JOURNAL OF VIROLOGY, Oct. 1991, p. 5149–5156
0022-538X/91/105149-08$02.00/0

Vol. 65, No. 10

The Herpes Simplex Virus 1 Gene Encoding a Protease Also Contains within Its Coding Domain the Gene Encoding the More Abundant Substrate

FENYONG LIU AND BERNARD ROIZMAN*

The Marjorie B. Kovler Viral Oncology Laboratories, The University of Chicago, 910 East 58th Street, Chicago, Illinois 60637

Received 25 April 1991/Accepted 20 June 1991

The herpes simplex virus 1 open reading frames U_L26 and $U_L26.5$ are 3′ coterminal. The larger, U_L26 open reading frame encodes a protein approximately 80,000 in apparent molecular weight and contains the promoter and coding sequence of the $U_L26.5$ gene, which specifies a capsid protein designated infected cell protein 35. The larger product contains in its entirety the amino acid sequence of the smaller protein. We report that the U_L26 gene encodes a protease which catalyzes its own cleavage and that of the more abundant product of $U_L26.5$. By inserting the coding sequence of an epitope to a cytomegalovirus monoclonal antibody and homologs of the immunoglobulin G binding domain of staphylococcal protein A into the 3′ termini of the coding domains of the two open reading frames, we identified both products of the cleavage and determined that the cleavage site is approximately 20 amino acids from the carboxyl termini of both proteins.

In this report, we describe the identification of a herpes simplex virus (HSV) protease and its substrate. This is the first herpesvirus protease identified to date. The remarkable properties of this protease are that (i) it catalyzes its own cleavage as well as that of a more abundant, smaller substrate and (ii) the gene encoding the smaller substrate, although transcribed and translated independently of the protease, is nevertheless entirely contained within the gene encoding the latter. The relevant background and experimental design of this study may be summarized as follows.

In the course of studies on the proteins forming the herpes simplex virus capsid, it was noted that a protein present in empty capsids lacking DNA, designated virion protein (VP22), was replaced in full capsids containing DNA by an apparently smaller, faster-migrating protein with similar characteristics and designated VP22A (8, 9). In the course of enumeration of infected cell proteins, VP22 and its putative product, VP22A, were designated the infected cell protein 35 (ICP35) (3, 4). ICP35 has been extensively studied: it consists of a family of highly heterogeneous proteins which form numerous bands in both one- and two-dimensional separations. Pulse-chase experiments suggested that the major translational products of the gene encoding these proteins form bands designated ICP35c and -d, whereas the post-translationally modified products form bands ICP35e and -f (4). It has recently been suggested that ICP35 functions as a scaffolding protein (15).

In the exposition of the nucleotide sequence of the HSV-1 genome, McGeoch et al. (14) assigned ICP35 to the open reading frame U_L26. The assignment was based on the observation that a temperature-sensitive mutation affecting the processing of ICP35 and accumulation of mature capsids mapped in the U_L26 open reading frame (18). However, when studies on this gene were initiated (12), it became apparent that this open reading frame predicted to encode 635 amino acids was much too big to encode ICP35 with an apparent molecular weight of approximately 45,000. Our studies (12) showed that U_L26 consists of two open reading frames, each encoded by its own transcript (Fig. 1, lines 4 and 5). Thus, U_L26 encodes a protein approximately 80,000 in molecular weight. In anticipation of the results to be presented in this report, we shall refer to this protein as the protease, or Pr. The smaller reading frame, schematically shown in Fig. 1, line 5, was designated $U_L26.5$. The two open reading frames share all of the amino acid sequences encoded in $U_L26.5$, and the promoter of the smaller open reading frame is contained in the 5′ coding and noncoding domains of U_L26.

In this report, we show that the product of U_L26 is a protease. In designing these studies, we took advantage of three observations: (i) a monoclonal antibody derived by Lenore Pereira and analyzed jointly in our laboratories reacts with the HSV-1 products of both U_L26 and $U_L26.5$ but not with those of the HSV-2, (ii) $U_L26.5$ is abundantly expressed in transfected cells, and (iii) HSV-1 strain F [HSV-1(F)] is temperature sensitive in the α4 gene (10), and at the nonpermissive temperature it induces α gene promoters (17) and expresses primarily α genes (1). We noted that HSV-1(F) at that temperature transactivates the U_L26 and $U_L26.5$ open reading frames transfected into the same cells but not its own U_L26 or $U_L26.5$ open reading frames (12). To systematically analyze the site and attribute the cleavage reaction to the product of the U_L26 open reading frame, we used a movable epitope, that is, a sequence encoding 20 amino acids specifying the epitope of a monoclonal antibody against a human cytomegalovirus (CMV) protein. This epitope was inserted wherever necessary to trace the products of the reaction.

MATERIALS AND METHODS

Virus and cells. The properties of HSV-1(F) and HSV-2(G), the prototype HSV-1 and HSV-2 strains, respectively, used in this laboratory and the maintenance and propagation of the thymidine kinase-minus baby hamster kidney (BHK) cells have been described previously (1, 6, 20).

Monoclonal antibodies. Monoclonal antibodies H725 and

* Corresponding author.

