JOURNAL OF VIROLOGY, July 1996, p. 4819–4824 Vol. 70, No. 7
0022-538X/96/$04.00+0

Characterization of a Soluble Stable Human Cytomegalovirus Protease and Inhibition by M-Site Peptide Mimics

ROBERT L. LAFEMINA,[1]* KALPANA BAKSHI,[1] WILLIAM J. LONG,[1] BARNALI PRAMANIK,[1] CHARLOTTE A. VELOSKI,[1] BOHDAN S. WOLANSKI,[1] ALICE I. MARCY,[2] AND DARIA J. HAZUDA[1]

Department of Antiviral Research, Merck Research Laboratories, West Point, Pennsylvania 19486,[1] and Department of Molecular Design and Diversity, Merck Research Laboratories, Rahway, New Jersey 07065[2]

Received 21 December 1995/Accepted 2 April 1996

The human cytomegalovirus (HCMV) protease is a potential target for antiviral chemotherapeutics; however, autoprocessing at internal sites, particularly at positions 143 and 209, hinders the production of large quantities of stable enzyme for either screening or structural studies. Using peptides encompassing the sequence of the natural M-site substrate (P5-P5′, GVVNA/SCRLA), we previously demonstrated that substitution of glycine for valine at the P3 position in the substrate abrogates processing by the recombinant protease in vitro. We now demonstrate that introduction of the V-to-G substitution in the P3 positions of the two major internal processing sites, positions 143 and 209, in the mature HCMV protease renders the enzyme stable to autoprocessing. When expressed in *Escherichia coli*, the doubly substituted protease was produced almost exclusively as the 30-kDa full-length protein. The full-length V141G, V207G (V-to-G changes at positions 141 and 207) protease was purified as a soluble protein by a simple two-step procedure, ammonium sulfate precipitation followed by DEAE ion-exchange chromatography, resulting in 10 to 15 mg of greater than 95% pure enzyme per liter. The stabilized enzyme was characterized kinetically and was indistinguishable from the wild-type recombinant protease, exhibiting K_m and catalytic constant values of 0.578 mM and 13.18/min, respectively, for the maturation site (M-site) peptide substrate, GVVNASCRLARR (underlined residues indicate additions to or substitutions from peptides derived from the wild-type substrate). This enzyme was also used to perform inhibition studies with a series of truncated and/or substituted maturation site peptides. Short nonsubstrate M-site-derived peptides were demonstrated to be competitive inhibitors of cleavage in vitro, and these analyses defined amino acids VVNA, P4 through P1 in the substrate, as the minimal substrate binding and recognition sequence for the HCMV protease.

At least seven major virally encoded proteins involved in herpesvirus capsid assembly have been identified (32, 40, 43). Among these proteins is a major structural protein, the assembly protein (16, 17), the product of the human cytomegalovirus (HCMV) UL80.5 gene (44). In both HCMV and the related herpes simplex virus (HSV), the assembly protein interacts with the major capsid protein and acts as a scaffold in the nucleus around which the capsid assembles (19, 28, 30, 33, 42). The presence of the assembly protein inside the immature B capsid is, however, incompatible with the incorporation of the infectious genome (10, 11, 23). Therefore, the assembly protein must be removed concordantly with or prior to the encapsidation of the viral DNA. The removal of the assembly protein during the process of capsid maturation requires specific cleavage by a virally encoded protease, assemblin (46). The importance of the herpesvirus protease in capsid maturation prior to the encapsidation of the DNA genome has been described for HSV by a number of investigators using temperature-sensitive, deletion, and point mutations in the protease (8, 10, 34). The HCMV protease provides a potential target for the development of an anti-HCMV agent, thus requiring the development of a system to provide sufficient quantities of enzyme for structural and enzymological studies as well as for high-throughput inhibitor screens.

The HCMV protease, UL80, is composed of an N-terminal 256-amino-acid proteolytic domain, a linker region, and a C-terminal structural domain which is colinear with the assembly protein, UL80.5 (44, 46) (see Fig. 1). A similar organization is displayed by other related herpesviruses (1, 2, 7, 9, 12, 13, 27, 29, 29a, 37–39, 41, 44, 45), including HSV with its UL26 protease and UL26.5 (ICP35 family) assembly protein (24–26, 29, 35). Genetic linkage of the protease and the assembly protein substrate within the UL80 open reading frame may facilitate capsid maturation by localizing the protease to the inside of the capsid through interactions specified within the assembly protein structural domain. The assemblin protease processes at two related sites, the maturation or M site (P3-P1′, VNA/S) and the release or R site (P3-P1′, VKA/S) (46). The M site is present in both the UL80 and UL80.5 gene products, whereas the R site is present only within the UL80 protease domain. Processing at the M site liberates the bulk of the assembly protein from the C-terminal residues (11, 35) which associate with the major capsid protein (19, 28, 30, 42). R-site processing releases the mature N-terminal 256-amino-acid proteolytic domain (46). After processing, the assembly protein is lost from the capsid. In contrast, the mature protease is retained inside the capsid, suggesting either a potential second structural or functional role for this enzyme in the virus growth cycle.

