Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 10792–10796, December 1991
Biochemistry

A herpesvirus maturational proteinase, assemblin: Identification of its gene, putative active site domain, and cleavage site

ANTHONY R. WELCH*, AMINA S. WOODS†, LISA M. MCNALLY*, ROBERT J. COTTER†, AND WADE GIBSON*‡

*Virology Laboratories and †Middle Atlantic Mass Spectrometry Laboratory, Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, Baltimore, MD 21205

Communicated by Paul Talalay, September 5, 1991 (received for review August 12, 1991)

ABSTRACT **A herpesvirus proteinase activity has been identified and partially characterized by using the cloned enzyme and substrate genes in transient transfection assays. Evidence is presented that the proteinase gene of cytomegalovirus strain Colburn encodes a 590-amino acid protein whose N-terminal 249 residues contain the proteolytic activity and two domains that are highly conserved in the homologous protein of other herpesviruses. Insertion of a short amino acid sequence between these domains abolished proteinase activity, suggesting that this region constitutes part or all of the enzyme active site. Plasma desorption mass spectrometry was used to identify the C terminus of the mature assembly protein as alanine, enabling the recognition of a consensus proteinase cleavage sequence of V/L-X-A ↓ S/V, near the C-terminal end of all herpesvirus assembly protein homologs. Interestingly, the proteinase and its substrate, the assembly protein precursor, are encoded by opposite halves of the same open reading frame.**

Herpes group viruses have a common intranuclear assembly intermediate referred to as the B- or intermediate capsid (1, 2). These particles lack DNA and accumulate in the nucleus of infected cells that have been blocked chemically (ref. 3, and references therein) or genetically (4–6) at the level of viral DNA synthesis. Like bacteriophage proheads, herpesvirus B-capsids or procapsids contain an abundant protein constituent that is not found in mature virions (1, 7, 8). This species is referred to as the scaffolding protein in bacteriophage (9) and as the assembly protein in herpesviruses (10). Unlike the other B-capsid proteins, the assembly protein is phosphorylated (3, 11, 12) and undergoes proteolytic processing (1, 5, 12) that eliminates its C-terminal end (13, 14), but the enzymes responsible have not been identified.

The assembly protein coding sequence is contained within a family of four in-frame overlapping 3′-coterminal genes, referred to as the assembly protein nested genes (APNGs), each of which gives rise to a separate transcript and protein (15). The largest of these four nested genes was designated as APNG1 and the sequence encoding the assembly protein precursor was designated as APNG.5 (15). Homologs of APNG1 and APNG.5 have been shown to be similarly organized (15) and expressed (16) in other herpesviruses. The study described here was prompted by the finding that the 5′ half of the longest open reading frame in each of the homologous genes encodes a highly conserved region with similarities to proteinase active site domains (see below). Experiments were done to determine whether the assembly protein precursor undergoes proteolytic processing in the absence of other viral genes, to test the possibility that its larger related form (i.e., the product of APNG1) may be the proteinase required for assembly protein maturational cleavage, and to determine the C terminus of the mature assembly protein and thereby deduce the proteinase cleavage site.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Cells and Virus. Human foreskin fibroblasts (HFFs) were prepared, grown, and infected with cytomegalovirus (CMV; strain Colburn) as described (7). Human embryonal kidney (HEK) cell line 293 (American Type Culture Collection) was grown in Dulbecco's modified Eagle's medium containing 10% (vol/vol) fetal calf serum. Transfections were done in two-well chambers (177380, Nunc) containing approximately 1×10^6 cells in 1 ml of medium per well.

Cloning, Plasmid Construction, and *in Vitro* Transcription/Translation. Standard techniques were used to construct, clone, and propagate the plasmids (17). The longest gene in the APNG family (APNG1) and the gene encoding the assembly protein itself (APNG.5) (15) were amplified from a plasmid containing the viral genomic *Xba* I R fragment (18) by using the PCR (19), and cloned into the eukaryotic expression vector pRSV.5(neo) (20) to give AW1 and AW4, respectively. APNG1 and APNG.5 were subcloned from AW1 and AW4 into pGEM-4Z (Promega) for *in vitro* transcription (17). The sequence DNPASTTNKDKLHYV, containing the 12-amino acid (underlined) poliovirus epitope C3 (21), was cloned into the *Dra* III site of AW1 and AW4 to make LM1 and LM2, respectively. LM3 contains the sequence VDNPASTTNKDKLH (C3 epitope, underlined) at the *Eco*RV site of AW4. Two deletion constructs of APNG1, LM7 and LM8, were made by PCR cloning as described above. A third deletion construct, AW5, was made by removing the *Sal* I–*Eco*47III fragment from AW4, which resulted in a noncoded sequence, Ile-Gln-Thr, being added to the C-terminal Ala-179.