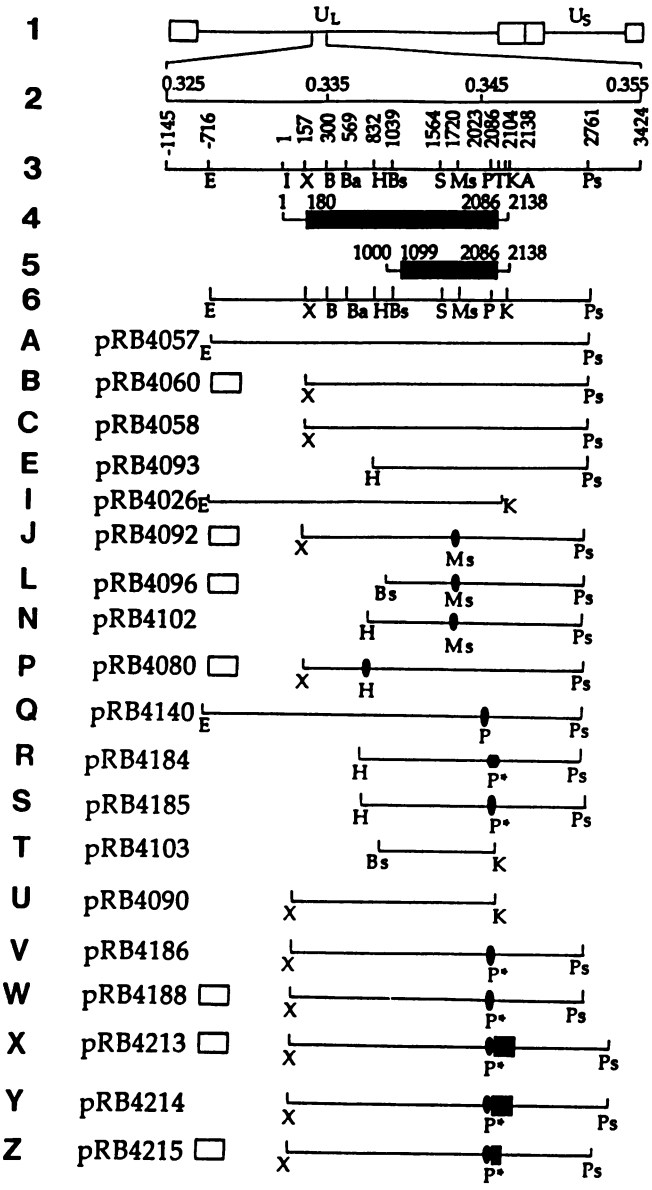

FIG. 1. Sequence arrangement of the HSV-1 genome, positions of the U_L26 and $U_L26.5$ open reading frames and their transcripts, and structures of the test plasmids constructed for these studies. Line 1, schematic representation of the sequence arrangement of the HSV-1 genome. U_L and U_S refer to the long unique and short unique sequences flanked by terminal inverted repeats shown as open rectangles. Lines 2 and 3, genome map positions, nucleotide numbers relative to the approximate transcription initiation site of U_L26 indicated by letter I at +1, and restriction endonuclease sites of the HSV-1 *Eco*NI-*Pst*I DNA fragment. Also shown are the positions of the translational termination codon (T) and the single poly(A) signal (A) which serve both the U_L26 and $U_L26.5$ RNAs. Lines 4 and 5, coding domains of the U_L26 and $U_L26.5$ open reading frames (filled rectangles). The numbers refer to the positions of the transcription initiation site, the translation initiation and termination codons, and the poly(A) signal for both open reading frames relative to nucleotide +1 of U_L26. Line 6, restriction endonuclease map drawn to scale with reference to lines A through Z, which are schematic representations of the HSV-1 sequences contained in the plasmid constructs used in the studies described in this report. Construction of the plasmids shown schematically in lines A through Z is described in Materials and Methods. The source of the α4 promoter (open rectangle) shown in plasmids B, J, L, P, W, X, and Z was a *Bam*HI Z DNA fragment (17) inserted in proper transcriptional orientation. The CMV epitope is shown as a filled oval. Oligonucleotide C with its complement is shown as a filled hexagon, and the new created *Pml*I site is marked (P*). Restriction endonuclease sites: 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.

CH28-2 to ICP35 and CMV glycoprotein B, respectively, were obtained from Lenore Pereira and have been described previously (3, 4, 12).

Construction of plasmids. Construction of plasmids I (pRB4026), A (pRB4057), C (pRB4058), B (pRB4060), E (pRB4093), L (pRB4096), N (pRB4102), and P (pRB4080) has been described previously (12). Two short double-stranded DNA fragments were made by synthesizing complementary strands of each in an Applied Biosystems 380B DNA synthesizer and annealing the two strands. Sequence A (5′-AGGGACAGAAGCCCAACCTGCTAGACCGACTG CGACACCGCAAAAACGGGTACCGACAC-3′) encodes the epitope of monoclonal antibody CH28-2 (12, 16). Plasmid Q (pRB4140) was derived by inserting oligonucleotide A with its complement into the unique *Pml*I site of plasmid A. Sequence C (5′-TCGACGTTGACACGGCCCGCGCCGCC GATTTGTTCGTCTCTCAGATGATGGGGGCCCGCC ACGTGTGA-3′) encodes the authentic U_L26 sequence from the *Pml*I site to the translation termination site but with the addition of a new *Pml*I site between the carboxyl-terminal amino acid and the stop codon of the U_L26 open reading frame. In addition, the sequence GTG at the authentic *Pml*I site has been changed to GTC. Insertion of sequence C into the unique *Pml*I site of the U_L26 open reading frame in the proper orientation resulted in the creation of a new *Pml*I site between the carboxyl-terminal amino acid and the stop codon of U_L26. The original authentic *Pml*I site was destroyed without changing the U_L26 amino acid sequence since the codons GTG and GTC encode the same amino acid. Therefore, the net effect of the insertion of oligonucleotide C with its complement to the U_L26 open reading frame was that U_L26 had two additional amino acids between its authentic carboxyl-terminal amino acid and its stop codon which were encoded by the new created *Pml*I recognition sequence. Plasmid R (pRB4184) was derived by inserting sequence C into the *Pml*I site of plasmid E (pRB4093), and plasmid S (pRB4185) was derived by inserting sequence A into the *Pml*I site of plasmid R. Plasmids T, U, V, and W (pRB4103, pRB4090, pRB4186, and pRB4188) were derived from plasmids B and S. Plasmids X, Y, and Z (pRB4213, pRB4214, and pRB4215) were constructed by inserting in frame into the U_L26 open reading frame at the site between the 3′ end of the CMV sequence and the stop codon either a sequence encoding 256 amino acids comprising five homologs of immunoglobulin G (IgG) binding domains of staphylococcal protein A or a sequence encoding 129 amino acids and comprising two such domains. These sequences were the *Bcl*I-*Hin*cII and *Hin*dIII-*Hin*cII fragments, respectively, of the protein A gene fusion vector pRIT5 (Pharmacia, Piscataway, N.J.). The vector for plasmids T, U, V, and Y was derived from pGEM3Zf(+) (Promega, Madison, Wis.); these plasmids could be used as templates for in vitro transcription by T7 or SP6 RNA polymerase. The vectors for all other plasmids were derived from pUC18. All insertion sites of sequences A and C into the plasmids were sequenced to verify that the CMV epitope and the amino acid sequence encoded by sequence C were inserted in frame with the U_L26 open reading frame.