At late times during infection, the HCMV protease is also processed at internal sites, predominantly at alanine 143 and to a lesser extent at alanine 209, both of which retain valine in the P3 position (valines 141 and 207, respectively) (3, 18). The internal site at position 143 is conserved in the proteases of simian and murine CMVs (27, 44, 45) but not in the proteases of human herpesvirus 6 (12), human herpesvirus 7 (29a), or the α or γ herpesviruses (1, 2, 9, 13, 29, 37, 38); the internal site at position 209 is unique to HCMV (5, 20). It has been proposed

* Corresponding author. Mailing address: Merck Research Laboratories, Building 16, Room 205, West Point, PA 19486. Phone: (215) 652-7419. Fax: (215) 652-0994. Electronic mail address: robert_lafemina@merck.com.

that the apparent conservation of internal processing sites by the CMV proteases might suggest a mechanism by which to regulate proteolytic activity (3). However, with the demonstration that the two-chain recombinant HCMV protease cleaved at position 143 retains catalytic function (15), the significance of autoprocessing remains unclear. When the mature HCMV protease is expressed in recombinant systems (4, 15, 31, 39), processing occurs at the internal sites, making the large-scale production of enzyme difficult. In addition, the purified wild-type enzyme is unstable and degrades into a two-chain form, which although active (15) is not optimal for biophysical studies.

Previously, we demonstrated that the valine in the P3 position is the most critical residue affecting processing of 10-amino-acid M-site peptide mimics in vitro (37). Substitution of the P3 valine with glycine resulted in a peptide substrate that was not detectably cleaved by the recombinantly expressed HCMV enzyme. In the present study, we have used these peptide cleavage studies as a guide to introduce amino acid substitutions in both of the internal protease processing sites. The resulting protease, V141G, V207G (V-to-G changes at positions 141 and 207), was stably expressed in *Escherichia coli* and was readily purified as an enzymatically active protein from the soluble fraction rather than the particulate fraction as previously described (37). Although others have introduced substitutions in and around the internal position 143 cleavage site, changes around the internal 209 site have not been examined, and in all cases the enzyme was either expressed as a fusion protein and/or purified as an insoluble protein and subsequently refolded (9, 15, 31, 39). The soluble V141G, V207G enzyme was characterized in in vitro cleavage assays and retained kinetic properties with respect to M-site peptide substrates similar to those of the wild-type protease. In addition, the enzyme was used to examine the ability of modified M-site peptides to act as competitive inhibitors of cleavage.

Expression of a stabilized HCMV protease. The organization of the HCMV UL80 and UL80.5 open reading frames (44, 46) is depicted schematically in Fig. 1. The locations of the release or R sites (UL80, position 256) and maturation or M sites (UL80, position 643, and UL80.5, position 308), as well as the internal processing or I sites (UL80, positions 143 and 209), are indicated. The location of the peptide (amino acids 4 to 15) used to generate the rabbit 177 antiserum for these analyses is also noted. When the wild-type mature HCMV protease (amino acids 1 to 256) is expressed in *E. coli*, this antiserum reacts with the full-length mature protease which migrates with an apparent molecular mass of 30 kDa, as well as with a major 16-kDa N-terminal cleavage product, the result of internal processing at position 143. As shown in Fig. 2, Western (immunoblot) analysis demonstrated that 50% of the recombinant wild-type enzyme expressed in *E. coli* was processed internally at this site (Fig. 2, lane 2). Since internal processing complicates the production and purification of large quantities of enzyme for structural studies and inhibitor screening, we and others have taken the approach that amino acid substitutions which alter the residues surrounding the major internal cleavage site(s) may limit the extent of autoprocessing and facilitate the production of a stable enzyme in vitro (9, 15, 21, 31). Substitutions at single amino acids surrounding the position 143 site have been shown to reduce autoprocessing in the recombinantly expressed enzyme purified in an insoluble form or as a fusion protein from *E. coli*. However, since internal processing occurs predominantly at position 143 and to a lesser extent independently at position 209 during a virus infection (3, 18), mutation of both sites within the mature protease may be necessary for producing a recombinant enzyme with significantly enhanced stability.

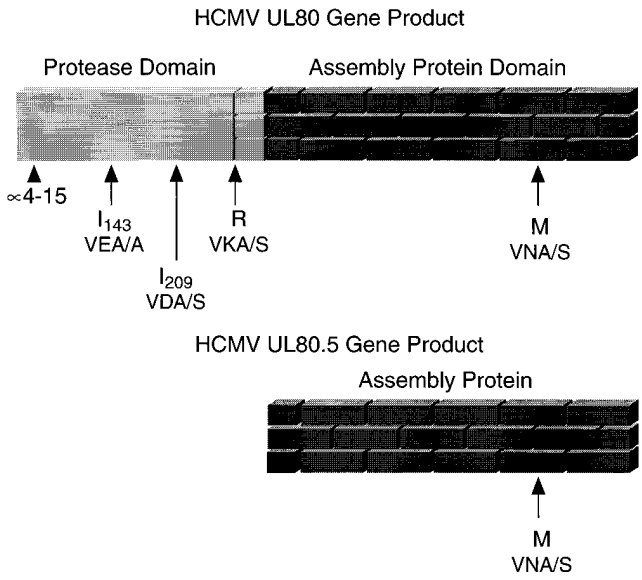

FIG. 1. Organization of the HCMV UL80 and UL80.5 open reading frames. The UL80 gene product contains a protease domain (solid block) and a domain colinear with the UL80.5 structural assembly protein domain (dark bricks), tethered by a linker domain (light bricks). Antiserum was raised against an N-terminal epitope comprising amino acids 4 to 15 (α4-15). The P3-P1′ residues for the I-143 and I-209 sites, as well as the R site at position 256 and the M site at position 643 of the UL80 gene product, are indicated by arrows. The M site is also depicted by the arrow at position 308 of the UL80.5 assembly protein.

