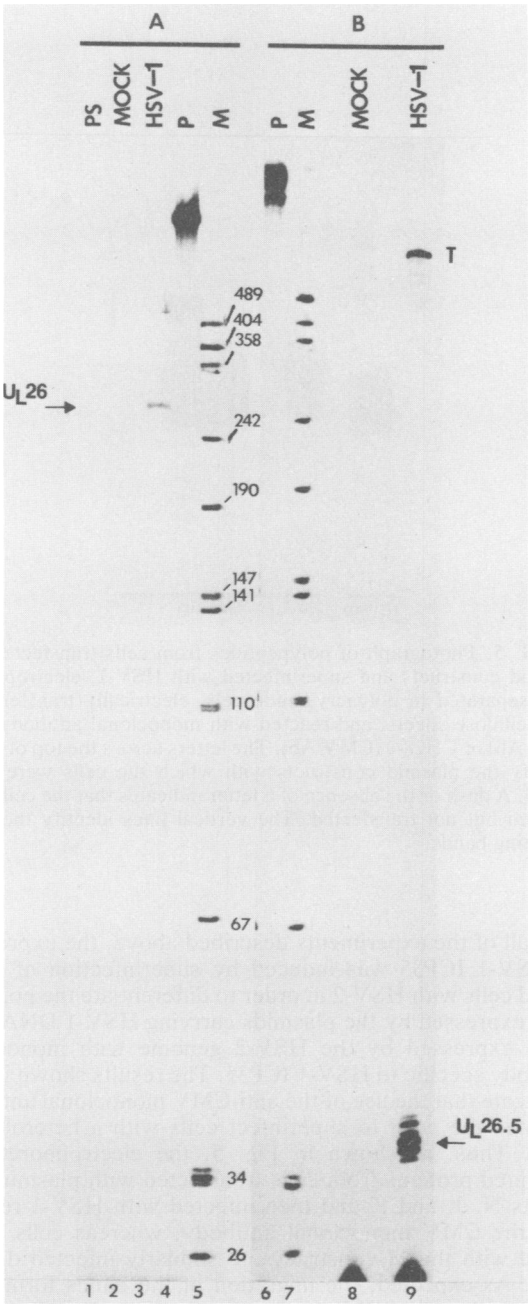

FIG. 2. Autoradiographic image of DNA probe 1 (A) and probe 2 (B) hybridized to total cytoplasmic RNA from mock-infected and 12-h-infected Vero cells and digested with S1 nuclease. The RNAs were prepared as described in Materials and Methods. Lane 1 (PS), S1-digested probe 1 under the same condition of hybridization and digestion as those shown in lanes MOCK and HSV-1; lanes 2 and 8 (MOCK), RNA extracted from cells 12 h after mock infection; lanes 3 and 9 (HSV-1), RNA extracted from cells infected with HSV-1(F) and maintained for 12 h; lanes 4 and 7 (P), positions of the undigested probes (probe 1 or 2); lanes 5 and 7 (M), 5′ end-labeled fragments obtained from digestion of pGEM3Z DNA with *Msp*I. Arrows indicate the protected 5′ termini of U_L26 (A) and $U_L26.5$ (B) RNAs. T, Position of the HSV-1 sequences in probe 2 protected by the U_L26 RNA.