In vitro transcription and translation. Plasmid DNA templates (5 μg) were prepared and transcribed in the presence of capped analog GppG (New England BioLabs, Beverly, Mass.) with SP6 or T7 RNA polymerase as recommended by Promega; 1 μg of either synthetic RNAs or brome mosaic virus RNA (supplied by Promega) was translated in a 50-μl reaction mixture containing nuclease-treated rabbit reticulocyte lysate and [^{35}S]methionine (New England Nuclear, Boston, Mass.) by using a kit from Promega. After incubation of the translation reaction mixture at 37°C for different time intervals as stated in Results, the translation reaction was terminated by the addition of either disruption buffer (0.05 M Tris [pH 7.0], 8.5% [vol/vol] sucrose, 5% [vol/vol] 2-β-mercaptoethanol, 2% [vol/vol] sodium dodecyl sulfate [SDS]) or cycloheximide (Sigma, St. Louis, Mo.) to a final concentration of 100 μg/ml. In some experiments, the translation mixture was diluted 10-fold in phosphate-buffered saline containing cycloheximide at a final concentration of 100 μg/ml. In all cases, the contents of the reaction mixtures were solubilized in disruption buffer and boiled for 1 min before electrophoretic separation in denaturing gels.

Transfections and superinfection of cells transfected with plasmid DNAs. Transfections were done as described by Kristie and Roizman (11). In most experiments, the transfected cells were exposed 18 to 20 h posttransfection to 10 PFU of HSV-1(F) or HSV-2(G) per cell as stated in Results. After 2 h of exposure of cells to virus at 10°C, the inoculum was replaced with Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and the cells were incubated at 34, 37, or 39°C for 20 h as stated in Results. In the experiments which did not involve viral infection, the cells were harvested 40 to 42 h posttransfection. The harvested cells were washed once with phosphate-buffered saline, pelleted by centrifugation at 4,000 rpm for 5 min in a Sorvall SS34 rotor spun in a Dupont centrifuge, suspended in the disruption buffer, sonicated for 20 s in ice, and boiled for 1 min before electrophoretic separation in denaturing gels.

Electrophoretic separation and staining of ICPs with monoclonal antibody. The denatured, solubilized polypeptides from cell lysates or in vitro translation were separated on 9.5 or 12% (vol/vol) SDS-polyacrylamide gels cross-linked with *N,N'*-diallyltartardiamide (4, 8, 9). The separated polypeptides from BHK cells were transferred electrically to nitrocellulose membranes and reacted in an enzyme-linked immunoassay only with anti-mouse IgG conjugated with horseradish peroxidase (Amersham, Arlington Heights, Ill.) or with this anti-mouse IgG in addition to the monoclonal antibody H725 against HSV-1 ICP35 or CH28-2 against the CMV epitope, as previously described (4, 12). The gels containing the separated polypeptides translated from the reticulocyte lysate were dried and exposed to Kodak X-Omat film.

RESULTS

Experimental design. We used two different monoclonal antibodies to trace the synthesis and processing of ICP35. The first was a monoclonal antibody specific for ICP35. The epitope for this antibody mapped at or near the 5′ terminus of the coding domain of the U_L26.5 open reading frame. The second tracer was a monoclonal antibody to a CMV epitope described previously (12). An oligonucleotide sequence encoding 21 amino acids described in Materials and Methods was inserted into the coding domains of the U_L26 and U_L26.5 at sites shown in Fig. 1. Plasmids A through P were described previously (12). All others are described in Materials and Methods. These plasmids were used to determine the site and requirement for the processing of ICP35 by cleavage.

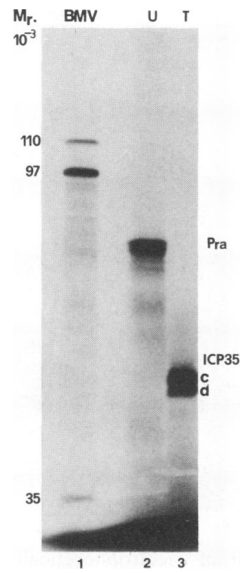

FIG. 2. Autoradiographic image of [^{35}S]methionine-labeled polypeptides translated in a nuclease-treated rabbit reticulocyte lysate and electrophoretically separated in a 9.5% denaturing polyacrylamide gel. Lanes: 1, translation products of brome mosaic virus templates provided with the kit (Promega) and transcribed according to the manufacturer's suggestions; 2, translation product of the U_L26 open reading frame in plasmid U; 3, translation product of the U_L26.5 open reading frame in plasmid T. Molecular weights are indicated in thousands on the left.

In vitro translation of U_L26 and U_L26.5 open reading frames. Braun et al. (4) reported that ICP35 proteins are processed posttranslationally into at least six species (ICP35a to -f) differing in electrophoretic mobility. Recently we demonstrated that ICP35 is encoded by U_L26.5 (12). To identify the unprocessed species of ICP35, both the U_L26.5 and U_L26 open reading frames were cloned into pGEM3Z-f(+) to derive plasmids T and U, respectively (Fig. 1). RNAs corresponding to the mRNAs of U_L26 and U_L26.5 were transcribed by SP6 RNA polymerase and translated in nuclease-treated rabbit reticulocyte lysates. The results (Fig. 2) indicated that U_L26 and U_L26.5 specify proteins each of which forms double bands with apparent molecular weights of 80,000 (Pra) and 45,000 (ICP35d and -c), respectively. The two species of U_L26.5 (ICP35) protein synthesized in vitro were found to comigrate with ICP35c and -d synthesized in vivo in HSV-1(F)-infected cells (data not shown).