To produce a source of the intact HCMV protease, we therefore introduced amino acid substitutions around both of the potential internal cleavage sites in an attempt to prevent autoprocessing. Both the wild-type and the substituted HCMV 256-amino-acid proteases were expressed in *E. coli* by using the T7 system (36), and the lysates were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting. As shown by Western blot analysis, substitution of the P1 alanine at position 143 with serine reduced processing at this site (Fig. 2, lane 3) and resulted in an increase in the amount of the 30-kDa protein relative to the 16-kDa internal cleavage product in comparison to the wild-type protein (Fig. 2, compare lanes 2 and 3). In contrast, substitution of a glycine at this position had only a minor effect (Fig. 2, lane 4). Introduction of glycine for the P3 valine at position 141 essentially abolished cleavage at the 143 site; however, several smaller immunoreactive polypeptides with apparent molecular masses of 21, 25, and 27 kDa, respectively, were still detected (Fig. 2, lane 5). Although these species would be consistent with processing at the alanines located at positions 189 (P3-P1′, LTAA), 209 (P3-P1′, VDAS), and 228 (P3-P1′, VDAL), processing at these sites may be autocatalytic or may be the result of cleavage by endogenous *E. coli* proteases.

Processing at the 209 position has been observed to occur in CMV-infected cells (18), while autoprocessing at positions 189 and 228 has not been described, suggesting that only the 209 site was the result of autoprocessing in *E. coli*. The 209 site was therefore altered in the V141G background, again changing the P3 residue at position 207 from valine to glycine. This additional substitution resulted in a significant increase in the amount of polypeptide produced with an apparent molecular weight consistent with the size expected for the full-length protein (Fig. 2, lane 6). In comparison with either the wild-type enzyme or the singly substituted V141G protein, only minor

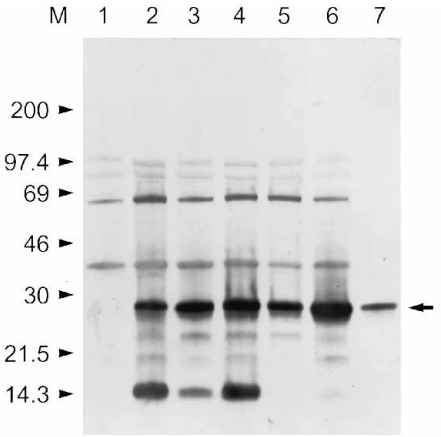

FIG. 2. Western blot analysis of the mature wild-type and substituted HCMV proteases expressed in *E. coli*. HCMV AD169 DNA was prepared as described elsewhere (37). Cloning of the wild-type mature 256-amino-acid protease from the UL80 open reading frame of HCMV AD169 (5) by PCR (14) and ligation into the T7 expression vector pET3c (36) have been described previously (37). Amino acid substitutions were introduced around the internal cleavage site at amino acid 143 by standard PCR mutagenesis techniques (14). The amino acid substitutions and the oligonucleotides used for the mutagenesis, with the mutagenic nucleotides underlined, are as follows: A143S, 5′GACGTGGAGTCCGC GACGTCG; A143G, 5′GACGTGGAGGGCGCGACGTCG; V141G, 5′GACG GGGAGGCCGCGACGTCG. A second substitution, V207G, was introduced into the V141G protease by PCR mutagenesis by using the oligonucleotide 5′GCAGCACTGCTGGCGACGCG. As with the wild-type enzyme, all enzymes were cloned into pET3c for expression in *E. coli* BL21(DE3). Except where noted, *E. coli* lysates representing 0.1 ml of bacteria at an optical density at 600 nm equal to 1.0 were electrophoresed through 10% Tricine gels (Novex, Encinitas, Calif.), transferred to polyvinylidene difluoride, and analyzed by using the 177 antiserum and goat anti-rabbit immunoglobulin G-coupled alkaline phosphatase and colorimetric development with Nitro Blue Tetrazolium and 5-bromo-4-chloro-3-indolylphosphate toluidinium (Bio-Rad, Richmond, Calif.). M, molecular size markers (in kilodaltons); lane 1, *E. coli* lysate; lane 2, *E. coli* expressing the wild-type mature 256-amino-acid protease domain; lane 3, *E. coli* expressing A143S protease; lane 4, *E. coli* expressing A143G protease; lane 5, *E. coli* expressing V141G protease; lane 6, *E. coli* expressing V141G, V207G protease (0.067 ml); lane 7, 1 μg of purified V141G, V207G protease. Relative mobility of the mature 256-amino-acid protease is indicated by the arrow.