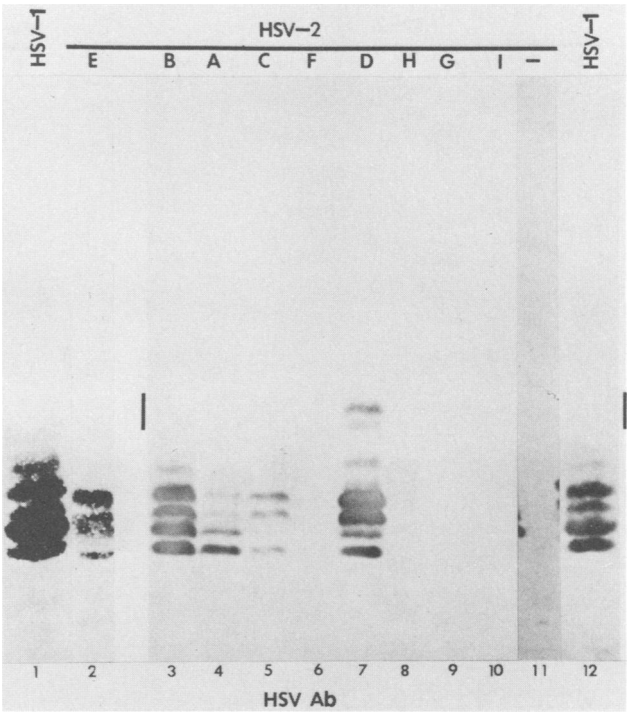

FIG. 3. Photograph of polypeptides from cells transfected with plasmid constructs and superinfected with virus, electrophoretically separated in polyacrylamide gels, electrically transferred to a nitrocellulose sheet, reacted with monoclonal antibody H725 to HSV-1 ICP35 (HSV Ab), and stained with goat anti-mouse immunoglobulin antibody coupled to peroxidase. Experimental details were described in Materials and Methods. The letters across the top of the gel identify the plasmid constructs with which the cells were transfected. A dash or the absence of a letter indicates that the cells were infected but not transfected. The vertical lines identify the slow-migrating bands.

sequences contained between *Hpa*I and *Pml*I restriction endonuclease sites (+832 to +2023, respectively) and the sequences downstream from the *Kpn*I site (+2104). The poly(A) site which serves the U_L26 open reading frame is at position +2138 (13).

(iv) The electrophoretically separated proteins in lane 7 contain additional, slower-migrating bands that reacted with the H725 monoclonal antibody. We should note that in plasmid construct D, the α4 promoter was fused to nucleotide +832 of the U_L26 open reading frame. In this fragment, the methionine codons suitable for initiation of translation would be at positions +841 and +937. Translation initiation at those positions could give rise to the slower-migrating bands in lane 7.

ICP35 open reading frame overlaps and is in frame with U_L26. In the preceding section we showed that the sequences which encode ICP35 correspond to those represented in $U_L26.5$ RNA and overlap the domain of the U_L26 open reading frame. As shown in Fig. 1, line 4, the reported sequence of the U_L26 open reading frame (13) predicts several methionine codons. To determine whether the translation of ICP35 is initiated from an internal codon of the U_L26 open reading frame or whether it represents another open reading frame, we inserted a CMV epitope into the *Mst*II restriction endonuclease site. Nucleotide sequence analyses (not shown) indicated that the CMV epitope is in

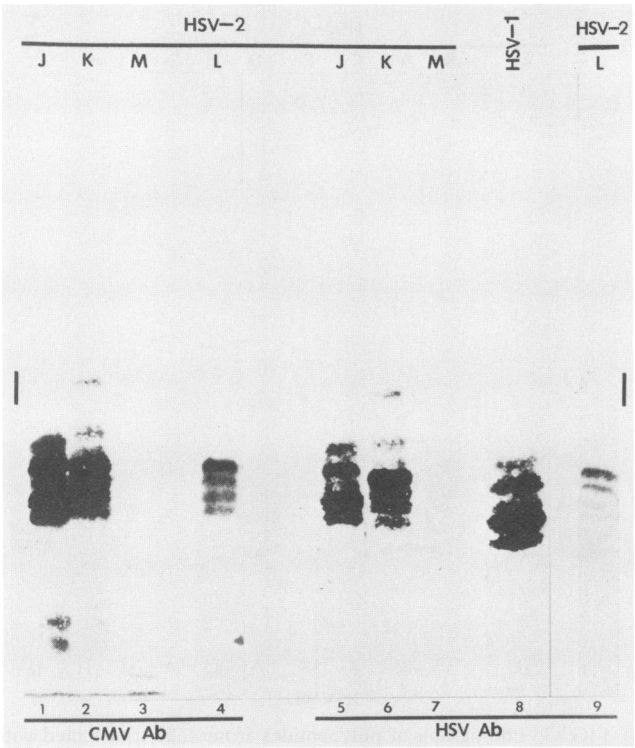

FIG. 4. Photograph of polypeptides from cells transfected with plasmid constructs and superinfected with virus, electrophoretically separated in polyacrylamide gels, electrically transferred to nitrocellulose sheets, and reacted with monoclonal antibody H725 (HSV Ab) or CH28-2 (CMV Ab). The letters across the top of the gel identify the plasmid constructs with which the cells were transfected. A dash or the absence of a letter indicates that the cells were infected but not transfected. The vertical lines identify the slow-migrating bands.