Transfection Assay. Approximately 1×10^6 cells were transfected (22) with 1 μg (A_{260}) of each indicated DNA, rinsed with phosphate-buffered saline (no Ca^{2+} or Mg^{2+}) 18–24 hr later, scraped from the slide in 50 μl of solubilizing buffer (23), boiled for 3 min, and analyzed.

Protein Assays. Proteins were separated by SDS/polyacrylamide gel electrophoresis (23). Towbin immunoassays (24) were done as described (14): rabbit antisera anti-N1 and anti-C1 (see Fig. 1) were used at dilutions of 1:50 and 1:80, respectively, and ^{125}I-labeled protein A was used to visualize bound antibodies.

Amino Acid Sequence Analyses. The assembly protein was isolated by HPLC from Colburn CMV B-capsids that had been disrupted in 6 M guanidine hydrochloride/10% (vol/vol) 2-mercaptoethanol and heated for 3 min at 45°C (13). Preparations of assembly protein were digested with endoproteinase Lys-C or endoproteinase Glu-C; resulting fragments were separated by HPLC; plasma desorption mass spectrometry (PDMS) was done with a Bio-Ion Bin-10K spectrometer (Nordic, Lausanne, Switzerland); and "on foil" carboxypeptidase P treatments were done, all as described (25).

Abbreviations: CMV, cytomegalovirus; PDMS, plasma desorption mass spectrometry; APNG, assembly protein nested gene; CD, conserved domain.

‡To whom reprint requests should be addressed.

RESULTS

The genetic constructs used and their protein products are summarized in Fig. 1. Altered forms of the wild-type proteins are indicated by the prefix Δ followed by the numerical designation of the construct.

Two Highly Conserved Sequences Are Present at the Amino End of the APNG1 Protein. When the predicted amino acid sequence of the APNG1 protein pNP1 was compared with those of its homologs in other herpesviruses (15), two highly conserved domains, CD1 and CD2, separated by 80 ± 4 residues were identified (Table 1). The conservation of CD1 and CD2 sequences suggests that they are functionally important, and the presence of the indicated (His, Asp/Glu, Cys) and (His, Asp/Glu, Ser) triads raised the possibility that this region may contain the active site of a cysteine- or serine-like proteinase (27, 28).

APNG1 Encodes Proteinase Able to Cleave Assembly Protein Precursor. When APNG.5 and APNG1 were transcribed *in vitro* and the resulting RNAs were translated, APNG.5 gave rise to a single protein (Fig. 2*A*) that comigrated with the infected-cell assembly protein precursor pAP (Fig. 2*B*), and APNG1 yielded the protein pNP1 (Fig. 2*A*) that comigrated with the lower band of the infected-cell 85/91-kDa doublet (Fig. 2*B*). Cleavage of the assembly protein precursor to its mature form was not observed in these preparations, whether they were kept separate or combined and incubated for up to 18 hr at 37°C (data not shown; see *Discussion*).

To determine whether the proteins would behave differently in living cells, transfections were done with AW1 and AW4 (Fig. 1). In AW1-transfected cells only the assembly protein precursor was detected (Fig. 2*C*, lane 2). In AW4-transfected cells two proteins were detected: a faint ≈82-kDa species, NP1, and a prominent ≈45-kDa species, $NP1_C$ (Fig. 2*C*, lane 4). In contrast to the results described above, AW1/AW4-cotransfected cells did show cleavage of the assembly protein precursor to its mature form (Fig. 2*C*, lane 3). None of the proteins expressed in AW1/AW4-transfected cells was detected by anti-C1 (e.g., Fig. 3*C*, lane 6), indicating that each had lost its C-terminal end, as would be expected for correct maturational cleavage.

The assembly protein precursor cleavage seen in AW1/AW4 cotransfections is consistent with APNG1 encoding the maturational proteinase. Furthermore, the finding that the proteins detected in AW4-transfected cells were smaller than the 85-kDa pNP1 synthesized *in vitro* and did not react with anti-C1 suggested that pNP1 itself was cleaved.