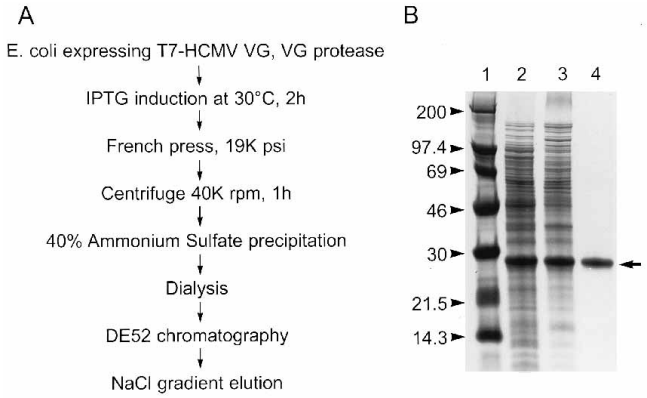

FIG. 3. Purification of the soluble HCMV V141G, V207G protease. (A) Flow chart of protease purification. For large-scale production of the V141G, V207G HCMV protease (VG, VG protease), cultures were grown in M9 medium supplemented with 0.1% Casamino Acids to an A_{600} of 1.0 at 37°C. For induction, 0.1 mM IPTG was added and the cultures were incubated for 2 h at 30°C. Under these conditions, 60 to 70% of the protease partitions in the soluble fraction of *E. coli* upon lysis. For purification, packed cells from 6-liter cultures were resuspended in 40 ml of buffer A (50 mM Tris-HCl, pH 7.8; 25 mM NaCl; 1 mM dithiothreitol; 1 mM EDTA; 10% glycerol) with 1 μg each of leupeptin, aprotinin, pepstatin, and bestatin per ml. Cells were lysed by a single passage through a French press at 19,000 lb/in^2. The debris was clarified by centrifugation (40,000 rpm, 1 h [Ti 70; Beckman, Fullerton, Calif.]), and the S100 soluble fraction was made 40% ammonium sulfate by the slow addition of a saturated solution. After incubation on ice for 30 min, the ammonium sulfate precipitate was pelleted by centrifugation (12,000 rpm, 15 min). The pellet was resuspended in and then dialyzed for 4 h against buffer B (50 mM Tris-HCl, pH 7.8; 10% glycerol; 5 mM 2-mercaptoethanol). The dialyzed protease was loaded onto an 8-ml column of DE52 (Whatman) previously equilibrated in buffer B. The column was washed with 10 volumes of buffer B plus 25 mM NaCl, and then the protease was eluted by using a linear NaCl gradient (25 to 200 mM) in buffer B. Protease fractions were identified following fractionation, by SDS-PAGE with Coomassie staining and by Western analysis. Peak fractions were pooled and concentrated to 10 mg/ml by using a Centriprep concentrator (Amicon, Beverly, Mass.). Purity was assessed by SDS-PAGE and isoelectric focusing. (B) Protein samples were subjected to SDS-PAGE, and proteins were stained with Coomassie blue. Lane 1, molecular size markers with sizes (in kilodaltons) indicated to the left; lane 2, total soluble fraction of lysed *E. coli* expressing the V141G, V207G protease; lane 3, pellet following 40% ammonium sulfate precipitation; lane 4, 5 μg of purified soluble HCMV V141G, V207G protease. The arrow indicates the relative mobility of the mature 256-amino-acid protease.

immunoreactive species of lower molecular weights were detectable for the doubly substituted enzyme (Fig. 2, compare lanes 2 and 5 with lane 6). Since Fig. 2 depicts equivalent loadings of *E. coli* lysate containing the wild-type, A143S, A143G, and V141G enzymes, while the V141G, V207G lysate was reduced by one-third, we conclude that the introduction of the V-to-G substitution at the second internal processing site resulted in a marked increase in the amount of full-length 30-kDa enzyme. We estimate that the HCMV V141G, V207G protease is expressed to a level of approximately 15 to 30 mg/liter in *E. coli.*

Purification of the stabilized V141G, V207G protease. Functional wild-type protease has previously been recovered and purified from the particulate fraction of lysed *E. coli* (37). However, purification could potentially be improved by purifying the enzyme from the soluble fraction, eliminating denaturation and the requirement for renaturation and refolding via dialysis. Therefore, in an attempt to increase the fraction of soluble enzyme expressed in *E. coli*, we examined the induction conditions for the V141G, V207G protease with regard to time, temperature, and concentration of isopropyl-β-D-thiogalactopyranoside (IPTG). As a result of these studies, we observed that induction for 2 h at 30°C with 0.1 mM IPTG resulted in greater than 60% of the protease partitioning into the soluble fraction upon lysis (20).

Large-scale cultures of *E. coli* expressing the V141G, V207G protease were grown, induced, and lysed under these modified conditions. With the low-temperature induction, the V141G, V207G protease was produced at high levels in the soluble fraction (Fig. 3B, lane 2). As outlined in Fig. 3A, a simple two-step procedure involving precipitation with 40% ammonium sulfate (Fig. 3B, lane 3) followed by DE52 anion-exchange chromatography (Fig. 3B, lane 4) resulted in HCMV protease that was purified to greater than 95% homogeneity and a yield of greater than 10 mg/liter. The V141G, V207G enzyme was stable both structurally and functionally when subjected to repeated freeze-thaw cycles and when incubated for 5 days at room temperature and at least 7 days at 4°C (22), similar to what others have observed for position 143 singly substituted protease (31). Thus, the V141G, V207G protease is suitable for both screening and structural studies.

Recombinant HCMV proteases containing amino acid substitutions at internal processing sites are indistinguishable from the wild-type enzyme in vitro. Although the V141G, V207G protease was expressed to high levels, was easily purified in quantities sufficient for structural studies, and exhibited desirable characteristics with regard to stability, it was essential to compare the enzymatic activity of the doubly substituted protease with that of the authentic, wild-type enzyme. In earlier studies using the P5-P5′ M-site peptide, GVVNASCRLA,

TABLE 1. Kinetic analysis of HCMV protease variants

Enzyme	K_m (mM) (± SD)	k_{cat} (min^{-1}) (± SD)
Wild type	0.515 ± 0.05[a]	12.37 ± 0.87
A143S	0.488 ± 0.068[b]	10.89 ± 0.036
A143G	0.633 ± 0.065[b]	ND[c]
V141G	0.414 ± 0.035[b]	14.14 ± 0.13
V141G, V207G		
Refolded	0.590 ± 0.043[b]	14.50 ± 0.08
Soluble	0.578 ± 0.053[b]	13.18 ± 0.06

[a] Substrate, GVVNA/SCRLA.
[b] Substrate, GVVNA/SCRLARR.
[c] ND, not determined.

the wild-type recombinant enzyme purified from the particulate fraction was shown to have a K_m of 0.515 mM and a catalytic constant (k_{cat}) of 12.37/min in in vitro peptide cleavage assays (37).

In the present study, we used an M-site mimic, GVVNAS CRLARR (P5-P5′ RR) (underlined residues in peptide substrates indicate additions to or substitutions from peptides derived from the wild-type substrate), as described below. K_m analysis for the singly and doubly substituted I-site protease mutants is presented in Table 1. The A143S, A143G, and V141G enzymes were prepared from the particulate fraction of lysed *E. coli* as previously described. For comparative purposes the V141G, V207G protease was prepared both from the particulate fraction, as described previously (37), and from the soluble fraction, as described above. As shown in Table 1, with the M-site mimic peptide, the K_m values exhibited by all four of the substituted enzymes were indistinguishable. The values of 0.578 and 0.590 mM obtained in these studies for the doubly substituted enzyme are also comparable to the previous values determined for the wild-type protease by using the P5-P5′ substrate (37). Others have observed similar K_m values for the enzyme in in vitro cleavage assays, despite minor differences in the lengths of the respective peptide substrates used by various investigators (4, 15, 31, 39).

The k_{cat} values for these enzymes were also determined and were also seen to be equivalent for the five enzymes tested: wild type; A143S; A143G; V141G; and V141G, V207G. The values of 14.5/min and 13.18/min for the refolded and soluble V141G, V207G enzymes were nearly identical to those previously published for the wild-type enzyme (4, 37). Therefore, mutation of the internal processing sites had no apparent effect on the activity of the recombinant HCMV protease in peptide cleavage assays in vitro. The observation that equivalent kinetic parameters are observed for wild-type and singly and doubly substituted proteases suggests that the substitutions around internal sites 143 and 209 do not affect the active site. Elucidation of the relationship of these amino acids relative to the active-site catalytic residues histidine 63 (45), serine 132 (39, 45), and glutamic acid 122 (6) awaits the solution of the three-dimensional structure of the enzyme. The availability of large amounts of active, recombinant enzyme should facilitate these studies.

It should be noted that the published k_{cat} values for HCMV proteases expressed as fusion proteins are somewhat lower (31, 39), and this decreased activity may reflect on the importance of the native N terminus as defined by previous deletion mutagenesis studies (3). However, the ease of the purification and the yields which we have achieved by expressing the stabilized protease in a soluble form make purification by means of affinity tags unnecessary.

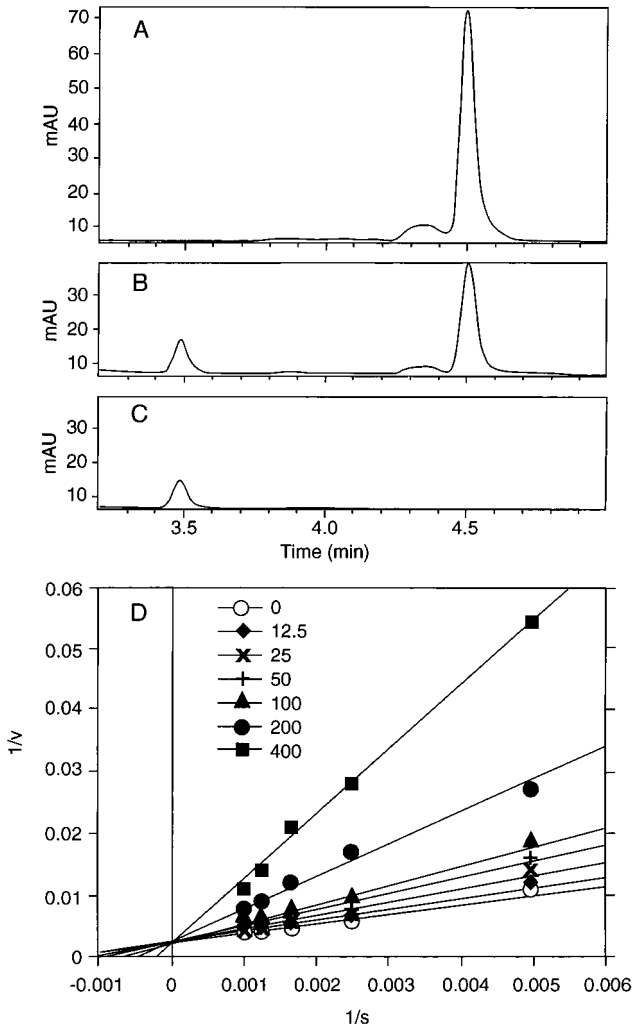

FIG. 4. Analysis of inhibition by the peptide GVVNAACRLA. The in vitro peptide cleavage assays were performed essentially as described previously (37) by using 50 nM protease in 100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 7.8)–25 mM NaCl–1 mM EDTA for 30 min at ambient temperature. For competitive peptide inhibition studies, the substrate GVVNA SCRLARR was titrated from 0.2 to 1 mM. The peptide inhibitor GVVNAA CRLA was dissolved in dimethyl sulfoxide (DMSO) and titrated by twofold dilution from 400 to 12.5 μM. All reactions were performed in a final concentration of 10% DMSO. This DMSO concentration was shown to have no effect on protease activity in vitro. HPLC analysis using a Western Nova-Pak C_{18} column (Millipore, Milford, Mass.) was performed as described previously (37). Shown are elution profiles from 3.2 to 5 min. The P side cleavage product, GVVNA, elutes in the solvent front, and unprocessed substrate elutes at 14 min. (A) With 1 mM substrate, the peak at 4.5 min represents the P′ cleavage product SCRLARR. (B) With 1 mM substrate and 200 μM competitor, the peak at 4.5 min represents the P′ cleavage product SCRLARR and the competitor peptide P′ cleavage product, ACRLA, has a retention time of 3.5 min. (C) With 200 μM competitor alone, the peak at 3.5 min represents P′ cleavage product ACRLA. mAU, milliunits of absorbance at 280 nm. (D) Lineweaver-Burk analysis using the substrate GVVNASCRLARR and 0 to 400 μM competitor peptide GV VNAACRLA as indicated by the various symbols.

Short M-site peptides which are not processed can act as competitive inhibitors. As described above, the M-site peptide mimic, P5-P5′ RR, was processed by the soluble V141G, V207G protease with kinetics equivalent to those observed for the wild-type enzyme using the authentic P5-P5′ M-site peptide. The presence of the two additional arginine residues in the M-site mimic, however, allows the P′ side cleavage product, SCRLARR, to be separated by high-performance liquid chro-

TABLE 2. Inhibition of HCMV protease by M-site peptides[a]

Sequence	K_m (mM)	K_i (mM)
GVVVNA/SCRLA	0.515	0.225
GVVNA/ACRLA	0.122	0.072
GGGNA/SCRLA	NC[b]	*[c]
aVVNA/SCRL	0.225	0.390
aVVNA/SCR	1.45	0.628
aVVNA/SC	ND[d]	1.52
aVVNA/S	NC[b]	2.17
aVVNA/A	NC[b]	2.94
aVVNA	NC[b]	1.36
VVNA/SCRL	1.07	0.699
VNA/SCRL	7.5	*
NA/SCRL	NC[b]	*
aVVNA/SCRL	0.225	0.390
aVNA/SCRL	2.16	*
aNA/SCRL	NC[b]	*

[a] Peptide competition studies were performed as described in the legend to Fig. 4.

[b] NC, no cleavage detected. K_m values could not be calculated in the absence of cleavage.

[c] *, inhibition observed only at >10 mM.

[d] ND, value could not be determined because of solubility problems at high concentrations.

matography (HPLC) from the M-site P′ product, ACRLA, as shown in Fig. 4. HPLC analyses of proteolysis reaction mixtures containing either the GVVNASCRLARR substrate alone, the substrate plus GVVNAACRLA (10-amino-acid peptide with a P1′A substitution), or the GVVNAACRLA peptide alone are presented in Fig. 4A, B, and C, respectively. The unprocessed substrate, GVVNASCRLARR, has a retention time of 14 min. In the presence of the protease, specific processing of the GVVNASCRLARR substrate is indicated by the generation of a new peak at 4.5 min with the retention time expected for the cleavage product SCRLARR (Fig. 4A). Processing of the substrate GVVNAACRLA generates a peak at 3.5 min, indicative of the product ACRLA (Fig. 4C). The differential separation of the two products permits the P5-P5′RR M-site mimic peptide to be used as the substrate in inhibition studies wherein M-site-derived peptides are used as potential competitors. As shown in Fig. 4B, when 1 mM GVV NAACRLA is added to the reaction mixture containing the GVVNASCRLARR substrate, the peak at 4.5 min characteristic of the SCRLARR cleavage product is significantly reduced. Amounts of substrate and competitor were varied relative to each other, with the amount of enzyme remaining constant, and the SCRLARR cleavage product was quantified; the analysis of these results is shown in Fig. 4D. The P1′ A-substituted M-site peptide competitively inhibited cleavage of the GVVNASCRLARR substrate by the HCMV protease with a K_i of 0.072 mM.

These analyses can be used to examine inhibition by M-site peptides independently of cleavage and can, therefore, potentially be used to distinguish the requirements for substrate binding from the requirements for catalysis. We examined a series of substituted and/or truncated M-site peptides using this approach. The K_m values are compared with the K_i values for these peptides used as inhibitors of M-site mimic P5-P5′RR peptide processing. The inclusion of an acetyl group on the amino terminus (acetylated peptides are indicated by an amino-terminal "a") was observed to decrease both the K_m and K_i values for all peptides analyzed. Therefore, for truncations from P5 to P3 the data for both the acetylated and nonacetylated peptides are presented.

As shown in Table 2, truncations in P4 and P3 resulted in both an increase in the K_m for cleavage and a loss of competitive inhibition. Comparison of either peptides VVNA/SCRL, VNA/SCRL, and NA/SCRL or aVVNA/SCRL, aVNA/SCRL, and aNA/SCRL showed that removal of the P4 residue resulted in a dramatic increase in the K_m, and poor inhibitory activity and elimination of the P3 residue essentially abolished processing as well as binding. This effect was not simply due to deletion of the amino acid in these positions. The 10-amino-acid peptide GGGNA/SCRLA, wherein the P3 and P4 valine resides were replaced with glycine, also neither was processed nor inhibited cleavage of the authentic P5-P5′RR substrate, suggesting that the P3 and P4 valines in the substrate are essential for binding. In contrast, with acetylated peptides which retain P4 to P1, although truncating the P3′ residue substantially reduced cleavage and elimination of the P2′ residue abolished cleavage, neither of these alterations significantly affected the ability of these peptides to competitively inhibit the reaction. In fact, peptides consisting exclusively of aVVNA were essentially indistinguishable from aVVNA/SCR in their ability to inhibit cleavage and were only 5- to 10-fold less effective than the P5-P5′ peptide. These studies demonstrate that the four amino acid residues specified by P4 to P1 in the substrate are probably sufficient for specific binding to the HCMV protease active site.

The amino acid requirements for substrate cleavage have been investigated in some detail previously, by using both HCMV M- and R-site peptide analogs (37). Efficient cleavage of peptide substrates by the HCMV protease in vitro requires amino acid residues from P4 to P4′, and substitution of any amino acid within this sequence has been shown to adversely affect proteolysis. In agreement with these data, it was also shown by Stevens et al. that R P5-P4′ and M P4-P5′ are the minimal peptide substrates spanning the release and maturation sites, respectively (39). In this paper, we now demonstrate that although larger peptides are required for efficient cleavage, smaller nonsubstrate peptides can be recognized and bind to the HCMV protease with affinities which are relatively good compared with those of the peptides which are efficient substrates for this enzyme. In particular, in competition studies the acetylated P4-P1 M-site peptide bound to the enzyme only four- or fivefold less well than did the acetylated P4-P4′ substrate, suggesting that recognition of the substrate may be determined predominantly by residues amino terminal to the cleavage site. Although these peptides are weak inhibitors of the enzyme in vitro, we note that given the best K_i value of 0.072 mM for peptide GVVNA/ACRLA and the difficulties of stability and cell penetration by peptides, the lack of antiviral activity exhibited by these peptides when used in cell culture infectivity studies is to be anticipated. However, these in vitro data would suggest that low-molecular-weight nonpeptide inhibitors of the protease based on the P4-P1 peptide can be designed. The identification of the P4-P1 peptide, VVNA, as a minimal substrate recognition sequence for the HCMV protease may therefore serve to facilitate any such future modeling efforts in this regard.

In conclusion, we have described the construction of a doubly substituted, stabilized HCMV protease; purified this enzyme as a soluble protein from *E. coli*; and demonstrated that this engineered enzyme is essentially indistinguishable from the wild-type enzyme in in vitro cleavage assays. In addition, we have used this enzyme to study the requirements for substrate binding and have identified a 4-amino-acid minimal element, VVNA, specified by P4 to P1 in the substrate which may serve as a core structure against which nonpeptide inhibitors may be modelled.

We thank Emilio Emini for his continued support of these studies. We acknowledge Vinod Sardana and Lawrence Kuo for helpful dis-

cussions and Kathy LeGrow for technical assistance. We also thank Vinod Sardana and Jill Wolfgang for initiating the protease stability studies. We are grateful to Jennifer Rotonda and Joseph Becker for helpful discussions and for performing the isoelectric focusing and Mohinder Sardana for amino acid analysis. Finally, we thank Dolores Wilson for manuscript preparation.

REFERENCES

1. **Albrecht, J.-C., J. Nicholas, D. Biller, K. R. Cameron, B. Biesinger, C. Newman, S. Wittmann, M. A. Craxton, H. Coleman, B. Fleckenstein, and R. W. Honess.** 1992. Primary structure of the herpesvirus saimiri genome. J. Virol. **66:**5047–5058.
2. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, T. P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310:**207–211.
3. **Baum, E. Z., G. A. Bebernitz, J. D. Hulmes, V. P. Muzithras, T. R. Jones, and Y. Gluzman.** 1993. Expression and analysis of the human cytomegalovirus UL80-encoded protease: identification of autoproteolytic sites. J. Virol. **67:** 497–506.
4. **Burck, P. J., D. H. Berg, T. P. Luk, L. M. Sassmannshausen, M. Wakulchik, D. P. Smith, H. M. Hsiung, G. W. Becker, W. Gibson, and E. C. Villarreal.** 1994. Human cytomegalovirus maturational proteinase: expression in *Escherichia coli*, purification, and enzymatic characterization by using peptide substrate mimics of natural cleavage sites. J. Virol. **68:**2937–2946.
5. **Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchison III, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. M. 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.
6. **Cox, G. A., M. Wakulchik, L. M. Sassmannshausen, W. Gibson, and E. C. Villarreal.** 1995. Human cytomegalovirus proteinase: candidate glutamic acid identified as third member of putative active-site triad. J. Virol. **69:** 4524–4528.
7. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella-zoster virus. J. Gen. Virol. **67:**1759–1816.
8. **Desai, P., S. C. Watkins, and S. Person.** 1994. The size and symmetry of B capsids of herpes simplex virus type 1 are determined by the gene products of the UL26 open reading frame. J. Virol. **68:**5365–5374.
9. **Donaghy, G., and R. Jupp.** 1995. Characterization of the Epstein-Barr virus proteinase and comparison with the human cytomegalovirus proteinase. J. Virol. **69:**1265–1270.
10. **Gao, M., L. Matusick-Kumar, W. Hurlburt, S. F. DiTusa, W. W. Newcomb, J. C. Brown, P. J. McCann III, I. Deckman, and R. J. Colonno.** 1994. The protease of herpes simplex virus type 1 is essential for functional capsid formation and viral growth. J. Virol. **68:**3702–3712.
11. **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.
12. **Gompels, U. A., J. Nicholas, G. Lawrence, M. Jones, B. J. Thomson, M. E. D. Martin, S. Efstathiou, M. Craxton, and H. A. Macaulay.** 1995. The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. Virology **209:**29–51.
13. **Griffin, A. M.** 1990. The complete sequence of the capsid p40 gene from infectious laryngotracheitis virus. Nucleic Acids Res. **18:**3664.
14. **Higuchi, R.** 1990. Recombinant PCR, p. 177–183. *In* M. A. Innis, D. H. Gelfand, J. T. Sninsky, and T. J. White (ed.), PCR protocols. Academic Press, San Diego, Calif.
15. **Holwerda, B. C., A. J. Wittwer, K. L. Duffin, C. Smith, M. V. Toth, L. S. Carr, R. C. Wiegand, and M. L. Bryant.** 1994. Activity of two-chain recombinant human cytomegalovirus protease. J. Biol. Chem. **269:**25911–25915.
16. **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.
17. **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.
18. **Jones, T. R., L. Sun, G. A. Bebernitz, V. P. Muzithras, H.-J. Kim, S. H. Johnston, and E. Z. Baum.** 1994. Proteolytic activity of human cytomegalovirus UL80 protease cleavage site mutants. J. Virol. **68:**3742–3752.
19. **Kennard, J., F. J. Rixon, I. M. McDougall, J. D. Tatman, and V. G. Preston.** 1995. The 25 amino-acid-residues at the carboxy-terminus of the herpes-simplex virus type-1 UL26.5 protein are required for the formation of the capsid shell around the scaffold. J. Gen. Virol. **76:**1611–1621.
20. **LaFemina, R.** Unpublished data.
21. **LaFemina, R. L., B. Pramanik, C. Veloski, W. Long, K. LeGrow, B. Wolanski, K. Bakshi, A. I. Marcy, and D. J. Hazuda.** 1995. Inhibition of the human cytomegalovirus UL80 mature protease by small peptides. Scand. J. Infect. Dis. **99**(Suppl.)**:**112.
22. **LaFemina, R. L., V. Sardana, and J. Wolfgang.** Unpublished data.
23. **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.
24. **Liu, F., and B. Roizman.** 1991. The herpes simplex virus type 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. J. Virol. **65:**5149–5156.
25. **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.
26. **Liu, F., and B. Roizman.** 1993. Characterization of the protease and other products of amino-terminus-proximal cleavage of the herpes simplex virus type 1 UL26 protein. J. Virol. **67:**1300–1309.
27. **Loutsch, J. M., N. J. Galvin, M. L. Bryant, and B. C. Holwerda.** 1994. Cloning and sequence analysis of murine cytomegalovirus protease and capsid assembly protein genes. Biochem. Biophys. Res. Commun. **203:**472–478.
28. **Matusick-Kumar, L., W. W. Newcomb, J. C. Brown, P. J. McCann III, W. Hurlburt, S. P. Weinheimer, and M. Gao.** 1995. The C-terminal 25 amino acids of the protease and its substrate ICP35 of herpes simplex virus type 1 are involved in the formation of sealed capsids. J. Virol. **69:**4347–4356.
29. **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.

29a. **Nicholas, J.** 1996. Determination and analysis of the complete nucleotide sequence of human herpesvirus-7. GenBank accession number U43400.

30. **Nicholson, P., C. Addison, A. M. Cross, J. Kennard, V. G. Preston, and F. J. Rixon.** 1994. Localization of the herpes-simplex virus type-1 major capsid protein VP5 to the cell-nucleus requires the abundant scaffolding protein VP22A. J. Gen. Virol. **75:**1091–1099.
31. **O'Boyle, D. R., II, K. Wager-Smith, J. T. Stevens III, and S. P. Weinheimer.** 1995. The effect of internal autocleavage on kinetic properties of the human cytomegalovirus protease catalytic domain. J. Biol. Chem. **270:**4753–4758.
32. **Person, S., S. Laquerre, D. Prashant, and J. Hempel.** 1993. Herpes simplex virus type 1 capsid protein, VP21, originates within the UL26 open reading frames. J. Gen. Virol. **74:**2269–2273.
33. **Preston, V. G., M. F. Alkobaisi, I. M. McDougall, and F. J. Rixon.** 1994. The herpes-simplex virus gene UL26 proteinase in the presence of the UL26.5 gene-product promotes the formation of scaffold-like structures. J. Gen. Virol. **75:**2355–2366.
34. **Preston, V. G., J. A. V. Coates, and F. J. Rixon.** 1983. Identification and characterization of a herpes simplex virus gene product required for incapsidation of virus DNA. J. Virol. **45:**1056–1064.
35. **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.
36. **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 DNAs by T7 RNA polymerase. Gene **56:**125–135.
37. **Sardana, V. V., J. A. Wolfgang, C. A. Veloski, W. J. Long, K. LeGrow, B. Wolanski, E. A. Emini, and R. L. LaFemina.** 1994. Peptide substrate cleavage specificity of the human cytomegalovirus protease. J. Biol. Chem. **269:** 14337–14340.
38. **Steffy, K. R., S. Schoen, and C. M. Chen.** 1995. Nucleotide sequence of the herpes-simplex virus type-2 gene encoding the protease and capsid protein ICP35. J. Gen. Virol. **76:**1069–1072.
39. **Stevens, J. T., C. Mapelli, J. Tsao, M. Hall, D. O'Boyle II, S. P. Weinheimer, and C. L. Diianni.** 1994. *In vitro* proteolytic activity and active-site identification of the human cytomegalovirus protease. Eur. J. Biochem. **226:**361–367.
40. **Tatman, J. D., V. G. Preston, P. Nicholson, R. M. Elliott, and F. J. Rixon.** 1994. Assembly of herpes-simplex virus type-1 capsids using a panel of recombinant baculoviruses. J. Gen. Virol. **75:**1101–1113.
41. **Telford, E. A. R., M. S. Watson, K. McBride, and A. J. Davison.** 1992. The DNA sequence of equine herpesvirus-1. Virology **189:**304–316.
42. **Thomsen, D. R., W. W. Newcomb, J. C. Brown, and F. L. Homa.** 1995. Assembly of the herpes simplex virus capsid: requirement for the carboxyl-terminal twenty-five amino acids of the proteins encoded by the UL26 and UL26.5 genes. J. Virol. **69:**3690–3703.
43. **Thomsen, D. R., L. L. Roof, and F. L. Homa.** 1994. Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. J. Virol. **68:**2442–2457.
44. **Welch, A. R., L. M. McNally, and W. Gibson.** 1991. Cytomegalovirus assembly protein nested gene family: four 3′-coterminal transcripts encode four in-frame, overlapping proteins. J. Virol. **65:**4091–4100.
45. **Welch, A. R., L. M. McNally, M. R. T. Hall, and W. Gibson.** 1993. Herpesvirus proteinase: site-directed mutagenesis used to study maturational, release, and inactivation cleavage sites of precursor and to identify a possible catalytic site serine and histidine. J. