

Expression and Analysis of the Human Cytomegalovirus UL80-Encoded Protease: Identification of Autoproteolytic Sites

ELLEN Z. BAUM,* GERALDINE A. BEBERNITZ, JEFFREY D. HULMES,
VIERA P. MUZITHRAS, THOMAS R. JONES, AND YAKOV GLUZMAN

*Molecular Biology Section, Medical Research Division, Lederle Laboratories,
American Cyanamid Company, Pearl River, New York 10965*

Received 22 July 1992/Accepted 23 October 1992

The 45-kDa assembly protein of human cytomegalovirus is encoded by the C-terminal portion of the UL80 open reading frame (ORF). For herpes simplex virus, packaging of DNA is accompanied by cleavage of its assembly protein precursor at a site near its C terminus, by a protease encoded by the N-terminal region of the same ORF (F. Liu and B. Roizman, *J. Virol.* 65:5149-5156, 1991). By analogy with herpes simplex virus, we investigated whether a protease is contained within the N-terminal portion of the human cytomegalovirus UL80 ORF. The entire UL80 ORF was expressed in *Escherichia coli*, under the control of the phage T7 promoter. UL80 should encode a protein of 85 kDa. Instead, the wild-type construct produces a set of proteins with molecular masses of 50, 30, 16, 13, and 5 kDa. In contrast, when mutant UL80 is deleted of the first 14 amino acids, it produces only an 85-kDa protein. These results suggest that the UL80 polyprotein undergoes autoproteolysis. We demonstrate by deletional analysis and by N-terminal sequencing that the 30-kDa protein is the protease and that it originates from the N terminus of UL80. The UL80 polyprotein is cleaved at the following three sites: (i) at the C terminus of the assembly protein domain, (ii) between the 30- and 50-kDa proteins, and (iii) within the 30-kDa protease itself, which yields the 16- and 13-kDa proteins and may be a mechanism to inactivate the protease.

Human cytomegalovirus (HCMV) is a clinically serious pathogen in immunocompromised adults, including patients with AIDS and recipients of organ or bone marrow transplants. This betaherpesvirus has a 230-kb double-stranded DNA genome which was recently sequenced in its entirety and contains over 200 significant open reading frames (ORFs) (4). Although HCMV has been studied widely for several decades, the functions of most of the protein products of these genes are unknown.

In 1983, Preston et al. (24) reported on a temperature-sensitive mutant, *ts1201*, of herpes simplex virus type 1 (HSV-1), an alphaherpesvirus. At the nonpermissive temperature, progeny viral DNA was not packaged into virions, resulting in empty nucleocapsids. Protein analyses revealed that an ~45-kDa precursor of a viral protein, called ICP35 (or VP22a or p40) (3, 33), was not processed to its ~40-kDa mature form, suggesting a linkage between ICP35 and the maturation of nucleocapsids. Therefore, ICP35 has been referred to as the assembly protein and is proposed to be analogous to the bacteriophage T4 scaffolding protein (7, 10, 11, 14). The assembly protein is associated with capsids which lack DNA (29). Marker rescue studies of *ts1201* indicated that the defect was in the N-terminal region of the UL26 ORF (24, 25), which could encode an ~80-kDa protein (15, 20). Recently, Liu and Roizman (16) reported that the UL26 ORF encodes a protease and its assembly protein precursor substrate from its N- and C-terminal regions, respectively. They demonstrated that the UL26 protease is responsible for the cleavage of the assembly protein near its C terminus. The assembly protein and protease are translated from separate UL26-derived mRNAs which are 3'

coterminous (9, 15), and the amino acid sequence of the assembly protein is identical to that of the C-terminal region of the ~80-kDa protease.

The HCMV UL80 ORF was previously reported to have homology with the HSV-1 UL26 ORF (4). The HCMV UL80 ORF also has homology with the assembly protein nested genes (APNGs) from simian cytomegalovirus (SCMV) (26, 30). Like HSV-1 UL26, several 3'-coterminous mRNAs are transcribed from both HCMV UL80 and SCMV APNGs at late times postinfection; multiple proteins were shown to be encoded by these transcripts (30). Gibson and coworkers have demonstrated that a protein corresponding to the C-terminal region of the SCMV ORF is posttranslationally cleaved near its C terminus (8, 28). The corresponding region of HCMV UL80 encodes a protein which is proteolytically processed in a similar fashion (8, 28), and both the HCMV and the SCMV proteins may be functionally analogous to the HSV-1 assembly protein.

On the basis of the identification of the HSV-1 UL26 assembly protein protease (16), it was of interest to determine whether HCMV UL80 also encodes a protease function. Recently, by transient expression in eukaryotic cells, the SCMV APNG1 ORF was shown to encode a protease (31). Our strategy was to express the UL80 ORF and several truncated derivatives in *Escherichia coli*. We report that the N-terminal domain of UL80 encodes a 30-kDa protease. The protease cleaves the UL80 polyprotein at three sites, producing five different products. The UL80-encoded protease was partially purified and was shown to cleave the HCMV assembly protein precursor in *trans*. This cleavage was sensitive to inhibitors of cysteine proteases. A model for proteolytic processing and its possible regulatory role in virus replication are discussed.

* Corresponding author.

MATERIALS AND METHODS

Cells and virus. Growth conditions for human foreskin fibroblasts (HFF) were described previously (13). HCMV strain AD169 was obtained from the American Type Culture Collection and propagated according to standard protocols (13).

Plasmid constructions. The numerical base designation of the HCMV strain AD169 is according to the system described by Chee et al. (4). The HCMV ORF UL80 (4) was placed under the control of the T7 promoter in the bacterial expression vector pT7K (23, 27) to yield pCPRT3 as follows. pHindIII-L, which contains the HCMV strain AD169 genomic HindIII-L DNA fragment in pAT153 (22), was digested with *SspI* and *XbaI*, and the UL80 DNA fragment corresponding to bases 115078 to 117485 was made blunt with Klenow enzyme, which was followed by partial digestion with *AccI* at base 115238. A synthetic oligonucleotide duplex which begins with an *NdeI* restriction site was used to rebuild the coding sequence of UL80 from the initiating methionine codon to the *AccI* site. The UL80 fragment plus oligonucleotides was ligated into the *NdeI* and *BamHI* sites (the latter was made blunt with Klenow enzyme) of pT7K. The ligation mixture was transformed into *E. coli* DH5 α . To create pAcc33, which has the first 14 amino acids of UL80 deleted, the *AccI* site at base 115238 was made blunt with Klenow enzyme and ligated into the *EcoRI* site of a derivative of pT7K previously made blunt with Klenow enzyme, placing Tyr-15 of UL80 immediately downstream of an in-frame Met-Glu-Phe of the vector.

The 3'-translational truncations of UL80 were constructed by linearizing pCPRT3 or pAcc33 at the indicated restriction sites (Fig. 1), making blunt ends as necessary, and ligating the *XbaI* nonsense linker CTAGTCTAGACTAG (New England Biolabs).

To construct p30k, which encodes the 30-kDa protein encoded by the N-terminal region of UL80, pCPRT3 was digested with *SpeI* (at Leu-244) and *BamHI* (in the vector) and a synthetic oligonucleotide duplex was inserted which regenerates the sequence from the *SpeI* site to Ala-256, followed by a translation termination codon and a *BamHI* site. Several insertions and deletions of the C terminus of the 30-kDa protein were also generated. p30k(-8) has nucleotide G115943 deleted, resulting in a frameshift and subsequent translation termination. Thus, the C terminus of this protein is Thr-248 followed by the frameshift sequence Asp-Ala-Ser-His-Thr. p30k(-3) is deleted of nucleotide G115947, resulting in a translation termination codon immediately after Tyr-253. The C terminus of the protein encoded by p30k(+12) is Ala-256 followed by Gly-Gly-Glu-Gln-Lys-Leu-Ile-Ser-Glu-Glu-Asp-Leu; the last 10 amino acids are a *c-myc* epitope (6).