The unprocessed forms of U_L26.5 ICP35c and -d can be processed into ICP35e and -f. Earlier experiments suggested that ICP35c and -d were the precursors of ICP35e and -f (4, 18). To test this hypothesis, BHK cells were transfected with plasmid E containing the U_L26.5 gene (Fig. 1) and superinfected with HSV-1(F) at 39°C. As note in the introduction, this virus is temperature sensitive and at 39°C does not express its own U_L26 and U_L26.5 open reading frames (12). The results (Fig. 3) show the following.

(i) As expected, the ICP35 gene resident in the viral genome was expressed at 34°C (lane 1) but not at 39°C (lane 2), as evidenced by the presence and absence, respectively, of the ICP35 bands reactive with monoclonal antibody H725 to ICP35.

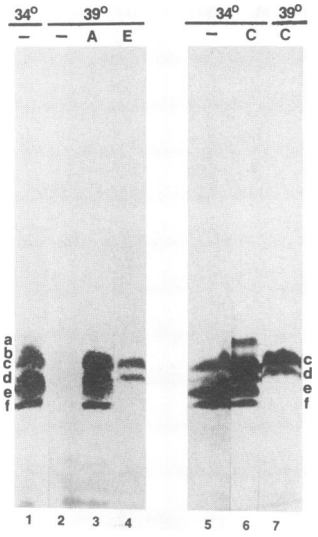

FIG. 3. Photograph of electrophoretically separated polypeptides from cells transfected with plasmid constructs, superinfected with HSV-1(F) at either 34°C (34°) or 39°C (39°), electrophoretically separated in polyacrylamide gels, electrically transferred to a nitrocellulose sheet, reacted with monoclonal antibody H725 to HSV-1 ICP35, and stained with goat anti-mouse IgG antibody coupled to peroxidase. Experimental details are described in Materials and Methods. Letters above the lanes identify the plasmid constructs with which the cells were transfected. A dash indicates that the cells were infected but not transfected. Letters at the sides refer to the different species of ICP35 as designated by Braun et al. (4).

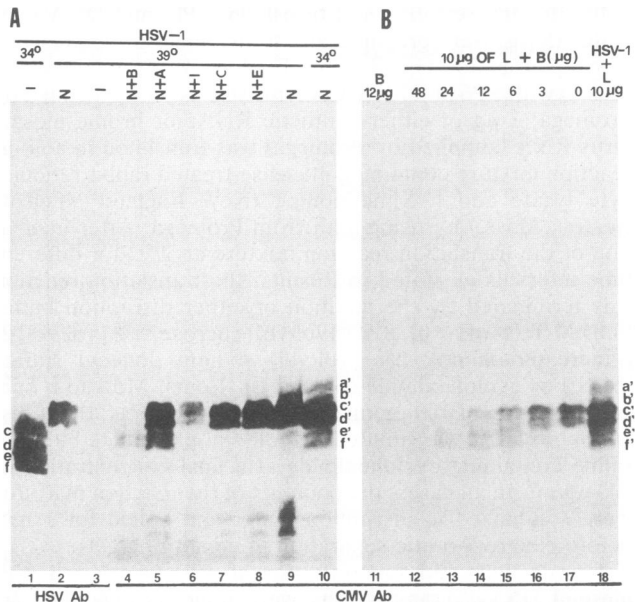

FIG. 4. Photograph of electrophoretically separated polypeptides from cells transfected with plasmids and either mock infected or superinfected with HSV-1(F) (HSV-1) at 34°C (34°; lanes 1 and 10), 39°C (39°; lanes 2 to 9), or 37°C (lanes 11 to 18), electrophoretically separated in polyacrylamide gels, electrically transferred to a nitrocellulose sheet, reacted with monoclonal antibody H725 (HSV Ab) or CH28-2 (CMV Ab), and stained with goat anti-mouse IgG antibody coupled to peroxidase. Letters above the lanes identify the plasmid constructs with which the cells were transfected. A dash indicates that the cells were infected but not transfected. Letters at the sides refer to ICP35 bands as designated by Braun et al. (4). c', d', e', and f' bands identify proteins with decreased electrophoretic mobility relative to the corresponding authentic proteins due to the insertion of the CMV epitope.

(ii) ICP35c and -d were the only two species of ICP35 expressed from the $U_L26.5$ open reading frame in plasmid E at 39°C (lane 4), whereas at least ICP35c to -f could be detected in lysates of productively infected cells maintained at 34°C (lane 1).

We conclude that (i) ICP35c and -d are the unprocessed forms of the ICP35 proteins, (ii) they can be processed into ICP35e and -f, and (iii) the processing requires a *trans*-acting factor since processing did not occur in the absence of HSV-1(F) late gene expression.

Localization of the DNA sequences in the viral genome required for the processing of ICP35c and -d into ICP35e and -f. BHK cells were transfected with a series of plasmids containing different lengths of HSV-1 DNA sequences each containing an intact ICP35 gene and superinfected with HSV-1(F) at 39°C. Figure 3 shows that BHK cells transfected with plasmid A containing the intact U_L26 gene (Fig. 1) generated ICP35e and -f in addition to ICP35c and -d (lane 3), whereas cells transfected with plasmid C in which the promoter region of the U_L26 gene was deleted and only the coding sequence of $U_L26.5$ was included (Fig. 1) generated only the unprocessed ICP35c and -d (lane 7). These results suggested that the gene product of U_L26 was required for the processing of ICP35c and -d into ICP35e and -f.

U_L26 can act in *trans* to process ICP35c and -d into ICP35e and -f. To determine whether U_L26 acts in *trans* or in *cis*, BHK cells were transfected with plasmid N as the substrate for processing and with a series of plasmids containing deletions in the U_L26 open reading frame, infected with HSV-1(F), and maintained at 39°C. The results (Fig. 4A) showed the following.

(i) ICP35c and -d did not autocatalyze their processing into ICP35e and -f, inasmuch as the lysates of cells cotransfected with plasmids N and E (Fig. 1) did not contain ICP35e and -f reactive with the CMV monoclonal antibody (lane 8).

(ii) ICP35c and -d were not processed in BHK cells cotransfected with plasmids N and C or I (lanes 7 and 6). Plasmids C and I contain deletions in the promoter region and at the polyadenylation site of the U_L26 open reading frame, respectively (Fig. 1).