Table 1. Herpesvirus proteinase homologs have two highly conserved domains, CD1 and CD2

Virus	Sequence	
	CD2	CD1
SCMV	PLPLNVNHDESATVGYV	.D.FKHVALCSVGRRRGTLAVYG.S
HCMV	ALPLNINHDDTAVVGHV	.D.FKHVALCSVGRRRGTLAVYG.S
HSV-1	PLPINVDHRAGCEVGRV	.E.FAHVALCAIGRRLGTIVTYD.S
VZV	KIPINIDHRKDCVVGEV	.E.FTHVALCVVGRRVGTVVNYD.S
EBV	PLPLTVEHLPDAPVGSV	.E.FDHVSICALGRRRGTTAVYG.S
ILTV	TIPINIDHESSCVVGTV	.E.FAHVALCELGRREGTVAIYG.S
HHV-6		.D.FHHV
T4	GVPMLESHDGKDLGLKP	.E.IPGVSSRGLGSLTDTNKGYR.S
	H←55±2→D/E⟨40±3⟩C	
	H←55±2→D/E←97±11→S	

Partial sequence for HHV-6 was kindly provided by T. Dambaugh (Dupont); the T4 proteinase (gp21) and other sequences are from GenBank (see refs. 15 and 26). Absolutely conserved residues are underlined. Number of residues separating the potential catalytic triads (His, Asp/Glu, Cys) and (His, Asp/Glu, Ser) are indicated. SCMV, simian CMV strain Colburn; HCMV, human CMV strain AD169; HSV-1, herpes simplex virus type 1; VZV, varicella zoster virus; EBV, Epstein–Barr virus; ILTV, infectious laryngotracheitis virus; HHV-6, human herpes virus 6; and T4, bacteriophage T4.

Proteinase Activity Abolished by Inserting a Peptide Between CD1 and CD2. Additional evidence that APNG1 encodes the maturational proteinase was obtained with the LM3 construct (Fig. 1) that showed three major differences from AW4 in transfection assays. (*i*) The largest protein, Δ3pNP1 doublet, was remarkably more abundant than NP1 and no $NP1_C$ was detected (Fig. 2*C*, lane 10). (*ii*) Unlike NP1, the Δ3pNP1 doublet reacted with anti-C1 (Fig. 3*C*, lanes 7 and 12, respectively). (*iii*) There was no cleavage of the assembly protein precursor in cotransfections with AW1 (Fig. 2*C*, lane 11). These results suggest that the putative active site domain of pNP1 participates in at least three separate cleavages: (*i*) pNP1 → $NP1_N$ + $NP1_C$; (*ii*) pNP1 → NP1; and (*iii*) pAP → AP (see Fig. 5).

$NP1_C$ Is the C-terminal Portion of pNP1. The only differences detected between transfections with LM2 (Fig. 1) and AW4 were the sizes of their protein counterparts. LM2 gave rise to slightly larger forms of pNP1 (Δ2pNP1 at ≈87 kDa; data not shown), NP1 (Δ2pNP1 at ≈85 kDa), and $NP1_C$ (Δ2$NP1_C$ at ≈48 kDa), due to the insertion (Fig. 2*C*, lane 8). The increased size of Δ2$NP1_C$ identified it as the C-terminal portion of Δ2pNP1; this conclusion is supported by the LM3/LM7 and LM3/LM8 cotransfections described below.

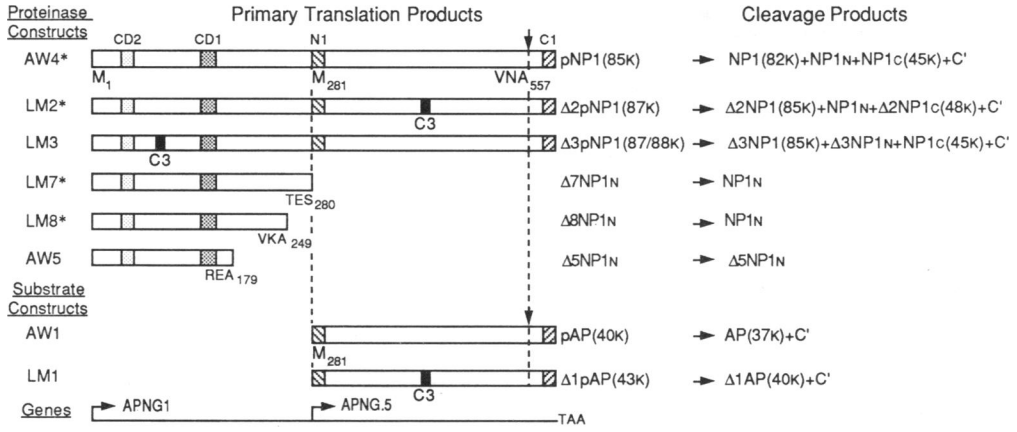

FIG. 1. Genetic constructs and antisera used. The genetic constructs (asterisk denotes active proteinase) and their protein products (size in kDa) are shown. Also indicated are conserved domains 1 and 2 (CD1 and CD2, respectively; shaded boxes); translational start methionines for pNP1 (M-1) and pAP (M-281); C-terminal residues (amino acid triplets), AW5 contains the C-terminal noncoded sequence Ile-Gln-Thr; maturational cleavage site (arrows); poliovirus C3 epitope (21) (solid boxes); and N1 and C1 (hatched regions), the peptides recognized by sera anti-N1 and anti-C1 (11, 12), respectively. C' is the 33-amino acid fragment eliminated by maturational cleavage.