Plasmid pBX49, encoding the C-terminal portion of UL80 and containing the putative assembly protein cleavage site (16), was constructed by ligating the *BamHI-XbaI* fragment of UL80 (bases 116508 to 117485) into an *EcoRI* site (previously made blunt with Klenow enzyme) downstream of an in-frame methionine in a derivative of pT7K.

***E. coli* expression.** UL80 plasmids were transformed into the expression strain, *E. coli* BL21pLysS (27). This strain contains a chromosomal copy of T7 RNA polymerase under the control of an inducible *lacUV5* promoter and a chloramphenicol resistance plasmid encoding T7 lysozyme, which inhibits transcription by low levels of T7 RNA polymerase in the absence of induction (21). For expression of UL80

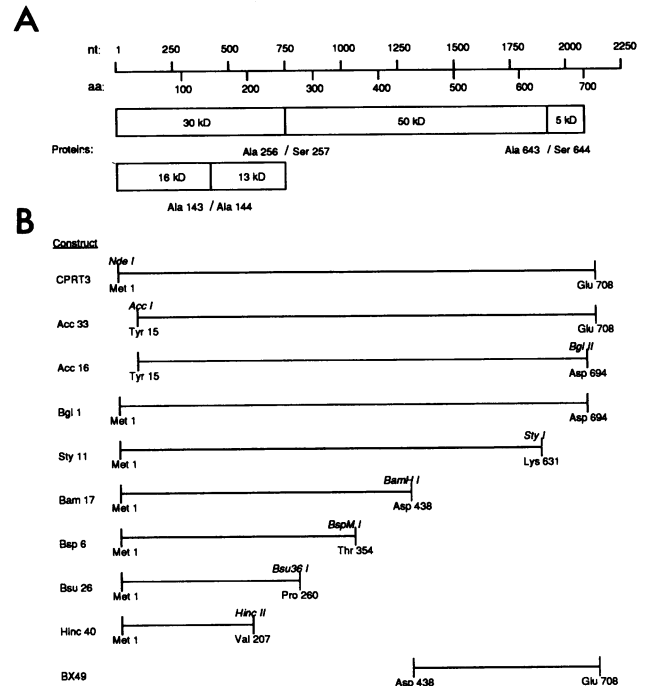


FIG. 1. (A) Map of HCMV UL80 proteins. The 16-kDa-13-kDa and 30-kDa-50-kDa cleavage sites were determined by N-terminal sequencing of the 13- and 50-kDa proteins, respectively. The 50-kDa-5-kDa site was determined by Welch et al. (31) by C-terminal sequencing of the HCMV mature assembly protein. aa, amino acid; nt, nucleotide. (B) UL80 constructs used in this study. Each region of UL80 pictured is under the control of the phage T7 promoter in the bacterial expression plasmid pT7K.

proteins, cells were induced with isopropyl- β -D-thiogalactopyranoside (IPTG) as described previously (23).

Radiolabelling of *E. coli*. *E. coli* cells harboring UL80 expression plasmids were metabolically labelled with [³⁵S]cysteine and/or [³⁵S]methionine as described previously (23). For the pulse-chase experiment (see Fig. 6), *E. coli* cells (2 ml) were grown with shaking at 37°C to an A_{595} of 0.5 in Luria broth (18) containing kanamycin (50 μ g/ml) and chloramphenicol (30 μ g/ml). IPTG (1 mM) was then added, and the incubation was continued for 30 min. Cells were pelleted and resuspended in 2 ml of M9 minimal medium (18) containing 1 mM IPTG. After continued incubation for 30 min, rifampin (20 μ g/ml) was added to suppress transcription and subsequent translation from *E. coli* (but not T7) promoters. After 30 min, the cells were pulsed with 200 μ Ci of [³⁵S]methionine (1,000 Ci/mmol), followed by a chase 25 s later with 0.2 ml of unlabelled 1% methionine. Aliquots (0.25 ml each) were withdrawn at the indicated times and precipitated with 0.5 ml of 10% trichloroacetic acid on ice for 30 min, and the precipitates were pelleted in a microcentrifuge. Samples were washed with cold acetone, dried, and resuspended in Laemmli sample buffer (18), and the suspensions were boiled prior to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography.

HCMV protease purification. Crude *E. coli* lysates were used as a source of protease for the experiment depicted in Fig. 4. *E. coli* BL21pLysS cells (1 ml) harboring the indicated plasmids were grown to an A_{595} of 0.5 in Luria broth, induced with 1 mM IPTG for 2 h, pelleted, and resuspended

in 0.1 volume of 10 mM Tris-Cl (pH 7.5)–1 mM dithiothreitol–50 mM NaCl (TDN). The cells were lysed either by sonication or by four cycles of freezing on dry ice followed by thawing in ice water. After centrifugation, the clear supernatant was used as a source of protease. To prepare substrate, a similar procedure was followed, except that the cells were labelled with [³⁵S]methionine. Some proteins (e.g., the protein from bacteria harboring pAcc16) did not partition into the supernatant but rather into inclusion bodies in the pellet fraction. In these cases, the pellet was extracted with 8 M urea–10 mM Tris-Cl (pH 7.5)–1 mM dithiothreitol–1 M NaCl. To renature the extracted proteins, this material was dialyzed against 10 mM Tris-Cl (pH 7.5)–1 mM dithiothreitol–1 M NaCl–10% glycerol and then against TDN–10% glycerol.

Protease assays. Protease and radiolabelled substrate were incubated in TDN buffer at 37°C for 2 h, prior to SDS-PAGE and autoradiography.

Protein sequence analysis. Proteins were electroblotted onto polyvinylidene difluoride membranes (Trans-Blot; Bio-Rad) (19) and visualized by Coomassie blue staining. The N-terminal sequence of the membrane-bound protein was determined by automated Edman degradation, using a Blott cartridge in an Applied Biosystems model 477A protein sequencer.

Antibodies. Preparative reverse-phase high-performance liquid chromatography on crude protease was performed with a C4 column (Vydac; 250 by 10 mm) and a gradient of acetonitrile in 0.1% trifluoroacetic acid (0 to 28% acetonitrile in 10 min; 28 to 56% acetonitrile in 60 min; 56 to 80% acetonitrile in 10 min at 1.5 ml/min). This procedure resulted in the separation of the 13-kDa protein from the 30- and 16-kDa proteins. After lyophilization, the 30- plus 16-kDa protein fraction was coupled to keyhole limpet hemocyanin and injected into New Zealand White rabbits for antibody production by standard techniques (Pocono Rabbit Farm and Laboratory, Canadensis, Pa.).

Immunoprecipitation of HCMV proteins. Uninfected or infected-cell proteins were radiolabelled and immunoprecipitated as described elsewhere (13).