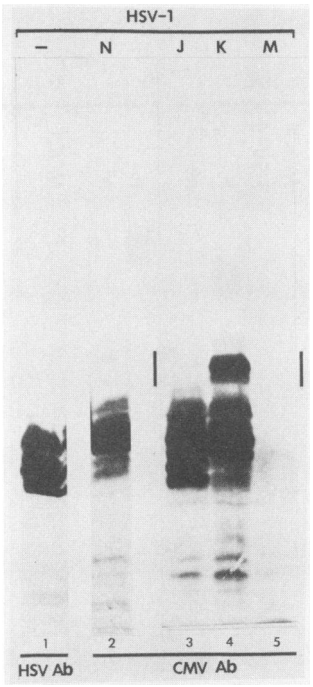

FIG. 5. Photograph of polypeptides from cells transfected with plasmid constructs and superinfected with HSV-1, electrophoretically separated in polyacrylamide gels, electrically transferred to nitrocellulose sheets, and reacted with monoclonal antibody H725 (HSV Ab) or CH28-2 (CMV Ab). The letters across the top of the gel identify the plasmid constructs with which the cells were transfected. A dash or the absence of a letter indicates that the cells were infected but not transfected. The vertical lines identify the slow-migrating bands.

frame with U_L26. Figure 4 shows that BHK cells transfected with construct J, K, or L made a family of proteins which reacted with both anti-HSV-1 ICP35 (H725) and anti-CMV (CH28-2) monoclonal antibodies. The formation of the characteristic four ICP35 bands by the products of transfection of plasmid construct L indicates that the initiating methionine codon for ICP35 is at position 1099. We should note that in plasmid M the first methionine codon available for initiation would be at position +1624 and the protein product could be expected to consist of approximately 154 HSV and 21 CMV amino acids. However, the methionine codon at position +1624 is in the wrong context and not likely to function as an initiating codon very efficiently. We should also note that a slower-migrating band reactive with both monoclonal antibodies was present in lysates of cells transfected with plasmid construct K (lanes 2 and 6). In this instance, too, the *Hpa*I-*Pst*I fragment was fused to an α4 promoter and the nearest available methionine initiation codons would be at positions +841 and +937.

We conclude from these results that the $U_L26.5$ coding sequences specifying ICP35 constitute a part of and are in frame with those of U_L26. In the preceding section we have shown that ICP35 could be expressed by transactivation of the DNA sequences contained in the *Hpa*I-*Pst*I fragment. Because plasmid construct E could be transactivated by HSV-2, we also conclude that the coding sequences of U_L26 include both the coding sequences and the promoter domain of the gene specifying ICP35.

In all of the experiments described above, the expression of HSV-1 ICP35 was induced by superinfection of transfected cells with HSV-2 in order to differentiate the polypeptides expressed by the plasmids carrying HSV-1 DNA from those expressed by the HSV-2 genome with monoclonal antibody specific to HSV-1 ICP35. The results shown in Fig. 5 indicate that the use of the anti-CMV monoclonal antibody obviates the need to superinfect cells with a heterologous virus. Thus, as shown in Fig. 5, the electrophoretically separated proteins from cells transfected with plasmid constructs N, J, and K and then infected with HSV-1 reacted with the CMV monoclonal antibody, whereas cells transfected with the M construct and similarly infected did not react. As expected, the migration of the bands formed by ICP35 without the CMV epitope were faster than those with the insertion. In this instance, too, the *Hpa*I-*Pst*I fragment (plasmid construct N) expressed ICP35. We should also note that as in Fig. 4, the cells transfected with construct K specified a slower-migrating protein band reactive with the CMV antibody. As noted above, this band may represent the induction of transcription and translation from a methionine codon upstream from the initiating codon for the ICP35.

Intact U_L26 open reading frame specifies a protein which migrates more slowly than ICP35. In the preceding section we showed that the sequences encoding ICP35 overlapped only a portion of the sequence designated the U_L26 open reading frame. The purpose of the studies described below was to identify the product of the full-length U_L26 open reading frame. BHK cells were transfected with plasmid

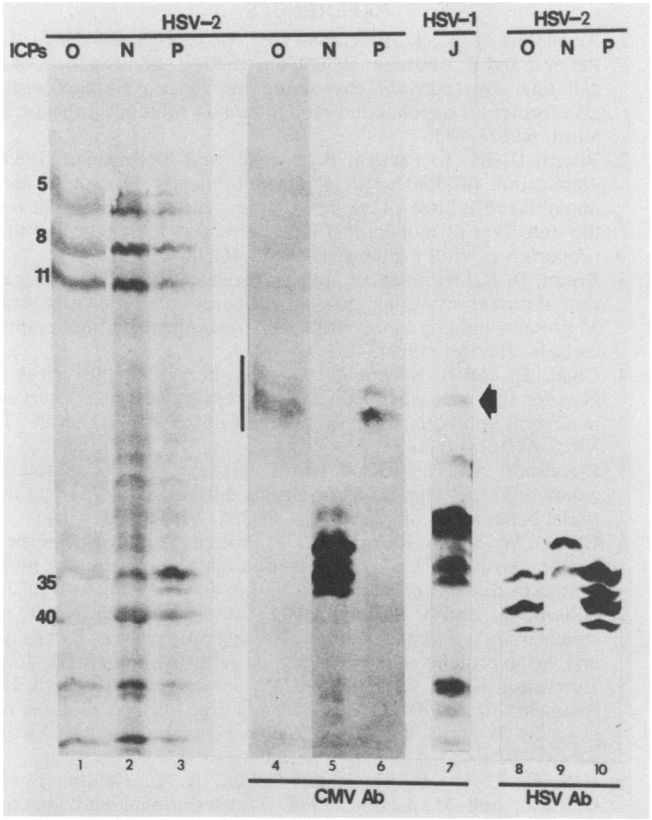

FIG. 6. Autoradiographic images and photograph of polypeptides from cells transfected with plasmid constructs and superinfected with virus, electrophoretically separated in polyacrylamide gels, electrically transferred to nitrocellulose sheets, and reacted with monoclonal antibody H725 (HSV Ab) or CH28-2 (CMV Ab). The letters across the top of the gel identify the plasmid constructs with which the cells were transfected. The vertical line and arrow identify the product of the U_L26 protein. Lanes 1, 2, and 3 are autoradiographic images of proteins labeled with [^{35}S]methionine as described in Materials and Methods. The infected-cell proteins (ICPs) of HSV-2 were numbered according to Morse et al. (14). Lanes 8, 9, and 10 show lysates of the same cells as shown in lane 4, 5, and 6 but stained with H725 rather than with the CH28-2 monoclonal antibodies. Lane 7 shows the lysate of cells transfected with plasmid construct J, in which U_L26 is driven by the α4 promoter, infected with HSV-1(F) and maintained at 39°C. Under these conditions, only α and a few β proteins are expressed, but the ICP35 of HSV-1(F) is not expressed. The ICP35 encoded by the plasmid construct is expressed inasmuch as the transfected gene is regulated as a β gene (12).

construct O, N, or P (Fig. 1) and then infected with HSV-2. The electrophoretically separated proteins from cells transfected with plasmids O, N, and P were reacted with monoclonal antibodies to either HSV-1 (H725) or CMV (CH28-2) (Fig. 6, lanes 4 through 6 and 8 through 10) and then autoradiographed to provide molecular weight markers (lanes 1 through 3). The salient features of the results were as follows.

(i) The plasmid constructs O and P containing the CMV epitope inserted in frame with U_L26 specified proteins which formed two bands with electrophoretic mobilities corresponding approximately to proteins with apparent molecular weights of 75,000 to 78,000 (Fig. 6, lanes 4 and 6). The CMV monoclonal antibody did not react with ICP35 bands produced by plasmids O and P (lanes 4 and 6), as was expected, since the CMV epitope was inserted at the *Hpa*I restriction endonuclease site (+832), i.e., before the translation initiation site of ICP35 at position +1099.

(ii) All plasmid constructs made ICP35, which reacted with H725 monoclonal antibody against HSV-1 ICP35. The disparity in the electrophoretic mobilities of the ICP35 proteins made by plasmid constructs N and P reflect the insertion into plasmid construct N of the oligonucleotide encoding an additional 21 amino acids.

(iii) The abundance of proteins specified by the entire U_L26 open reading frame relative to that of the ICP35 protein may be deduced from the observation that while both the M_r 75,000 to 78,000 proteins and the ICP35 react with the same monoclonal antibody, H725, the reactivity or amount of the larger protein is considerably lower than that observed for ICP35.

(iv) Both Braun et al. (2, 3) and Zweig et al. (20) reported that monoclonal antibody to ICP35 reacted with a protein band with an apparent molecular weight of 76,000 (2, 3) to 80,000 (20), as would be expected if both ICP35 and the larger protein shared amino acid sequences. The significance of these observations at the time they were published was not apparent. The compelling evidence that the two proteins share amino acid sequences is based on the observation that construct J under conditions of overproduction of the U_L26 proteins yielded proteins which comigrated with both the larger protein and ICP35 and reacted with the CH28-2 monoclonal antibody (arrow, Fig. 6, lane 7).

DISCUSSION

In this report we present evidence that the sequences contained in the open reading frame U_L26 are contained in two transcriptional units. The larger unit specified a protein of approximately 75,000 apparent molecular weight. This protein formed a doublet in denaturing polyacrylamide gels. The second transcriptional unit yielded several immunologically related polypeptides which constitute the family 35 proteins formally designated ICP35. The amino acid sequence of ICP35 overlaps that of the protein with the higher (75,000-molecular-weight) protein. From the immunologic reactivity of the two sets of proteins we may deduce that the amounts of ICP35 produced are greatly in excess of those of the larger protein. Several features of the design and results should be noted.

(i) On completion of the studies, we noted that the designation by McGeoch et al. (13) of open reading frame U_L26 as the gene specifying ICP35 did not take into account the report by Holland et al. (9) which described the existence of two transcripts corresponding approximately in size to the two RNA transcripts described in this study. Holland et al. (9) did not identify the protein products of these transcripts. It is also of interest that both Zweig et al. (20) and Braun et al. (3) reported that a monoclonal antibody to ICP35 also reacted with a protein of approximately 76,000 to 80,000 apparent molecular weight. Although these data could have been interpreted to indicate that the two proteins arose from overlapping transcripts and shared amino acid sequences, they also lent themselves to alternative explanations.

(ii) The primary objective of our studies, that is, to map the open reading frame which specifies ICP35, could have been accomplished in a number of ways. The procedure we chose is similar to that used in another study from this laboratory (4) in that we inserted an epitope to a well-characterized monoclonal antibody into the sequence pre-

dicted to express the target protein for which a specific antibody was available. In this instance, the sequence of U_L26 was predicted from the sequencing studies of McGeoch et al. (13). The simplest protocol for mapping the coding sequences of ICP35 was to use the nucleotide sequence to insert the epitope in frame with that of the predicted open reading frame. Our mapping of the 5′ terminus of $U_L26.5$ coupled with the demonstration of the reactivity of proteins made by plasmids N and L with both the CMV and HSV monoclonal antibodies indicated that ICP35 maps in only a portion of open reading frame U_L26 and that the initiating codon for ICP35 is at position +1099. We have designated the coding sequence of ICP35 as the $U_L26.5$ open reading frame to differentiate it from the U_L26 open reading frame of McGeoch et al. (13).

The demonstration that U_L26 specifies a protein of 75,000 to 78,000 apparent molecular weight (M_r 75,000 protein) coupled with (a) the evidence that a plasmid carrying the CMV epitope in the $U_L26.5$ domain yields both the M_r 75,000 protein and ICP35, which react with the CMV monoclonal antibody CH28-2, and (b) the earlier reports that monoclonal antibodies to ICP35 react with a protein of 75,000 apparent molecular weight are compelling evidence that U_L26 specifies the M_r 75,000 protein, that the coding sequences of ICP35 and the M_r 75,000 protein overlap, and that the two proteins share amino acid sequences. The sequence of U_L26 is concordant with this conclusion.

(iii) A striking feature of the results is the observation that the plasmid constructs transfected into BHK cells and transactivated with either the homologous HSV-1 or the heterologous HSV-2 make readily detectable quantities of ICP35. We estimate that the quantities made are, in terms of antigenic mass, of the same order of magnitude as those expressed in productive infection by the gene resident in the viral genome. Of particular interest is the *HpaI-PstI* fragment, which contains only 168 bp 5′ of the transcription initiation site. The $U_L26.5$ promoter may turn out to be particularly strong and useful for expression of genes within the context of the viral genome.

(iv) The key features of the overlapping open reading frames, i.e., the sharing of amino acid sequences, the disparity in the amount of each protein made, and the location of the promoter regulatory domain of the shorter protein within the coding sequences of the largest protein, have not been reported previously for HSV genomes. A similar situation does exist, however, in the case of the pre-S1, pre-S2, and S genes for the hepatitis B virus (8, 15). In the case of hepatitis B virus, the functions of the three partially overlapping proteins are related. The specific functions of the M_r 75,000 and ICP35 proteins are not known. Inasmuch as the locations of the coding sequences are now known, studies on the functions of these proteins should be forthcoming.