Unlike LM3, LM2 yielded all three proteolytic activities noted for AW4, including cleavage of the assembly protein precursor in cotransfections with AW1 (Fig. 2*C*, lane 9): i.e., $\Delta 2NP1 \rightarrow \Delta 2NP1_C + NP1_N$; $\Delta 2pNP1 \rightarrow \Delta 2NP1$; and $pAP \rightarrow AP$.

Insertion of the same sequence in the assembly protein precursor (LM1, Fig. 1) increased its size by ≈3 kDa (Δ1pAP) but did not affect its ability in cotransfections to serve as a substrate for AW4 (Fig. 2*C*, lanes 6 and 7, respectively).

Proteinase Activity Lies Within the First 249 Amino Acids of pNP1. The region of APNG1 encoding the proteinase was determined by testing three deletion constructs in transfection assays. The shortest of these, AW5 (Fig. 1), did not produce cleavage of either the assembly protein precursor (Fig. 3, lane 10) or Δ3pNP1 (Fig. 3, lane 16). LM7 and LM8 (Fig. 1), however, did produce cleavage of both (Fig. 3, lanes 8 and 9; refs. 14 and 15—note appearance of Δ3NP1 in Fig. 3*B*). Because LM7 and LM8 do not code for the C-terminal portion of pNP1, the $NP1_C$ band seen in LM7 and LM8 cotransfections with LM3 (Fig. 3*A*, lanes 14 and 15) must come from Δ3pNP1.

Identification of the Assembly Protein Maturational Cleavage Site. The P_1 amino acid (see ref. 29) at the cleavage site was identified by determining the C-terminal amino acid sequence of the mature assembly protein by PDMS. The masses of 23 Lys-C and 20 Glu-C peptides from the purified assembly protein (Fig. 4*A*) were determined. The protonated molecular ions of Lys-C fraction 24 (Fig. 4*B*) and Glu-C fraction 23 (Fig. 4*D*) were 903.3 (Fig. 4*C*) and 616.2 (Fig. 4*E*), respectively. These masses agree well with those calculated for the peptides SAERGVVNA (902.5) and RGVVNA (614.4) and indicate that Ala-557 (in the pNP1 sequence) is the C terminus of the mature assembly protein. To verify this result, Lys-C fraction 24 and Glu-C fraction 23 were each subjected to digestion with carboxypeptidase P and then analyzed by PDMS. The spectrum of carboxypeptidase-treated Lys-C fraction 24 showed a significant reduction in the signal at m/z 903.3 and the appearance of a new peak at m/z 462.2, which corresponds well with the predicted mass of the original peptide minus the five C-terminal amino acids GVVNA (i.e., 903.3 − 440 = 463.3 atomic mass units). The spectrum of carboxypeptidase-treated Glu-C fragment 23 showed a new peak at m/z 547.6, which corresponds well with the predicted mass of the original peptide minus one alanine (i.e., 616.2 − 71 = 545.2 atomic mass units), confirming that the last residue of the assembly protein is alanine.