(iii) ICP35c and -d were processed into ICP35e and -f in BHK cells cotransfected with plasmids N and A or B. Plasmids A and B contain the intact U_L26 promoter and open reading frame and the U_L26 coding sequence driven by the α4 promoter, respectively (lanes 5 and 4). The α-transducing factor in HSV-1(F) induces the α4 promoter to a high level (2, 17) at 39°C. The high level of expression of U_L26 may explain the presence of the processed forms of ICP35 (forms e and f) in lysates of cells cotransfected with plasmids N and B (lane 4).

These results indicate that U_L26 encodes a protein involved in the processing of ICP35c and -d into ICP35e and -f.

U_L26 is competent and the only viral protein required for the processing of ICP35c and -d into ICP35e and -f. To determine whether U_L26 is the only viral protein required for this processing and to exclude the possibility that viral genes expressed by the HSV-1(F) genome at 39°C contribute to the catalysis of ICP35, BHK cells were cotransfected with a constant amount of plasmid L and different amounts of plasmid B as the genes encoding the substrate and the enzyme for the processing, respectively. In plasmid L (Fig.

1), the U_L26.5 open reading frame was regulated by the α4 promoter and the CMV epitope was inserted at the *Mst*II restriction endonuclease site, whereas plasmid B contained the intact U_L26 open reading frame driven by the same promoter. Since the α4 promoter is a strong eukaryotic promoter constitutively expressed in transfected cells (11, 17), expression of the U_L26.5 and U_L26 proteins in cells transfected with plasmids L and B did not require superinfection with HSV-1(F). The results (Figure 4B) were as follows.

(i) In the absence of viral infection, ICP35c and -d were the only two species expressed in cells transfected with plasmid L (lane 17). The epitopically marked ICP35 expressed by plasmid L was fully processed in cells superinfected with HSV-1(F) at the permissive temperature (lane 18). As expected, plasmid B did not produce products reactive with the anti-CMV antibody (lane 11).

(ii) In the presence of plasmid B containing U_L26, the epitopically marked ICP35c and -d expressed by plasmid L were processed into ICP35e and -f. At low concentrations of plasmid B, the extent of accumulation of ICP35e and -f was directly proportional to the amount of U_L26 plasmid DNA cotransfected with plasmid L into BHK cells (lanes 12 to 16). The decrease in the amounts of ICP35e and -f observed in the presence of the highest amounts of plasmid B may reflect competition between the two plasmids or reduced yield as a result of the toxicity caused by the high amounts of DNA.

We conclude from these studies that the product of U_L26 is the only viral factor both competent and sufficient to process ICP35c and -d into ICP35e and -f.

Processing of ICP35c and -d to ICP35e and -f involves carboxyl-terminal proteolytic cleavage. Inasmuch as ICP35e and -f specified by plasmid N comigrated in denaturing gels with ICP35c and -d produced in HSV-1-infected cells (Fig. 4A, lanes 1 and 5), it may be deduced that the portion of ICP35 cleaved during processing is roughly equivalent to the size of the CMV amino acid sequence inserted into plasmid N. To determine whether ICP35 processing involves carboxyl-terminal proteolytic cleavage, BHK cells were transfected with plasmids J, Q, R, S, and W and superinfected with HSV-2. Plasmid Q contained the CMV epitope (sequence A) inserted in the *Pml*I site of U_L26.5, whereas in plasmids S and W the insert was at the carboxyl-terminal amino acid (Fig. 1). Analyses of the electrophoretically separated, electrically transferred polypeptides with the anti-HSV-1 (H725) and anti-CMV (CH28-2) monoclonal antibodies revealed the following (Fig. 5).

(i) Cells transfected with plasmid J in which the CMV epitope was inserted at the *Mst*II site 122 amino acids upstream from the U_L26 stop codon made both the precursors ICP35c and -d and the processed products ICP35e and -f, which reacted with the CMV antibody (lane 8). The decrease in the electrophoretic mobility of ICP35c to -f relative to the wild-type proteins corresponds to the increase in the molecular weight due to the insertion of the CMV epitope.

(ii) Only ICP35c and -d were made in cells transfected with plasmid Q (lanes 3 and 6). In this plasmid, the CMV epitope was inserted into the *Pml*I site of U_L26, which is 21 amino acids upstream from the U_L26.5 stop codon. Identification of the ICP35c and -d forms was based on the observation that they comigrated with the corresponding forms specified by plasmid L, which expressed only ICP35c and -d in transfected cells (Fig. 4B, lane 17).

(iii) Insertion of sequence C into plasmid R at the *Pml*I restriction endonuclease site destroyed this site and created a new *Pml*I cleavage site between the carboxyl-terminal amino acid and the stop codon of U_L26 without changing the amino acid sequence of either U_L26 or U_L26.5. ICP35c to -f detected with monoclonal antibody H725 comigrated with the authentic proteins (lane 11), indicating that the insertion of sequence C had no effect on ICP35 expression and processing.

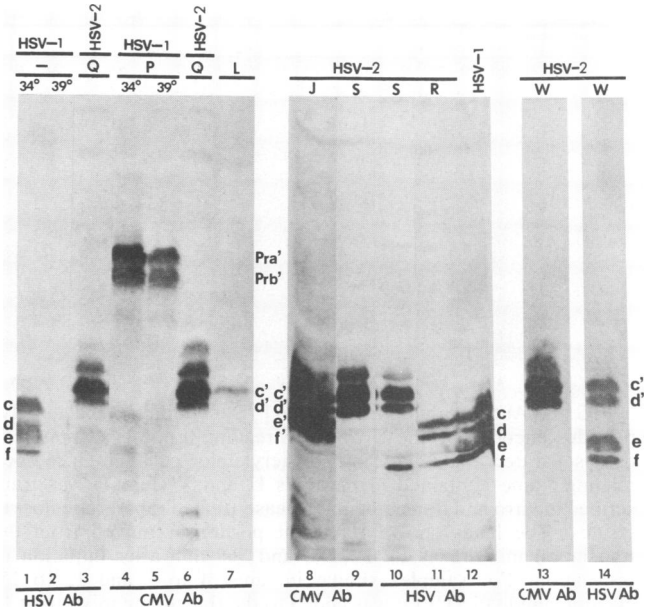

FIG. 5. Photograph of polypeptides from cells transfected with plasmids and either mock infected or superinfected with HSV-1(F) (HSV-1) or HSV-2(G) (HSV-2) at 34°C (34°; lanes 1 and 4), 39°C (39°; lanes 2 and 5), or 37°C (lanes 3 and 6 to 14), electrophoretically separated in denaturing polyacrylamide gels, electrically transferred to a nitrocellulose sheet, reacted with monoclonal antibody H725 (HSV Ab) or CH28-2 (CMV Ab), and stained with goat anti-mouse IgG antibody coupled to peroxidase. Letters above the lanes identify the plasmid constructs with which the cells were transfected. A dash indicates that the cells were infected but not transfected.