ACKNOWLEDGMENTS

We thank Richard J. Roller for invaluable advice, Amy E. Sears for making available the sequence of the CMV epitope prior to its publication, and Lenore Pereira for the monoclonal antibodies.

These studies were aided by Public Health Service grants from the National Cancer Institute (CA47451) and the National Institute for Allergy and Infectious Diseases (AI124009 and AI1588).

REFERENCES

1. **Arsenakis, M. J., J. Hubenthal-Voss, G. Campadelli-Fiume, L. Pereira, and B. Roizman.** 1986. Construction and properties of a cell line constitutively expressing the herpes simplex virus glycoprotein B dependent on functional α4 protein synthesis. J. Virol. **60**:674–682.
2. **Braun, D. K., L. Pereira, B. Norrild, and B. Roizman.** 1983. Application of denatured, electrophoretically separated, and immobilized lysates of herpes simplex virus-infected cells for the detection of monoclonal antibodies and for studies of the properties of viral proteins. J. Virol. **46**:103–112.
3. **Braun, D. K., B. Roizman, and L. Pereira.** 1984. Characterization of post-translational products of herpes simplex virus gene 35 proteins binding to the surfaces of full capsids but not empty capsids. J. Virol. **49**:142–153.
4. **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.
5. **Ejercito, P. M., E. D. Kieff, and B. Roizman.** 1968. Characterization of herpes simplex virus strains differing in their effect on social behavior of infected cells. J. Gen. Virol. **2**:357–364.
6. **Gibson, W., and B. Roizman.** 1972. Proteins specified by herpes simplex virus. VIII. Characterization and composition of multiple capsid forms of subtypes 1 and 2. J. Virol. **10**:1044–1052.
7. **Gibson, W., and B. Roizman.** 1974. Protein specified by herpes simplex virus. Staining and radiolabeling properties of B capsids and virion proteins in polyacrylamide gels. J. Virol. **13**:155–165.
8. **Heermann, K. H., U. Goldmann, W. Schwartz, T. Seyffarth, H. Baumgarten, and W. Gerlich.** 1984. Large surface proteins of hepatitis B virus containing the pre-S sequence. J. Virol. **52**:396–402.
9. **Holland, L. E., R. M. Sandri-Goldin, A. L. Goldin, J. C. Glorioso, and M. Levine.** 1984. Transcriptional and genetic analyses of the herpes simplex virus type 1 genome: coordinates 0.29 to 0.45. J. Virol. **49**:947–959.
10. **Jenkins, F. J., and M. K. Howett.** 1984. Characterization of the mRNAs mapping in the *BglII* N fragment of the herpes simplex virus type 2 genome. J. Virol. **52**:99–107.
11. **Kristie, T. M., and B. Roizman.** 1984. Separation of sequences defining basal expressing from those conferring α gene recognition within the regulatory domains of herpes simplex virus 1 α genes. Proc. Natl. Acad. Sci. USA **81**:4065–4069.
12. **Liu, F., and B. Roizman.** Unpublished data.
13. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**:1531–1574.
14. **Morse, L. S., L. Pereira, B. Roizman, and P. A. Schaffer.** 1978. Anatomy of herpes simplex virus (HSV) DNA. X. Mapping of viral genes by analysis of polypeptides and functions specified by HSV-1 × HSV-2 recombinants. J. Virol. **26**:389–410.
15. **Neurath, A. R., S. B. H. Kent, N. Strick, P. Taylor, and C. E. Stevens.** 1985. Hepatitis B virus contains pre-S gene-encoded domains. Nature (London) **315**:154–156.
16. **Preston, V. G., J. A. V. Coates, and F. J. Rixon.** 1983. Identification and characterization of a herpes simplex virus gene product required for encapsidation of virus DNA. J. Virol. **45**:1056–1064.
17. **Roizman, B., and P. G. Spear.** 1968. Preparation of herpes simplex virus of high titer. J. Virol. **2**:83–84.
18. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Sears, A. E., and L. Pereira.** Unpublished data.
20. **Zweig, M., C. J. Heilman, Jr., H. Rabin, and B. Hamper.** 1980. Shared antigenic determinants between two distinct classes of proteins in cells infected with herpes simplex virus. J. Virol. **35**:644–652.