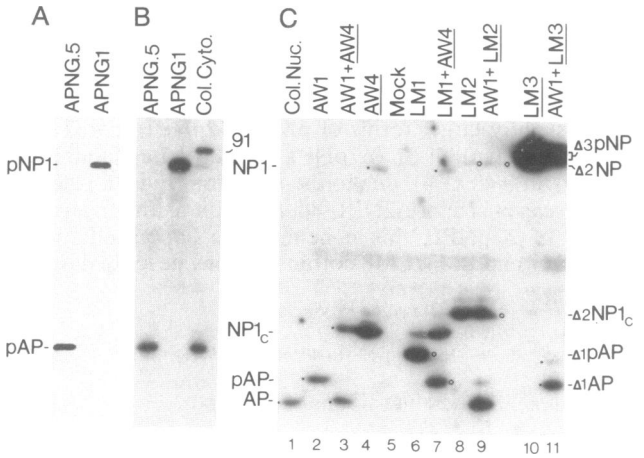

FIG. 2. Protein products of wild-type and insertion-mutated APNG1 and APNG.5. (*A*) [^{35}S]Methionine-labeled proteins synthesized from APNG1 and APNG.5 RNA in a rabbit reticulocyte lysate (Amersham). (*B*) An immunoblot probed with anti-C1 comparing the sizes of APNG1 and APNG.5 proteins made *in vitro*, with proteins in the cytoplasmic fraction of Colburn-infected HFF cells (Col. Cyto.). (*C*) An immunoblot showing proteins reactive with anti-N1 in HEK cells transfected with wild-type constructs AW1 and AW4 or insertion constructs LM1, LM2, and LM3, either alone or in combination. Underlined constructs are potential proteinase donor. Nuclear fraction of Colburn-infected HFF cells (Col. Nuc.) was used for comparison. Proteins are designated as described in Fig. 1. Proteins of altered genes are indicated in the right margin and by a circle to right of the corresponding band. Proteins of wild-type genes are indicated in the left margin and by dot to left of the corresponding band. Δ3pNP1 doublet is seen better in Fig. 3*B*.

DISCUSSION

Proteolytic processing of a highly conserved herpesvirus protein, referred to as the assembly protein, is believed to be an essential step in herpesvirus maturation (4, 5). The gene encoding this protein is the 3′ half of a larger overlapping in-frame gene (15, 16), referred to as APNG1 in CMV strain Colburn. We show here that the 5′ half of APNG1 encodes the proteinase required for maturational cleavage of the assembly protein precursor, that the enzyme cleaves between alanine and serine in the sequence VNAS, and that all herpesviruses examined have homologs of both this proteinase and its cleavage site.

Initial assays showed cleavage of the assembly protein precursor in cells cotransfected with APNG1 (in AW4) and the assembly protein precursor gene APNG.5 (in AW1) but not in cells transfected with either gene alone. This indicated that the assembly protein precursor is neither autoproteolytic

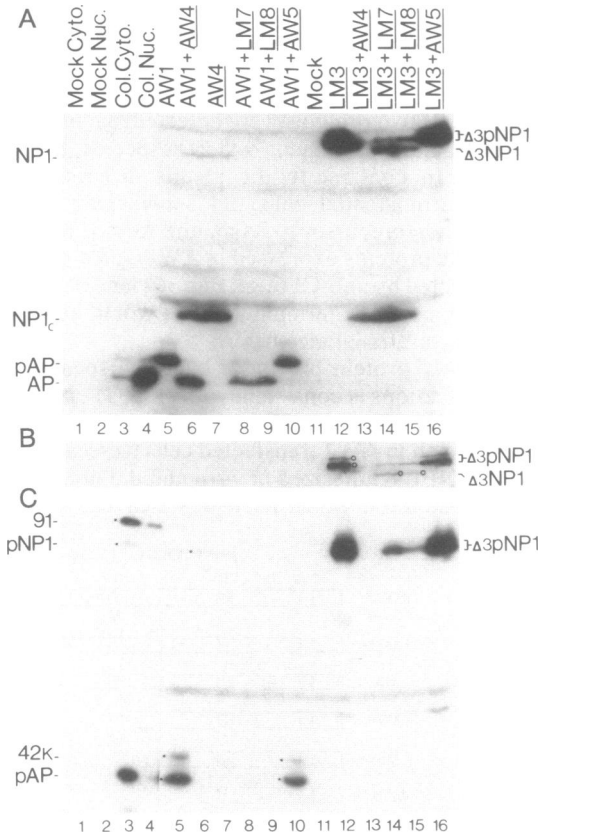

FIG. 3. Proteins in cells transfected with deletion constructs of APNG1. Fluorographic (*A* and *C*) and autoradiographic (*B*) images of immunoblots are shown, comparing proteins in transfected HEK cells that reacted with anti-N1 (*A* and *B*) or with anti-C1 (*C*). Protein designations are as in Fig. 1. The band designated 42K comigrated with a comparatively scarce infected-cell protein that was detected only after prolonged exposure. *A* and *B* are exposures of the same blot (Δ3pNP1 doublet and Δ3NP1 are shown more clearly in *B*); *C* shows a different immunoblot prepared from the same samples in the same order. Other details are explained in Fig. 2.