(iv) In plasmids S and W, the CMV epitope was inserted into the new *Pml*I site of plasmid R at the carboxyl terminus of U_L26.5. Cells transfected with these plasmids accumulated ICP35c to -f reactive with HSV-1 monoclonal antibody H725 (lanes 10 and 14), but only ICP35c and -d reacted with the CMV monoclonal antibody CH28-2 (lanes 9 and 13). The significant finding is that whereas ICP35c and -d of plasmid S comigrated with the corresponding forms of plasmid J, i.e., they were 21 amino acids longer than wild type, ICP35e and -f comigrated with the wild-type ICP35e and -f, indicating that the inserted amino acid sequence encoding the CMV epitope was removed (lanes 9 and 10). The products specified by plasmid W behaved in the same manner (lanes 13 and 14). ICP35e and -f specified by plasmid W were more abundant than those specified by plasmids S and R, possibly because in plasmid W the entire U_L26 open reading frame was reconstituted and more of the protein product was expressed and made available to process ICP35.

We conclude that the cleavage of the precursor ICP35 protein is approximately 20 amino acids from the carboxyl-terminal codon and that insertion of the CMV epitope 21 amino acids from the terminus interfered with the cleavage whereas insertion of the epitope at the carboxyl terminus enabled the cleavage to take effect.

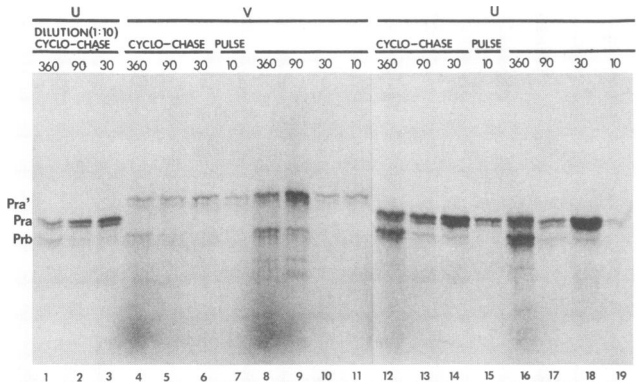

FIG. 6. Autoradiographic image of [^{35}S]methionine-labeled polypeptides encoded by the U_L26 open reading frame electrophoretically separated in a denaturing polyacrylamide gel. The U_L26 open reading frame contained in plasmids U and V (Fig. 1) was transcribed in vitro and translated in nuclease-treated rabbit reticulocyte lysates. The lanes shown represent portions removed from the translation mixture at 10, 30, 90, and 360 min after initiation of translation. For samples shown in lanes 4 to 7 and 12 to 15, cycloheximide (CYCLO) was added to the translation mixture at 10 min after initiation of translation to inhibit further translation. For the samples in lanes 1 to 3, the translation mixture at 10 min after initiation of translation was diluted 10-fold in phosphate-buffered saline containing cycloheximide (100 μg/ml).

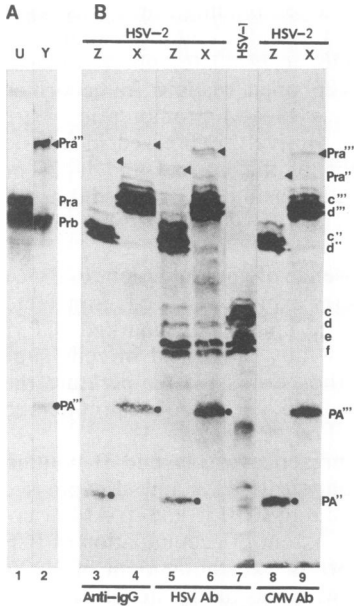

FIG. 7. Autoradiographic images (A) and photograph (B) of polypeptides either synthesized in vitro from sequences encoded in plasmid U or Y or contained in lysates of cells transfected with plasmid X or Z and superinfected with HSV-1(F) (HSV-1) or HSV-2(G) (HSV-2). The in vitro-synthesized polypeptides and those contained in cell lysates were electrophoretically separated in the same denaturing 12% polyacrylamide gel, electrically transferred to a nitrocellulose sheet, and reacted with goat anti-mouse IgG conjugated with horseradish peroxidase (Anti-IgG) only or with this anti-IgG antibody in addition to monoclonal antibody H725 (HSV Ab) or CH28-2 (CMV Ab). A dash indicates that the cells were infected but not transfected. The polypeptides shown in panel A were labeled with [^{35}S]methionine. Band designations are as follows: letters c to f without primes identify authentic ICP35 products of the $U_L26.5$ open reading frame; Pra and Prb are the translation-processed forms of the protease products of the U_L26 open reading frame; the double and triple primes indicate that the protein also contains the CMV epitope and the sequence encoding two IgG and five IgG binding domains, respectively, of staphylococcal protein A; PA″ and PA‴ are the carboxyl-terminal products of the cleavage of the ICP35c and -d and Pra proteins containing inserts of the CMV epitope and IgG binding domains.

Autoprocessing of U_L26 involves carboxyl-terminal proteolytic cleavage. In the preceding sections, we have demonstrated that U_L26 is the only viral factor responsible for the carboxyl-terminal proteolytic processing of ICP35. Recently, we also demonstrated that U_L26 and ICP35 share the same carboxyl-terminal amino acid sequence (12). The possibility that U_L26 cleaves itself emerged from the observation that BHK cells transfected with plasmid P (Fig. 1) and superinfected with HSV-1(F) at either 34 or 39°C expressed a doublet band of U_L26 (Fig. 5, lanes 4 and 5) which reacted with monoclonal antibody CH28-2. This observation suggested the possibility that U_L26 catalyzes its own cleavage since HSV-1(F) expresses primarily α genes at 39°C.

Additional evidence that U_L26 can catalyze its own cleavage emerged from in vitro studies. RNAs transcribed from plasmid U or V (Fig. 1) in vitro by SP6 or T7 RNA polymerase were translated in nuclease-treated rabbit reticulocyte lysate in the presence of [^{35}S]methionine. Analyses of the electrophoretically separated products of the translation reaction were as follows (Fig. 6).

(i) Incubation of the translation products of plasmid U in the presence of cycloheximide resulted in a gradual accumulation of the cleavage product (Prb) of the U_L26 protein. The amount of accumulated cleavage product was proportional to the duration of the incubation (lanes 12 to 15). (ii) Identical results were obtained with the translation products of plasmid V (lanes 4 to 7). The significance of this experiment stems from the presence of the CMV epitope at the carboxyl terminus of U_L26. As expected, the translation product Pra of U_L26 made from plasmid V migrated more slowly than the authentic protein derived from plasmid U. However, the processed form Prb of U_L26 synthesized from plasmid V comigrated with that of the authentic protein from plasmid U, indicating that U_L26 autoprocessing involves carboxyl-terminal proteolytic cleavage.

The cleavages of ICP35c, ICP35d, and Pra are sequence specific and at the same site. The results of the experiments presented in the preceding section predicted that the cleavage and processing of the products of U_L26 and $U_L26.5$ occurred at a site approximately 20 amino acids from the carboxyl terminus of the proteins. To demonstrate that the processing of these proteins occurs at the predicted site, it was necessary to demonstrate both products of the cleavage reaction on the same gel. To visualize both products, we inserted into the coding sequence at the predicted carboxyl terminus both the epitope for the CMV monoclonal antibody and the sequences encoding the IgG binding domains of staphylococcal protein A. Plasmids Z and X were constructed by inserting in frame the sequences coding for 129 and 256 amino acids comprising two and five IgG binding domains of protein A, respectively, between the 3′ terminus of the CMV epitope and the stop codon of U_L26 (Fig. 1). Two experiments were done. In the first, the HSV-1 open reading frames in plasmids U and Y were transcribed and translated for 6 h. The proteins translated in vitro were then electrophoretically separated in a denaturing gel (Fig. 7A). In the second experiment, BHK cells were transfected with

plasmid Z or X and then superinfected with HSV-2(G). The cell lysates were electrophoretically separated in the same gel as that used for the separation of the in vitro-translated protein, electrically transferred to a nitrocellulose sheet, and reacted with antibody to CMV or HSV or with anti-IgG antibody that would bind to IgG binding domains of protein A (Fig. 7B). The results were as follows.

(i) Autocatalytic processing of the in vitro-transcribed/translated HSV-1 sequences in plasmid U yielded, as expected, the protein bands designated Pra and Prb. Similar autocatalytic processing of the products of the Y plasmid yielded three bands. The first band migrated slower than the authentic precursor Pra band, as would be expected from the presence of the additional 256 amino acids of protein A and the 21 amino acids constituting the CMV epitope. The second band comigrated with the Prb band and is therefore the product of the autocatalytic cleavage of the translation product. The third band comigrated with the bands described below, which reacted with the CMV antibody as well as with the anti-IgG antibody. Cycloheximide chase experiments after a short pulse indicated that the first band is the precursor of the other two bands (data not shown).

(ii) The expected translation products of plasmid X were ICP35c and -d and Pra. It could be expected that the translation products would react with the CMV, HSV, and anti-IgG antibodies. The predicted translation products of plasmid Z should be similar except that because of the smaller inserts of the protein A sequences, these proteins should migrate correspondingly faster than those of plasmid X. This was in fact the case (Fig. 7; compare lanes 3, 5, and 8 with lanes 4, 6, and 9). It could also be predicted that if the cleavage of the ICP35c and -d occurs as expected 20 amino acids from the carboxyl terminus of the authentic protein, then the amino-terminal products of the cleavage reaction should comigrate with the authentic ICP35 and react only with the HSV-1 monoclonal antibody. This was in fact the case: ICP35e and -f produced by plasmids Z and X (lanes 5 and 6) comigrated with the authentic ICP35e and -f (lane 7) and were detectable solely by the HSV-1-specific monoclonal antibody. Conversely, it could be expected that the carboxyl-terminal products of the cleavage reaction should migrate in accordance with their size and should react with both anti-IgG and CMV antibodies. As shown in Fig. 7B, the bands reactive with the anti-IgG antibody from lysates of cell transfected with plasmid X migrated slower than the corresponding Z bands. However, since all of the carboxyl-terminal cleavage products contained the IgG binding domains of protein A, all of the protein products would be expected to react with IgG irrespective of specificity of the immunoglobulin (e.g., lanes 3 and 4).

Inasmuch as we detected both products, the results indicate that ICP35 (forms c and d) and Pra, the products of U_L26.5 and U_L26, respectively, are both posttranslationally processed by cleavage. Since the two proteins share amino acid sequences for the entire length of ICP35c and -d and since the products of the cleavage of the two proteins comigrate, the two proteins are cleaved at identical sites. Finally, the translational products of both open reading frames in vitro resolve into double bands. The double bands are particularly noticeable in the case of ICP35 (forms c and d). In all of the experiments done to date, including those shown in Fig. 7, the carboxyl-terminal product of the cleavage formed a single band. This observation is consistent with the hypothesis that the differences in the proteins which form the doublets are at the amino rather than carboxyl termini of the proteins.

DISCUSSION

The key finding presented in this report is that the product of the U_L26 open reading frame is both necessary and the sole viral protein that suffices to effect its own cleavage and that of the product of the U_L26.5 open reading frame. We have previously identified the product of U_L26.5 as ICP35. In this report, we have designated the product of the U_L26 as Pra. In cell-free systems, Pra cleaved itself to Prb, suggesting that the translation product, Pra, can function as a protease. Prb, the product of autocatalytic cleavage of Pra, is approximately 20 amino acids smaller. Experiments are in progress to determine whether it too exhibits proteolytic activity.

The distinguishing features of this protease are that (i) it catalyzes its own cleavage, (ii) the more abundant substrate on which it acts is encoded by a sequence entirely contained within the gene encoding the protease, and (iii) the protease and the substrate share amino acid sequences.

The substrate of this protease, ICP35, has been previously identified as a protein which changes structure in the transition from empty to full capsids (8, 9). It has been suggested that ICP35 functions as a scaffolding protein in the assembly of the capsid (15). Homologs of ICP35 have been detected in other herpesviruses; indeed, it has been reported recently that the CMV equivalent of the ICP35 protein is cleaved at the carboxyl terminus (7, 19, 21), but the CMV protease has not been identified as yet.

The protease reported in this study is the first identified for HSV-1. We have already noted that the CMV protein corresponding to ICP35 is cleaved at its carboxyl terminus, and therefore it is likely that this virus also encodes its own protease. The existence of a homologous open reading frame in the varicella-zoster virus genome (5, 14) suggests that the ICP35 equivalent and the corresponding protease are conserved among the various herpesviruses. Gibson et al. (7) noted the presence of the conserved amino acid sequence -Val-Asn-Ala-Ser- near the carboxyl terminus of the proteins predicted by the homologous open reading frames in the HSV-1 and human and simian CMV genomes. The serine in this sequence is 25 amino acids from the carboxyl terminus, approximately at or near the site of the cleavage predicted from the results shown in Fig. 5. The findings that the amino acid sequence of ICP35 is entirely contained in the carboxyl terminus of Pr and that ICP35 does not show demonstrable proteolytic activity lead us to predict that the proteolytic activity exhibited by Pr is expressed by the amino-terminal domain of the protein. Preliminary studies (13) indicate that deletion of sequences encoding approximately 150 amino acids at the carboxyl terminus does not affect the proteolytic activity of the mutated product of U_L26. It is of interest that in the varicella-zoster virus genome (5), the open reading frame corresponding to U_L26 of HSV-1 exhibit greater homology in amino acid sequence at the amino terminus than at the carboxyl terminus. The observation that the temperature-sensitive mutation of Preston et al. (18) maps at the 5′ terminus of the U_L26 open reading frame and that at the nonpermissive temperature ICP35c and -d are not processed further is congruent with but does not prove our prediction.

The only function of the protease identified to date is that it cleaves itself and ICP35. Because of its overlap with the more abundant ICP35, it is conceivable that it assembles with and is dedicated to the processing of this protein. Given the size and the number of open reading frames encoded in the HSV-1 genome, and should the protease identified in this study be indeed dedicated to the cleavage of itself and of

ICP35, it would not be too surprising if additional viral proteases were to be detected. Further studies will determine whether this is the case. Because processing of the ICP35 is an essential step in the assembly of the capsid, the protease appears to be a suitable target of antiviral drug research.

In this study, we made use of two monoclonal antibodies, one to a stationary epitope encoded by both U_L26 and $U_L26.5$ open reading frames and one reactive with a movable epitope. The latter was an indispensable tool in identification of the products of the two open reading frames, in determination of the function of the proteins, and in mapping of the cleavage site. Without the movable epitope, we would have had to rely solely on radioactive tracers or make antibodies to oligopeptides corresponding to various domains of the genes. The movable epitope offers instant antibody to the product of any open reading frame and, when used in the context described in this and preceding studies, can enormously facilitate identification of the function of the product of the gene into which it has been inserted.

ACKNOWLEDGMENTS

We thank Lenore Pereira for the invaluable gift of the monoclonal antibodies and Richard Roller for advice.

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 AI1588-11).

REFERENCES

1. **Arsenakis, M., 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. **Batterson, W., and B. Roizman.** 1983. Characterization of the herpes simplex virion-associated factor responsible for the induction of α genes. J. Virol. **46**:371–377.
3. **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.
4. **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.
5. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella-zoster virus. J. Gen. Virol. **67**:1759–1816.
6. **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.
7. **Gibson, W., A. I. Marcy, J. C. Comolli, and J. Lee.** 1990. Identification of precursor to cytomegalovirus capsid assembly protein and evidence that processing results in loss of its carboxyl-terminal end. J. Virol. **64**:1241–1249.
8. **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.
9. **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.
10. **Knipe, D. M., W. T. Ruyechan, B. Roizman, and I. W. Halliburton.** 1978. Molecular genetics of herpes simplex virus: demonstration of regions of obligatory and nonobligatory identity within diploid regions of the genome by sequence replacement and insertion. Proc. Nat. Acad. Sci. USA **75**:3896–3900.
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.** 1991. The promoter, transcriptional unit, and coding sequence of herpes simplex family 35 proteins are contained within and in frame with the U_L26 open reading frame. J. Virol. **65**:206–212.
13. **Liu, F., and B. Roizman.** Unpublished data.
14. **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.
15. **Newcomb, W. W., and J. C. Brown.** 1991. Structure of the herpes simplex virus capsid: effects of extraction with guanidine hydrochloride and partial reconstitution of extracted capsids. J. Virol. **65**:613–620.
16. **Pereira, L., and A. Sears.** Unpublished data.
17. **Post, L. E., S. Mackem, and B. Roizman.** 1981. The regulation of α genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with α gene promoters. Cell **24**:555–565.
18. **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.
19. **Robson, L., and W. Gibson.** 1989. Primate cytomegalovirus assembly protein: genome localization and nucleotide sequence. J. Virol. **63**:669–676.
20. **Roizman, B., and P. G. Spear.** 1968. Preparation of herpes simplex virus of high titer. J. Virol. **2**:83–84.
21. **Schenk, P., A. S. Woods, and W. Gibson.** 1991. The 45-kilodalton protein of cytomegalovirus (Colburn) B-capsids is an amino-terminal extension form of the assembly protein. J. Virol. **65**:1525–1529.