RESULTS

UL80 polyprotein expressed in *E. coli* undergoes autoproteolysis. To determine whether the HCMV UL80 ORF encodes a protease function analogous to that of its reported homolog in HSV-1, UL26 (16), UL80 was expressed in *E. coli* BL21pLysS under the control of the T7 promoter. We have previously reported the expression of functional human immunodeficiency virus and poliovirus proteases in this system (1, 2). pCPRT3 contains the entire UL80 ORF (Fig. 1B) and was expected to encode a protein with a molecular mass of approximately 80 kDa. Instead, upon metabolic labelling of *E. coli* BL21pLysS cells harboring pCPRT3 with [³⁵S]methionine, proteins with molecular masses of approximately 50, 30, 16, and 5 kDa were observed (Fig. 2A). When [³⁵S]cysteine was used instead of [³⁵S]methionine, the 50-, 30-, 16-, and 5-kDa proteins were again detected, as well as a 13-kDa protein which lacks a methionine residue. In contrast, bacteria harboring pAcc16 (Fig. 1B), which is deleted of both the first 14 amino acids of the N terminus and the last 14 amino acids of the C terminus of UL80, produced a single protein with a molecular mass of approximately 85 kDa (Fig. 2A). These results suggested that the UL80 protein encoded by pCPRT3 undergoes autoproteolytic cleavage, whereas the deletion protein encoded by pAcc16 is unable to

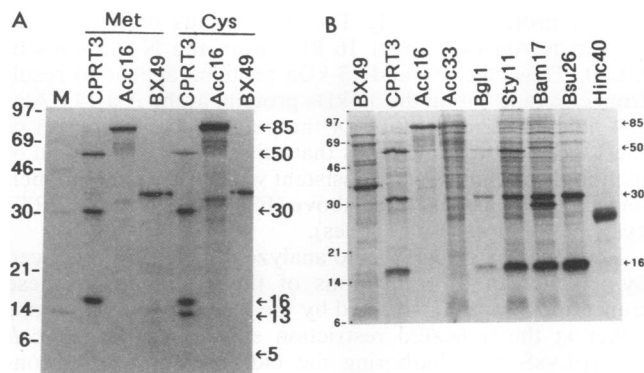


FIG. 2. (A) Metabolic labelling of UL80 constructs in *E. coli* BL21pLysS. Cells harboring the indicated plasmids were labelled with either [³⁵S]methionine (Met) or [³⁵S]cysteine (Cys) and then subjected to SDS-PAGE and autoradiography. The 5-kDa protein is not obvious at this exposure level but is clearly visible with longer exposure of the gel (data not shown) and comigrates with the 5-kDa protein generated by *in vitro* cleavage of the assembly protein-like precursor (protein encoded by pBX49; see Fig. 4, lane 3). (B) Labelling of UL80 truncation constructs. Cells harboring the indicated plasmids were labelled with [³⁵S]methionine as described for panel A. The positions of molecular mass markers (in kilodaltons) are shown to the left of each panel.

autodigest. A series of additional experiments described below support this view.

Mapping of the UL80 proteolytic products. Two approaches were used to localize the products of the apparent autoproteolytic cleavage of the UL80 protein. The protein map deduced from these experiments appears in Fig. 1A and is consistent with the independently obtained map for the simian homolog of this gene, APNG1 (31). In the first strategy, the 50-, 30-, 16-, and 13-kDa proteins were subjected to N-terminal sequence analysis (Table 1). The resulting amino acid sequences were then compared with the translated UL80 gene sequence (4) and aligned (Fig. 1A). The N terminus of the 30-kDa product is Met-1, the N terminus of UL80. The N terminus of the 50-kDa product is Ser-257, which is about 30 kDa from the N terminus of UL80. This suggested that the 50-kDa protein resulted from cleavage of the UL80 polyprotein precursor at the Ala-256–Ser-257 junction. Since C-terminal sequencing was not performed, the C terminus of the 30-kDa protein is inferred to be Ala-256, although this was not demonstrated directly.

Like that of the 30-kDa protein, the N terminus of the

TABLE 1. N-terminal sequences of HCMV proteins expressed from pCPRT3

Protein	Amt (pmol)	Sequence ^a	Amino acid position
50 kDa	10	X ₁ -V-X ₂ -P-E-A-A-X ₃ -D-I ^b -	Ser-257
30 kDa	200	M ^c -T-M-D-E-Q-Q-S-Q-A-	Met-1
16 kDa	400	M ^c -T-M-D-E-Q-Q-S-Q-A-	Met-1
13 kDa	350	A-T-S-L-S-G-S-E-T-T-	Ala-144

^a The first 10 amino acids are indicated.

^b From alignment of this peptide with the UL80 ORF, X₁ and X₂ are inferred to be Ser-257 and Ser-259; the amounts were below detectable levels. Cysteine residues are not detected by this methodology; X₃ is inferred to be Cys-264.

^c Approximately 20 to 25% of the protein contains no N-terminal methionine.

16-kDa protein is Met-1. The N terminus of the 13-kDa protein is Ala-143, about 16 kDa from the N terminus of UL80. Thus, the 16- and 13-kDa proteins appear to result from cleavage within the 30-kDa protein, at the Ala-143–Ala-144 junction. Examination of the putative coding region of the 13-kDa protein indicates that this sequence is devoid of methionine residues (4), consistent with the results of labeling experiments described above (Fig. 2A; compare CPRT3 cysteine and methionine lanes).

The second strategy was to analyze the proteins produced by a series of 3' truncations of the UL80 gene. These truncations were constructed by the insertion of a nonsense linker at the indicated restriction sites (Fig. 1B). *E. coli* BL21pLysS cells harboring the indicated truncation constructs were metabolically labelled with [³⁵S]methionine, and induced proteins were displayed by SDS-PAGE and then by autoradiography (Fig. 2B). pCPRT3 produced the 50-, 30-, and 16-kDa proteins. The 13-kDa band was not detected because it lacks methionine, but the 5-kDa band was detected upon prolonged exposure of the gel (data not shown). When the experiment depicted in Fig. 2B was repeated with [³⁵S]cysteine instead of [³⁵S]methionine, essentially identical results were obtained for all constructs, except that whenever the 16-kDa protein was produced, the 13-kDa protein was also detected (data not shown).

Construct pAcc16 again produced the ~85-kDa polyprotein, as did pAcc33, from which 14 amino acids at the N terminus are deleted but which has an intact C terminus (Fig. 1B). These data suggest that sequences at the N terminus are required for the putative autodigestion of the UL80 polyprotein. Furthermore, the lack of significant levels of discrete 50-, 30-, and 16-kDa cleavage products from the UL80 polyprotein encoded by both pAcc16 and pAcc33 indicates that this proteolysis is not carried out by an *E. coli* protease, since all of the proposed cleavage sites are present in these two mutant UL80 proteins. Some nonspecific degradation of the pAcc16 and pAcc33 polyproteins does occur (Fig. 2B) and is probably due to cell proteases.

The UL80 ORFs encoded by constructs pBgl1 and pSty11 should end at Asp-694 and Lys-631, respectively, compared with Glu-708 of pCPRT3 (Fig. 1B). Both the pBgl1 and the pSty11 constructs produced the 50-, 30-, and 16-kDa proteins (Fig. 2B). The 5-kDa protein was not detected for pBgl1 and pSty11, even upon prolonged exposure of a higher-percentage gel when the 5-kDa protein produced by pCPRT3 was visible (data not shown). The Ala-643–Ser-644 junction is predicted to be the site of assembly protein cleavage (8, 28, 31) and is also likely to be the junction between the 50- and 5-kDa proteins which we observe (Fig. 2A). Since the pBgl1 construct terminates at Asp-694, which is downstream of this cleavage site, the 50-kDa protein would be unaffected. The pSty11 construct terminates at Lys-631, 12 amino acids upstream of the putative 50-kDa–5-kDa junction. Therefore, the 5-kDa protein should not be made and the 50-kDa protein should be reduced in size by ~1 kDa, which is not a detectable difference in this gel system.

The pBam17 construct terminates at Asp-438 (which is within the 50-kDa protein) and produced the 30- and 16-kDa proteins but not the 50-kDa protein, as expected (Fig. 2B). In addition, an ~28-kDa protein, which may be the truncated version of the 50-kDa protein, is observed. The pBsu26 construct terminates at Pro-260, just downstream of the Ala-256–Ser-257 junction which apparently separates the 30- and 50-kDa proteins. pBsu26 also produced the 30- and 16-kDa proteins (as well as the 13-kDa protein if [³⁵S]cysteine is used; data not shown), suggesting that the protease

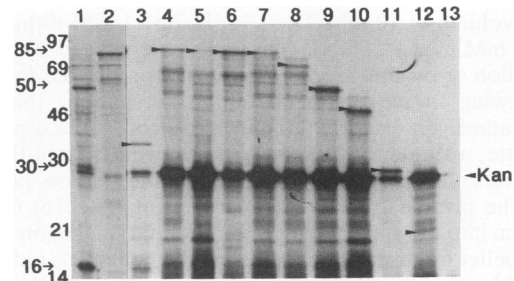


FIG. 3. Coupled transcription and translation of UL80 plasmids. The indicated plasmids (1 to 2 μ g) were incubated with *E. coli* S30 extract containing [³⁵S]methionine as described elsewhere (5) and then subjected to SDS-PAGE and autoradiography. Lanes: 3, pBX49; 4, pCPRT3; 5, pAcc16; 6, pAcc33; 7, pBgl1; 8, pSty11; 9, pBam17; 10, pBsp6; 11, pBsu26; 12, pHinc40; 13, no DNA. For comparison, [³⁵S]methionine-labelled proteins produced in BL21pLysS by constructs pCPRT3 (lane 1) and pAcc16 (lane 2) are shown. The plasmid-encoded kanamycin resistance protein (Kan) is indicated. Molecular mass markers (in kilodaltons) are shown on the left.

domain is found at the N terminus of UL80, analogously to the UL26 protease of HSV-1 (16). Consistent with this hypothesis is the observation that the pHinc40 construct, which contains a synthetic termination codon after Val-207, produced a 25-kDa protein, which is the size expected from this ORF. The 25-kDa protein apparently lacks autoproteolytic activity, since the 16-kDa protein was not produced, even though the Ala-143–Ala-144 cleavage site is present. The other putative cleavage product (~7 kDa, corresponding to Ala-144–Val-207 and containing two cysteine residues) was also not detected, consistent with lack of proteolytic activity. These experiments strongly suggest that the 30-kDa protein is the protease responsible for cleavage of the UL80 polyprotein; however, the possibility that the 13-kDa protein is actually the protease could not be excluded. From data to be discussed elsewhere, the 13- and 16-kDa proteins were shown to lack protease activity; pCPRT3, which has been mutated at the 16-kDa–13-kDa junction, still generates the 30-, 50-, and 5-kDa proteins, despite the absence of the 16- and 13-kDa products, suggesting that the 30-kDa species is the active protease (1).

To verify that all of these truncation constructs do in fact encode the expected ORFs, these plasmids were subjected to coupled *in vitro* transcription and translation in an *E. coli* S30 extract (Fig. 3). We had previously observed that a human immunodeficiency virus protease precursor which autodigests to completion when expressed in intact *E. coli* cells (1a) produces detectable amounts of the protease precursor when expressed in the S30 system (1). Similarly, pCPRT3 and the UL80 truncation constructs displayed less autoproteolysis in the S30 system than in intact *E. coli* cells (Fig. 3). For each construct, protein of the size expected for its ORF was detected. Welch et al. (31) also noted that the protein from the corresponding SCMV ORF, APNG1, failed to autodigest upon translation of its mRNA in the rabbit reticulocyte lysate.

Proteolysis *in trans* of the assembly protein cleavage site. To determine the abilities of various derivatives of UL80 protein to carry out proteolysis *in trans*, [³⁵S]methionine-labelled substrate was prepared from *E. coli* cells harboring pBX49 (Fig. 1B). pBX49 encodes a 35-kDa protein (Fig. 4, lane 1) corresponding to amino acids Asp-438–Glu-708 of UL80, containing the predicted assembly protein cleavage

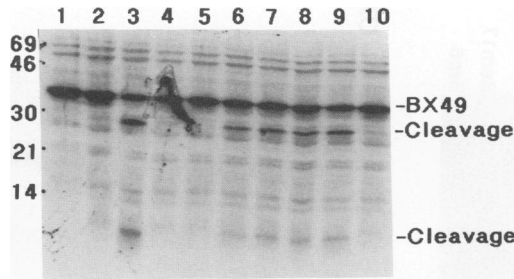


FIG. 4. Transdigestion of the substrate derived from pBX49 by UL80 proteins. Lysates of *E. coli* BL21pLysS cells harboring the indicated plasmids were incubated with [³⁵S]methionine-labelled pBX49-derived protein as described in Materials and Methods and then subjected to SDS-PAGE and autoradiography. Molecular mass markers (in kilodaltons) are shown to the left. The positions of intact pBX49-derived protein and of the 30- and 5-kDa cleavage products are shown on the right. Lane 1, unincubated pBX49-derived substrate. Labelled pBX49-derived substrate was incubated with lysates of bacteria harboring pBX49 (lane 2), pCPRT3 (lane 3), pAcc16 (lane 4), pAcc33 (lane 5), pBgl1 (lane 6), pSty11 (lane 7), pBam17 (lane 8), pBsu26 (lane 9), or pHinc40 (lane 10).

site at Ala-643-Ser-644 (which also appears to be the cleavage site between the 50- and 5-kDa proteins encoded by HCMV UL80; Fig. 1A). Note that the bona fide assembly protein corresponds to Met-336-Glu-708 and is presumably translated from the 1.1-kb UL80.5 mRNA, compared with the 2.1-kb UL80a mRNA encoding the entire UL80 ORF (30). Thus, the pBX49-derived substrate is an ~10-kDa N-terminal deletion of the HCMV assembly protein precursor. Lysates were prepared from induced, unlabelled *E. coli* cells harboring pCPRT3 or the truncation constructs (Fig. 1B) and incubated with the radiolabelled pBX49-derived substrate. Proteolytic cleavage of this 35-kDa substrate encoded by pBX49 was monitored by SDS-PAGE followed by autoradiography (Fig. 4). The lysate from *E. coli* cells

harboring pCPRT3 cleaved the 35-kDa pBX49-derived substrate into the expected 30- and 5-kDa digestion products (Fig. 4, lane 3). These products were not observed upon incubation of the pBX49-derived substrate alone (Fig. 4, lane 2) or with pAcc16- or pAcc33-derived lysate (lanes 4 and 5). In contrast, lysates from bacteria harboring truncation constructs pBgl1, pSty11, pBam17, and pBsu26 (Fig. 4, lanes 6 to 9) were able to cleave the pBX49-derived substrate into the 30- and 5-kDa products. However, the lysate from bacteria harboring pHinc40 failed to cleave the pBX49-derived substrate (Fig. 4, lane 10). These data suggest that the ability to cleave the HCMV assembly protein substrate *in trans* is retained by the N-terminal 30-kDa domain of the UL80 protein.

Inhibition of protease activity. HCMV protease was prepared from inclusion bodies from BL21pLysS cells harboring pBam17. This material is a mixture of the 30-, 16-, and 13-kDa proteins (Fig. 5A, lane 5). After the protease was titrated to determine the minimal amount needed to cleave the pBX49 substrate (Fig. 5B), a battery of protease inhibitors were tested for their abilities to inhibit the protease (Fig. 5C). Aprotinin (Fig. 5, lane 3 [1 μg/ml]), E-64 (lane 4 [0.5 μg/ml]), calpain inhibitors I (17 μg/ml) and II (7 μg/ml), TLCK (*N*-tosyl-L-lysine chloromethyl ketone; 40 μg/ml), TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone; 80 μg/ml), pepstatin A (0.7 μg/ml), leupeptin (0.5 μg/ml), cystatin (100 μg/ml), and EDTA (1 mM) did not inhibit the digestion of the pBX49-derived substrate (data not shown). However, ZnCl₂ (2.5 mM) completely inhibited the proteolysis of that substrate (Fig. 5, lane 6), suggesting that HCMV protease is a cysteine protease. Phenylmethylsulfonyl fluoride (100 μg/ml), which inhibits cysteine and serine proteases, was also slightly inhibitory toward HCMV protease (Fig. 5, lane 5). In contrast to our results with HCMV UL80 protease, Liu and Roizman (17) recently demonstrated that HSV-1 UL26 protease exhibits serine protease-like features.

The 30-kDa protein is not the precursor of the 16- and 13-kDa proteins. Since the N-terminal sequence analysis

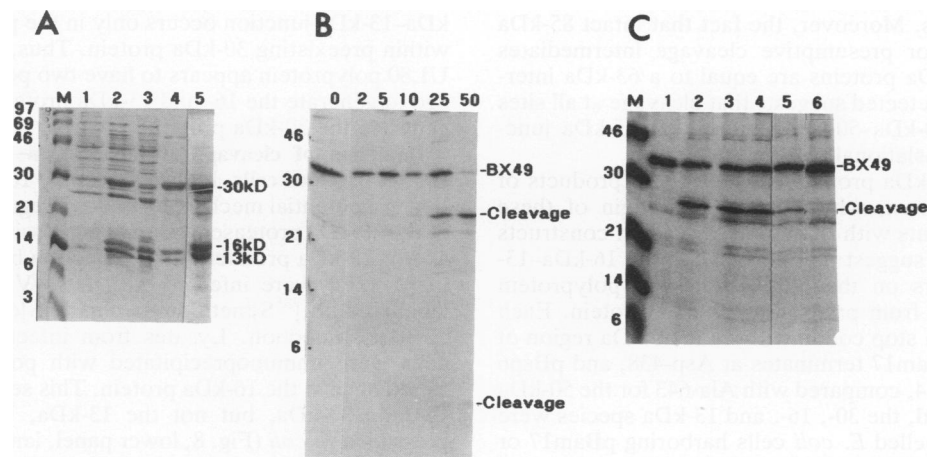


FIG. 5. (A) Purification of HCMV protease. *E. coli* BL21pLysS cells harboring pBam17 are shown prior to (lane 1) and 2 h after (lane 2) addition of IPTG. The induced cells were lysed in a French press, and the lysates were centrifuged to separate supernatant (lane 3) from pellet (lane 4). The pellet was solubilized with urea and dialyzed as described in Materials and Methods (lane 5). Lanes 1 to 4 contain 100-μl cell equivalents; lane 5 contains 500-μl cell equivalents. This is a Coomassie blue stain of a 15% polyacrylamide gel. (B) Titration of HCMV protease. The indicated quantities (in micrograms) of 30-kDa protease prepared from the pellet shown were incubated with pBX49-derived substrate and then subjected to SDS-PAGE and autoradiography. (C) Inhibitor test. A 20-μg sample of HCMV protease was incubated with pBX49-derived substrate in the presence of the indicated protease inhibitors and then subjected to SDS-PAGE and autoradiography. Lanes: 1, no protease; 2, no inhibitor; 3, aprotinin (1 μg/ml); 4, E-64 (0.5 μg/ml); 5, phenylmethylsulfonyl fluoride (100 μg/ml); 6, ZnCl₂ (2.5 mM).

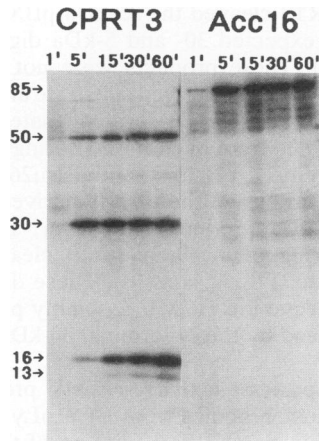


FIG. 6. Pulse-chase of *E. coli* BL21pLysS cells harboring pCPRT3 or pAcc16. Cells were pulsed with [³⁵S]methionine as described in Materials and Methods and chased with unlabelled methionine for the indicated times. The 13-kDa protein lacks methionine residues; the labelling of this protein probably results from reincorporation of label from [³⁵S]methionine into cysteine.

discussed above demonstrated that the 16- and 13-kDa proteins are organized in a linear fashion within the 30-kDa protein, it was of interest to determine whether the 30-kDa protein is the precursor of the two smaller proteins. Accordingly, a pulse-chase radiolabelling experiment was performed on *E. coli* cells harboring pCPRT3 or pAcc16 (Fig. 6). As expected, pCPRT3 produced [³⁵S]methionine-labelled proteins with molecular masses of 50, 30, and 16 kDa. A small amount of labelling of the 13-kDa protein was detected and probably results from degradation of methionine and reincorporation of the ³⁵S into cysteine. None of these proteins is present in the lysates of bacteria harboring the pAcc16 control. Between 5 and 60 min, the amounts of the 16- and 13-kDa proteins increased approximately 10-fold, yet the amount of 30-kDa protein stayed constant. These data suggest that the 30-kDa protein is not the precursor of the 16- and 13-kDa proteins. Moreover, the fact that intact 85-kDa UL80 polyprotein or presumptive cleavage intermediates (e.g., 13- plus 50-kDa proteins are equal to a 63-kDa intermediate) were not detected suggests that cleavage at all sites (16-kDa–13-kDa, 30-kDa–50-kDa, and 50-kDa–5-kDa junctions) occurs cotranslationally.

If the 16- and 13-kDa proteins are not direct products of the 30-kDa breakdown, what then is the origin of these proteins? Experiments with the UL80 truncation constructs pBam17 and pBsp6 suggest that cleavage at the 16-kDa–13-kDa junction occurs on the UL80 precursor polyprotein rather than directly from preexisting 30-kDa protein. Each construct contains a stop codon within the 50-kDa region of UL80 (Fig. 1B). pBam17 terminates at Asp-438, and pBsp6 terminates at Thr-354, compared with Ala-643 for the 50-kDa protein. As expected, the 30-, 16-, and 13-kDa species were detected in radiolabelled *E. coli* cells harboring pBam17 or pBsp6 (Fig. 7A). As noted for the experiment shown in Fig. 2B, an ~28-kDa protein which may be the truncated version of the 50-kDa protein was observed for pBam17; the corresponding truncation product for pBsp6 was not observed and may be unstable. A new protein which is not present from parental pCPRT3 was observed for each construct (Fig. 7A [arrows]). These bands are consistent with full-length polyproteins with molecular masses of ~52 kDa for pBam17 and

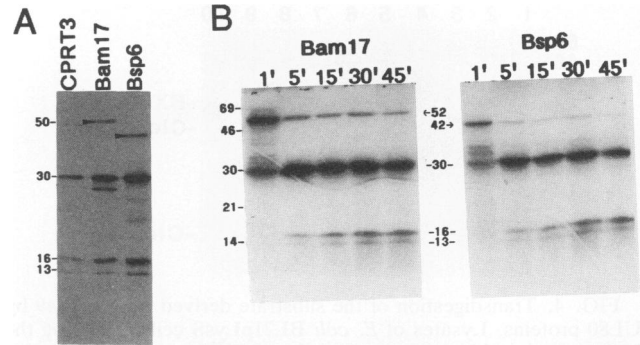


FIG. 7. (A) Metabolic labelling of UL80 constructs in *E. coli* BL21pLysS. Cells harboring the indicated plasmids were labelled with [³⁵S]methionine plus [³⁵S]cysteine and then subjected to SDS-PAGE and autoradiography. The putative full-length polyproteins from bacteria harboring the pBam17 and pBsp6 constructs are shown (arrows). (B) Pulse-chase radiolabelling of *E. coli* BL21pLysS cells harboring pBam17 or pBsp6. The experiment was performed as described in the legend to Fig. 6, except that both [³⁵S]methionine and [³⁵S]cysteine were used for the pulse and both methionine and cysteine were used for the chase. The times of chase are shown in minutes.

of ~42 kDa for pBsp6. Full-length 85-kDa polyprotein was not observed for pCPRT3, suggesting either that cleavage of pBam17 and pBsp6 polyprotein is less efficient than that of pCPRT3 or that the rate of synthesis of pBam17 and pBsp6 polyproteins is higher than that of pCPRT3 polyprotein.

A pulse-chase experiment with *E. coli* cells harboring pBam17 and pBsp6 was performed (Fig. 7B). The full-length polyprotein (Fig. 7B, arrows) decreased in abundance between 1 and 5 min, as the amounts of 30-, 16-, and 13-kDa proteins increased. As was seen in the pulse-chase experiment with *E. coli* cells harboring pCPRT3 (Fig. 6), the 16- and 13-kDa species did not appear at the expense of the 30-kDa species. The disappearance of the polyproteins during the chase coupled with the apparent stability of the 30-kDa species strongly suggests that cleavage at the 16-kDa–13-kDa junction occurs only in the polyprotein and not within preexisting 30-kDa protein. Thus, the N terminus of UL80 polyprotein appears to have two possible fates: cleavage to generate the 16- and 13-kDa proteins and cleavage to generate the 30-kDa protein.

Detection of cleavage at the 16-kDa–13-kDa junction in HCMV-infected cells. If cleavage at the 16-kDa–13-kDa junction is a potential mechanism to downregulate the expression of the 30-kDa protease, it should be possible to detect the 16- and/or 13-kDa product during infection by HCMV. Accordingly, HFF were infected with HCMV strain AD169 and labelled with [³⁵S]methionine plus [³⁵S]cysteine from 68 to 72 h postinfection. Lysates from infected and uninfected cells were immunoprecipitated with polyclonal antiserum raised against the 16-kDa protein. This serum recognizes the 16- and 30-kDa, but not the 13-kDa, UL80 proteins expressed in *E. coli* (Fig. 8, lower panel, lane E). Four proteins were specifically immunoprecipitated from HCMV-infected cells: ~85- and ~80-kDa proteins (Fig. 8, upper panel) and 30- and 16-kDa proteins (lower panel). These proteins were not immunoprecipitated from uninfected cells or with pre-immune serum.

The 30- and 16-kDa proteins from virus-infected cells comigrated with their counterparts expressed in *E. coli* cells (Fig. 8, lower panel). Thus, these proteins are produced

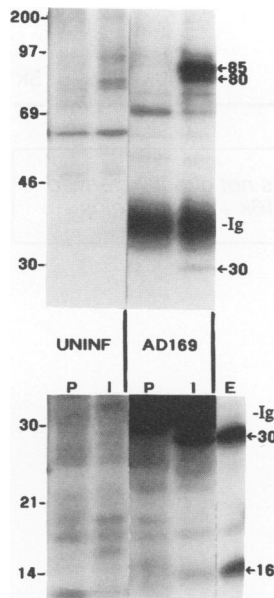


FIG. 8. UL80 proteins produced during HCMV infection. Uninfected HFF (UNINF) or HFF infected with HCMV strain AD169 were radiolabelled with [³⁵S]methionine and [³⁵S]cysteine from 68 to 72 h postinfection. Lysates were immunoprecipitated with either preimmune serum (P) or immune serum (I) raised against the 16- and 30-kDa UL80 proteins and subjected to SDS-PAGE and autoradiography. The upper panel is a 1-day exposure of the upper portion, and the lower panel is a 14-day exposure of the lower portion of the same gel. Lane E, immunoprecipitate of radiolabelled *E. coli* cells harboring pBam17. Molecular mass markers (in kilodaltons) are shown on the left. An unidentified protein (Ig) with a molecular mass of approximately 35 kDa is nonspecifically immunoprecipitated from AD169-infected cells by almost every serum sample tested and may be an immunoglobulin receptor-like protein (12a).

during the HCMV replicative cycle. The 85-kDa species is full-length UL80 polyprotein, and the 80-kDa species is probably UL80 polyprotein which has been cleaved at the assembly protein cleavage site (Ala-643-Ser-644; Fig. 1A), with 5 kDa removed from the C terminus as a result. The relative abundance of these four proteins at 68 to 72 h postinfection (85 ~ 80 ≫ 30 ≫ 16 kDa) is quite different from that of UL80 expression in *E. coli* (16 ~ 30 ≫ 80, 85 kDa) (from pCPRT3; Fig. 2A). This result may be due to differential cleavages at the various sites in the UL80 polyprotein in HCMV-infected cells or to differential stabilities of the various UL80-derived proteins. The results from Western blots (data not shown) are consistent with the immunoprecipitation data.

Separation of *cis* and *trans* activities by alteration of the C terminus of the 30-kDa protein. Several constructs which result in production of the 30-kDa protein with either small deletions or additions at its C terminus were generated. Constructs p30k(-8) and p30k(-3) are deleted of eight and three amino acids, respectively, from the C terminus of the 30-kDa protein. Similarly, p30k(+12) contains an extra 12 amino acids (a *c-myc* epitope) at the C terminus. The proteins produced by *E. coli* cells harboring these constructs were labelled with [³⁵S]methionine and [³⁵S]cysteine and examined by SDS-PAGE and autoradiography. p30k, encoding the wild-type 30-kDa protein, produced the 30-, 16-, and 13-kDa proteins in ratios comparable to those produced by pCPRT3 (Fig. 9A). In contrast, deletion of as few as three

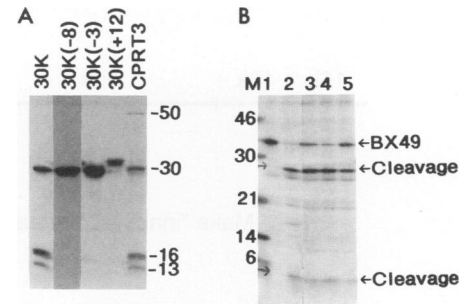


FIG. 9. (A) Proteins generated by the 30-kDa protease construct and by C-terminal mutants. *E. coli* BL21pLysS cells harboring the indicated plasmids (see text) were labelled with [³⁵S]methionine plus [³⁵S]cysteine and then subjected to SDS-PAGE and autoradiography. 30k, the wild-type protein which terminates at Ala-256; deletion or addition of amino acids at the C terminus is indicated in parentheses. Wild-type UL80 proteins encoded by pCPRT3 are shown as a control. (B) Cleavage of pBX49-derived substrate by 30-kDa protease and by C-terminal mutants. Protein (2 to 10 μg) from urea-solubilized inclusion bodies from *E. coli* BL21pLysS cells harboring the indicated plasmids was incubated with [³⁵S]methionine-labelled pBX49-derived substrate and then subjected to SDS-PAGE and autoradiography. Lanes: 1, substrate only; 2, pCPRT3; 3, p30k(-8); 4, p30k; 5, p30k(+12).

amino acids from the C terminus drastically reduced production of the 16- and 13-kDa proteins, by at least 90% [Fig. 9A, lane 30k(-3)]. Similar results are obtained upon deletion of 8 amino acids [Fig. 9A, lane 30k(-8)] or addition of 12 amino acids [lane 30k(+12)]. These data indicate that the authentic C terminus of the 30-kDa protein is required for efficient cleavage at the 16-kDa-13-kDa junction. This suggests that in *E. coli* cells harboring pCPRT3, the 30-kDa protease must be released from the precursor polyprotein for cleavage at the 16-kDa-13-kDa junction to occur. The 16- and 13-kDa bands are also observed for pBsu26 (Fig. 2B), which terminates at Pro-260, four amino acids downstream of the C terminus of the 30-kDa protein. This implies either that these extra four amino acids are cleaved from the C terminus or that they do not impede 16-kDa-13-kDa cleavage if present.

The abilities of the 30-kDa protease and of the C-terminal mutants to cleave pBX49-derived substrate *in trans* were examined. These mutants were able to cleave the 50-kDa-5-kDa junction of that substrate (Fig. 9B, lanes 3 to 5). Thus, mutations at the C terminus severely diminish protease activity at the 16-kDa-13-kDa bond but not at the 50-kDa-5-kDa bond *in trans*. Note that when unlabelled wild-type protease was added to labelled lysates of bacteria harboring p30k(-8), p30k(-3), or p30k(+12), the labelled 30-kDa mutant proteins were not further cleaved into the 16- and 13-kDa products (data not shown). It is possible that cleavage at the 16-kDa-13-kDa bond can occur only intramolecularly or that cleavage occurs cotranslationally and that the 16-kDa-13-kDa bond is inaccessible once the 30-kDa protein is released from the UL80 polyprotein. In either case, it appears that the wild-type C terminus of the 30-kDa protease is required for this cleavage.

DISCUSSION

We have used expression in *E. coli* to demonstrate that the HCMV UL80 ORF encodes a protease at its N terminus. The presence of a protease in HCMV UL80 was postulated by Liu and Roizman (17) from studies of HSV-1 ORF UL26,

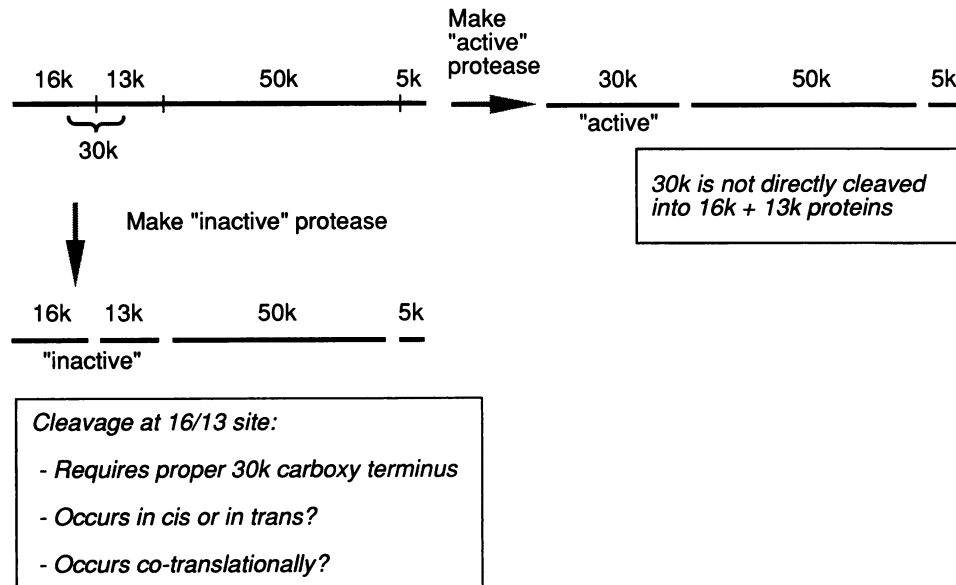


FIG. 10. Model of UL80 polyprotein-processing pathway. A two-branch pathway is postulated (see Discussion), which results in the production of active 30-kDa protease or inactive 16- and 13-kDa proteins. Once formed, the 30-kDa protease appears to be stable and is not further cleaved into the 16- and 13-kDa proteins.

which, like UL80, encodes the assembly protein. The SCMV homolog of UL80, APNG1, was recently shown to encode a protease (31). Previous studies with HSV, SCMV, and HCMV demonstrated that the assembly protein undergoes proteolytic cleavage near its C terminus (8, 24, 28); in HSV-1, this cleavage is presumably required for packaging of progeny DNA into the capsid (11, 14, 24, 25). We now show that a 30-kDa protein is generated from autocatalytic digestion of 85-kDa UL80 polyprotein and is able to cleave a substrate containing the assembly protein cleavage site. Thus, the 30-kDa protein encoded by the N-terminal region of UL80 appears to be the protease.

The 85-kDa UL80 polyprotein is cleaved at three different sites (the 16-kDa–13-kDa, 30-kDa–50-kDa, and 50-kDa–5-kDa junctions; Fig. 1), generating five major products (with molecular masses of 16, 13, 30, 50, and 5 kDa). Cleavage at the 30-kDa–50-kDa junction presumably releases the mature 30-kDa protease and a 50-kDa protein which is an N-terminal extension of the 40-kDa assembly protein. The 50-kDa protein is probably analogous to the SCMV 45-kDa protein observed by Schenk et al. (28), and its function, if any, is unknown.

For HCMV-infected cells at late times postinfection, we show that the 85-kDa species (intact UL80 polyprotein) is slightly more abundant (approximately threefold) than the 80-kDa species (UL80 polyprotein which has been cleaved at the assembly protein cleavage site) (Fig. 8). Together, the 85- and 80-kDa proteins are approximately 10- to 20-fold more abundant than the 30-kDa protein. Thus, the 30-kDa species is a relatively minor component in HCMV-infected cells. In contrast, for UL80 expressed in *E. coli*, cleavage at both the 50-kDa–5-kDa and the 30-kDa–50-kDa junctions is essentially complete, since no polyprotein precursors are detected. This suggests that the extent of cleavage at the 50-kDa–5-kDa and the 30-kDa–50-kDa junctions is regulated in virus-infected cells. Alternatively, the cleavage products may exhibit differential stabilities.

It is not known whether the 80- and 85-kDa UL80 poly-

proteins possess protease activity. Mutagenesis of the 30-kDa–50-kDa junction of the UL80 gene expressed in HCMV would address this point more completely. Cleavage at the 30-kDa–50-kDa site does appear to be required for the 30-kDa protein to acquire its full range of proteolytic activity, since fusion to the C terminus of the 30-kDa protease [e.g., Fig. 9, lane 30k(+12)] abolished the ability to cleave at the 16-kDa–13-kDa junction but not at the 50-kDa–5-kDa junction. This is reminiscent of the poliovirus 3CD (polymerase protease) polyprotein, which exhibits protease activity but has substrate specificity different from that of mature 3C protease (32).

A recent publication by Welch et al. (31) also demonstrated the existence of a 30-kDa protease at the N terminus of the UL80 homolog (APNG) from SCMV. By transfecting portions of the APNG ORF into mammalian cells, they deduced that cleavage occurs at the 30-kDa–50-kDa junction (Ala-256–Ser-257) and at the 50-kDa–5-kDa junction (Ala-643–Ser-644), in agreement with our results with prokaryotic cells. We have detected an additional cleavage event at Ala-143–Ala-144, which generates 16- and 13-kDa proteins whose sequences are entirely contained within the 30-kDa protein. The 16-kDa–13-kDa cleavage site (Ala-Ala) is ~50% cleaved in *E. coli*, compared with quantitative cleavage at the Ala-Ser sites of the 30-kDa–50-kDa and 50-kDa–5-kDa junctions. Computer alignment indicates that the SCMV APNG studied by Welch et al. (30, 31) lacks the exact sequence of the HCMV UL80 16-kDa–13-kDa cleavage site (data not shown); cleavage may occur elsewhere, or not at all, within the SCMV 30-kDa protease. Furthermore, since Welch et al. (31) monitored cleavage by epitope tagging at either the C terminus of APNG or within a region that apparently inactivated the protease, the 16- and 13-kDa proteins, if present, would not be detected.

We initially detected the 16- and 13-kDa proteins by expression of UL80 in *E. coli*, but, more importantly, this cleavage event also occurs during virus replication, as demonstrated by immunoprecipitation of the 16-kDa protein

by antibodies to the N terminus of UL80. However, at late times postinfection, the 16-kDa protein is relatively rare (~1/10 the abundance of the 30-kDa protein, which is itself ~1/10 the abundance of the 80- plus 85-kDa proteins). Thus, it remains to be determined whether the 16-kDa-13-kDa cleavage event plays a role in the virus life cycle. Cleavage at the 16-kDa-13-kDa site separates the two conserved domains, CD2 and CD1, which are present in all putative herpesvirus proteases and which may be important for function (31). Thus, this cleavage possibly is a mechanism to inactivate the protease. The individual 16- and 13-kDa proteins lack detectable protease activity (1). Additionally, the 16- and/or 13-kDa protein may serve some unknown function in HCMV replication; this is currently under investigation.

The production of the 16- and 13-kDa proteins has several interesting features. Although the 16- plus 13-kDa sequences constitute the 30-kDa protein, the resistance of the 30-kDa species to further cleavage and its stability in pulse-chase experiments suggest that the 30-kDa protein is not the direct precursor of the 16- and 13-kDa proteins. Instead, cleavage at the 16-kDa-13-kDa junction appears to occur on the UL80 polyprotein. Since the 30-kDa protein is stable, it is possible that cleavage at the 16-kDa-13-kDa junction occurs cotranslationally and that this site is inaccessible once the 30-kDa protein is released from the UL80 precursor. These features are summarized in the model diagrammed in Fig. 10. A two-branch pathway is postulated: the first cleavage occurs either at the 30-kDa-50-kDa junction, producing active 30-kDa protease, or at the 16-kDa-13-kDa junction, producing proteolytically inactive 16- and 13-kDa proteins. The two branches can be separated by mutation of the C terminus of the 30-kDa protease; whereas wild-type protease carries out both branches, C-terminal mutants are defective for cleavage at the 16-kDa-13-kDa junction.

In *E. coli*, the amount of 30-kDa protein produced is about the same as the amount of 16- or 13-kDa protein produced; this suggests that the two putative branches are utilized approximately equally. In contrast, in virus-infected cells at late times postinfection, the 30-kDa species is ~10-fold more abundant than the 16-kDa species, suggesting that the branch to generate active protease is utilized almost exclusively. If cleavage at the 16-kDa-13-kDa junction is a regulatory mechanism in HCMV, then a shift in the ratio of active to inactive branch utilization may occur during virus replication. To determine the relevance of cleavage at the 16-kDa-13-kDa junction and at the other UL80 junctions, mutations of these sites can be analyzed within the context of the virus via gene replacement (12, 13).

ACKNOWLEDGMENTS

We thank J. Morin, M. Vitek, S. Plotch, and L. Li for useful criticism of the manuscript.

REFERENCES

- Baum, E. Z., and G. A. Bebernitz. Unpublished data.
- Baum, E. Z., G. A. Bebernitz, and Y. Gluzman. 1990. Isolation of mutants of human immunodeficiency virus protease based on the toxicity of the enzyme in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **87**:5573-5577.
- Baum, E. Z., G. A. Bebernitz, O. Palant, T. Mueller, and S. J. Plotch. 1991. Purification, properties, and mutagenesis of poliovirus 3C protease. *Virology* **185**:140-150.
- Braun, D. K., B. Roizman, and L. Pereira. 1984. Characterization of posttranslational products of herpes simplex virus gene 35 products binding to the surface of full but not empty capsids. *J. Virol.* **46**:103-112.
- Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchinson, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. Weston, and B. G. Barrell. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.* **154**:125-169.
- Collins, J. 1979. Cell-free synthesis of proteins coding for mobilization functions of ColE1 and transposition functions of Tn3. *Gene* **6**:29-42.
- Evan, G. I., G. K. Lewis, G. Ramsay, and J. M. Bishop. 1985. Isolation of monoclonal antibodies specific for human *c-myc* proto-oncogene product. *Mol. Cell. Biol.* **5**:3610-3616.
- Gibson, W. 1983. Protein counterparts of human and simian cytomegaloviruses. *Virology* **128**:391-406.
- 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 carboxy-terminal end. *J. Virol.* **64**:1241-1249.
- 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.
- Irmiere, A., and W. Gibson. 1983. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology* **130**:118-133.
- Irmiere, A., and W. Gibson. 1985. Isolation of human cytomegalovirus intranuclear capsids, characterization of their protein constituents, and demonstration that the B-capsid assembly protein is also abundant in noninfectious enveloped particles. *J. Virol.* **56**:277-283.
- Jones, T. R., and V. P. Muzithras. 1992. A cluster of dispensable genes within the human cytomegalovirus genome short component: IRS1, US1 through US5, and the US6 family. *J. Virol.* **66**:2541-2546.
- Jones, T. R., and V. P. Muzithras. Unpublished data.
- Jones, T. R., V. P. Muzithras, and Y. Gluzman. 1991. Replacement mutagenesis of the human cytomegalovirus genome: US10 and US11 gene products are nonessential. *J. Virol.* **65**:5860-5872.
- Lee, J. Y., A. Irmiere, and W. Gibson. 1988. Primate cytomegalovirus assembly: evidence that DNA packaging occurs subsequent to B capsid assembly. *Virology* **167**:87-96.
- 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.
- Liu, F., and B. Roizman. 1991. The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. *J. Virol.* **65**:5149-5156.
- Liu, F., and B. Roizman. 1992. Differentiation of multiple domains in the herpes simplex virus 1 protease encoded by the UL26 gene. *Proc. Natl. Acad. Sci. USA* **89**:2076-2080.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**:10035-10038.
- 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.
- Moffatt, B. A., and F. W. Studier. 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49**:221-227.
- Oram, J. D., R. G. Downing, A. Alkrigg, A. A. Dollery, C. J. Duggleby, G. W. G. Wilkinson, and P. J. Greenaway. 1982. Use of recombinant plasmids to investigate the structure of the human cytomegalovirus genome. *J. Gen. Virol.* **59**:111-129.
- Plotch, S. J., O. Palant, and Y. Gluzman. 1989. Purification and properties of poliovirus RNA polymerase expressed in *Es-*

- cherichia coli*. J. Virol. **63**:216–225.
24. **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.
 25. **Preston, V. G., F. J. Rixon, I. M. McDougall, M. McGregor, and M. F. Al Kobaisi.** 1992. Processing of the herpes simplex virus assembly protein ICP35 near its carboxy terminal end requires the product of the whole of the UL26 reading frame. Virology **186**:87–98.
 26. **Robson, L., and W. Gibson.** 1989. Primate cytomegalovirus assembly protein: genome location and nucleotide sequence. J. Virol. **63**:669–676.
 27. **Rosenberg, A. H., B. N. Lade, D.-S. Chui, S.-W. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned cDNAs by T7 RNA polymerase. Gene **56**:125–135.
 28. **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.
 29. **Sherman, G., and S. L. Bachenheimer.** 1988. Characterization of intranuclear capsids made by ts morphogenic mutants of HSV-1. Virology **163**:471–480.
 30. **Welch, A. R., L. M. McNally, and W. Gibson.** 1991. Cytomegalovirus assembly protein nested gene family: four 3'-coterminally transcripts encode four in-frame, overlapping proteins. J. Virol. **65**:4091–4100.
 31. **Welch, A. R., A. S. Woods, L. M. McNally, R. J. Cotter, and W. Gibson.** 1991. A herpesvirus maturational proteinase, assemblin: identification of its gene, putative active site domain, and cleavage site. Proc. Natl. Acad. Sci. USA **88**:10792–10796.
 32. **Ypma-Wong, M. F., P. Gillis Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166**:265–270.
 33. **Zweig, M., C. J. Heilman, H. Rabin, and B. Hampar.** 1980. Shared antigenic determinants between two distinct classes of proteins in cells infected with herpes simplex virus. J. Virol. **35**:644–652.