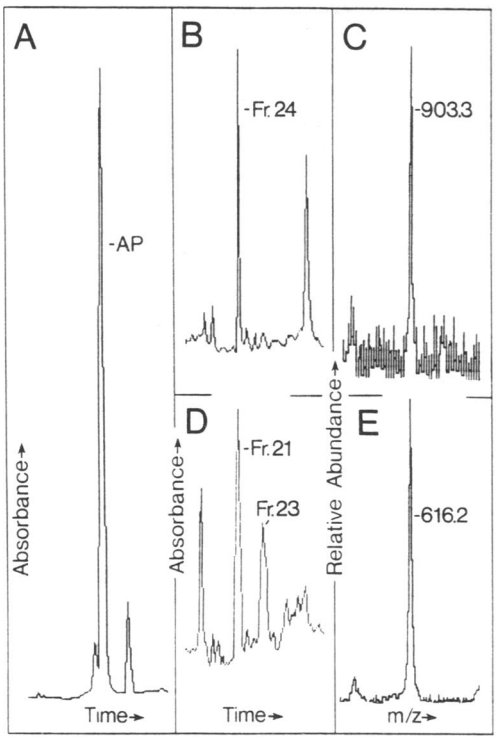

FIG. 4. Identification of assembly protein C-terminal peptides. Relevant portions of recordings obtained from the following procedures are shown. (*A*) HPLC separation of the B-capsid assembly protein, scanned at A_{214}. (*B* and *D*) HPLC separations of peptides resulting from Lys-C and Glu-C treatment, respectively, scanned at A_{214}. (*C* and *E*) PDMS of the Lys-C peptide in fraction 24 (*B*) and the Glu-C peptide in fraction 23 (*D*), respectively, accumulated for 30 min at 16 kV.

nor cleaved by a cellular proteinase and suggested that the proteinase responsible is coded for by APNG1. Insertional mutagenesis underscored the functional importance of the CD1/CD2 region by demonstrating a loss of proteolytic activity when 14 amino acids were inserted between CD1 and CD2 (i.e., LM3, Figs. 2 and 3). Deletional mutagenesis demonstrated that constructs coding for the first 249 (LM8) or 280 (LM7) amino acids of pNP1 retained full proteolytic activity, whereas a shorter one (AW5) coding for only the first 179 amino acids of pNP1 was inactive. Although the proteins of all three constructs contain CD1 and CD2, the shortest one does not contain Ser-195, which may be part of a serine proteinase active site triad (Table 1, and see below).

The proteinase makes two cleavages in its primary translation product, pNP1 (Fig. 5). One is identical to the assembly protein maturational cleavage (Fig. 1, arrow) and occurs because the sequence of the C-terminal half of pNP1 is identical to that of the assembly protein precursor (15). The second cleavage made in pNP1 is at a "release" site (see below) that divides the molecule into a ≈45-kDa C-terminal fragment ($NP1_C$) recognized by anti-N1, and a corresponding N-terminal fragment ($NP1_N$) that was immunologically invisible in these experiments because it is not recognized by either anti-C1 or anti-N1 (Fig. 1). As diagrammed in Fig. 5, this cleavage may be required to free the proteinase ($NP1_N$) from its C-terminal end, thereby activating or otherwise enabling it to cleave the assembly protein precursor. The following evidence supports this model: (*i*) AW4-transfected cells contained much more $NP1_C$ than its putative precursors pNP1 or NP1, indicating that most of the primary translation product had been cleaved. (*ii*) LM3 was unable to produce cleavage of either its own primary translation product or the assembly protein precursor (Fig. 2*C*, lanes 10 and 11), suggesting that the same proteinase catalyzes both cleavages. (*iii*) $NP1_C$ was present in cells cotransfected with LM3 ($NP1_C$ donor) and proteolytically active deletion constructs that lack their own $NP1_C$ coding sequence (i.e., LM7 or LM8), demonstrating that the proteinase can cleave its release site in trans. The possibility that cleavage at this site occurs cotranslationally in cis is mechanistically appealing and is included in Fig. 5. Because pNP1 exhibited no proteolytic activity toward the assembly protein precursor *in vitro*, this model includes a hypothetical activation step (Δ) that may be essential but not provided in cell-free lysates and may account for the electrophoretically slower 91-kDa species (Figs. 2 and 3, lane Col. Cyto.) as an activated form of pNP1. The requirement for such an activation step would distinguish this proteinase from those of positive-strand RNA viruses that typically do exhibit proteolytic activity *in vitro*.

The C-terminal end of the mature assembly protein was determined to be Ala-557. This residue is within a four-amino acid consensus sequence [V/L-X-A-S/V, where X is a polar amino acid, P_3 is usually valine (underlined), and P_1' is usually serine (underlined)] (Table 2). Because this sequence is highly conserved in a region near the C-terminal end of all herpesvirus assembly protein homologs analyzed, it is likely to be an important feature of the substrate recognition/cleavage domain. Considering the possibility that the proteinase may use a similar substrate recognition/cleavage site

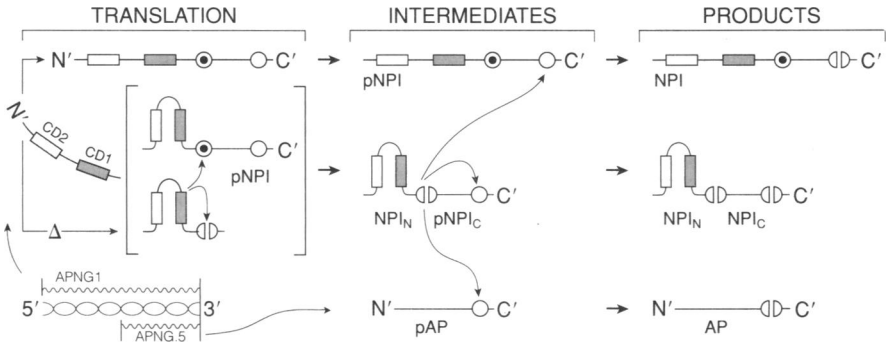

FIG. 5. Hypothetical model of the synthesis and function of the CMV proteinase assemblin. The APNG1 transcript (sawtooth line above double-stranded gene) encodes the proteinase at its 5′ end and can be translated completely to yield pNP1 (85 kDa; reactive with anti-N1 and anti-C1, Figs. 1–3). An undefined event (Δ) modifies the protein so that the proteinase is activated (indicated by changed shape in CD1/CD2 region). Cotranslationally activated proteinase could free itself by cleaving in cis at a specific "release" site (circle with dot, see text), and full-length pNP1 would not be made. Alternatively, cleavage at this site could be made in trans. Both possibilities are indicated (see brackets). Free proteinase ($NP1_N$) cleaves three substrates, all having the same sequence (open circle) at their C-terminal ends: pNP1, $pNP1_C$, and pAP. Cleavage of these substrates eliminates an identical 33-amino acid fragment (C′). This model is based on the results of transfection assays; additional factors may modulate activity of the proteinase in the context of an infected cell.

Table 2. Cleavage site of the CMV assembly protein precursor and its homologs in other herpesviruses

Virus	Recognition/cleavage domain
SCMV	SKSAERGVVNA ↓ SCRVAPP
HCMV	AERAQAGVVNA ↓ SCRLATA
HSV-1	SNAEAGALVNA ↓ SSAAHVD
VZV	HTDTVGQDVNA ↓ VEASSKA
EBV	GHHRGKKLVQA ↓ SASGVAQ
ILTV	NQESARETVDA ↓ SMPKRLK
HHV-6	AASPKPSILNA ↓ S-------

C terminus of Colburn assembly protein was determined by PDMS (Fig. 4); all others are predicted. Abbreviations are as in Table 1. The HHV-6 sequence was kindly provided by G. Lawrence (Cambridge Univ.); others are as in Table 1. Absolutely conserved residues are underlined; arrows denote cleavage sites.

for its own maturational release (Fig. 5), we looked for and found a similar consensus sequence (Y-V/L-Q/K-A-S/N) located between amino acids 234 and 262 in each of the herpesvirus homologs of pNP1 (Table 3). This putative proteinase release site is more highly conserved and differs from the C-terminal cleavage consensus site by having a conserved tyrosine at the P_4 position, a preference for leucine at P_3, and a preference for lysine or glutamine at P_2. Cleavage after the alanine in these sequences would yield the free proteinase, $NP1_N$, and a C-terminal fragment, $NP1_C$.

Although the nature of the catalytic site of this proteinase is presently undefined, the strong conservation of CD1 and CD2 suggests that these domains are essential for activity. Interestingly, 6 of the 15 amino acids absolutely conserved in CD1 and CD2 are also conserved in gp21, the maturational proteinase of bacteriophage T4. It is noteworthy that gp21 contains no cysteine but appears to share other characteristics in common with this herpesvirus proteinase, including the juxtaposition of its gene with that of its substrate, gp22. Further studies are needed to verify that this region is the catalytic site of the herpesvirus proteinase and to determine its mechanism of action.

The discovery of this herpesvirus proteinase is expected to lead to a better understanding of how the assembly protein functions in nucleocapsid formation and maturation and to provide tools and targets for developing inhibitors of herpes group viruses. It also reveals another striking similarity between the eukaryotic herpes group viruses and the prokaryotic bacteriophage T4, which may prove useful in further structural and functional characterizations of both viral enzymes. Data presented here for the CMV strain Colburn proteinase support the conclusions that it is encoded by the 5′ portion of APNG1 and that it is released from its primary translation product as $NP1_N$ by cleavage at a site that yields a complementary C-terminal fragment, $NP1_C$. If this site is YVKA ↓ S in the Colburn protein, as predicted, then the product of our LM8 construct should be the proteinase. The finding that the LM8 protein is proteolytically active is consistent with this prediction. For the sake of simplifying its future discussion, and in recognition of its assembly protein precursor substrate, we refer to this maturational proteinase as assemblin.

Table 3. Putative "release" site of the CMV maturational proteinase and its homologs in other herpesviruses

Virus	Recognition/cleavage domain
SCMV	VTARESYVKA ↓ SVSPAEQETC
HCMV	VTERESYVKA ↓ SVSPEARAIL
HSV-1	GIAGHTYLQA ↓ SEKFKMWGAE
VZV	GIMGHVYLQA ↓ STGYGLARIT
EBV	NIPAESYLKA ↓ SDAPDLQKPD
ILTV	AVYNPKYLQA ↓ NEVITIGIKE

Abbreviations and other notations are as in Table 2.

We thank B. Korant for enthusiastic encouragement of our interest in proteinases, M. A. Mullen for transfection advice, J.-F. Tomb for assistance with computer comparisons, T. Dambaugh, G. Lawrence, and U. Gomples for sharing sequence data prior to publication, and Pamela Talalay for editorial advice. A.R.W. and A.S.W. are students in the Pharmacology and Molecular Science Training Program and were supported by National Institutes of Health Grant GM0726. This research was aided by United States Public Health Service Grants AI13718 and AI22711 to W.G. and National Science Foundation Grant DIR 89-14549 to R.J.C.

1. Gibson, W. & Roizman, B. (1972) *J. Virol.* **10,** 1044–1052.
2. O'Callaghan, D. J. & Randall, C. C. (1976) *Prog. Med. Virol.* **22,** 152–210.
3. Lee, J. Y., Irmiere, A. & Gibson, W. (1988) *Virology* **167,** 87–96.
4. Ladin, B. F., Ihara, S., Hampl, H. & Ben-Porat, T. (1982) *Virology* **116,** 544–561.
5. Preston, V. G., Coates, J. A. V. & Rixon, F. J. (1983) *J. Virol.* **45,** 1056–1064.
6. Sherman, G. & Bachenheimer, S. (1988) *Virology* **163,** 471–480.
7. Gibson, W. (1981) *Virology* **111,** 516–537.
8. Rixon, F., Cross, A. M., Addison, C. & Preston, V. G. (1988) *J. Gen. Virol.* **11,** 2879–2891.
9. Casjens, S. & King, J. (1975) *Annu. Rev. Biochem.* **44,** 555–611.
10. Irmiere, A. & Gibson, W. (1985) *J. Virol.* **56,** 277–283.
11. Gibson, W. & Roizman, B. (1974) *J. Virol.* **13,** 155–165.
12. Braun, C., Roizman, B. & Pereira, L. (1984) *J. Virol.* **49,** 142–153.
13. Gibson, W., Marcy, A. I., Comolli, J. C. & Lee, L. (1990) *J. Virol.* **64,** 1241–1249.
14. Schenk, P., Woods, A. S. & Gibson, W. (1991) *J. Virol.* **65,** 1525–1529.
15. Welch, A. R., McNally, L. M. & Gibson, W. (1991) *J. Virol.* **65,** 4091–4100.
16. Liu, F. & Roizman, B. (1991) *J. Virol.* **65,** 206–212.
17. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
18. Robson, L. & Gibson, W. (1989) *J. Virol.* **63,** 669–676.
19. Mullis, K. B., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, E. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51,** 263–273.
20. Long, E. O., Rosen-Bronson, D. R., Karp, D. R., Malnati, M., Sekaly, R. P. & Jaraquemada, D. (1991) *Hum. Immunol.* **31,** 229–235.
21. Wychowski, C., van der Werf, S., Siffery, O., Crainic, R., Bruneau, P. & Girard, M. (1983) *EMBO J.* **2,** 2019–2020.
22. Chen, C. & Okayama, H. (1987) *Mol. Cell. Biol.* **7,** 2745–2752.
23. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–684.
24. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4350–4354.
25. Woods, A. S. & Cotter, R. J. (1991) *Int. J. Mass Spectrom. Ion Proc.*, in press.
26. Griffin, A. M. (1990) *Nucleic Acids Res.* **18,** 3664.
27. Bazan, J. F. & Fletterick, R. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7872–7876.
28. Gorbalenya, A. E., Donchenko, A. P., Blinov, V. M. & Koonin, E. V. (1989) *FEBS Lett.* **243,** 103–114.
29. Schechter, I. & Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162.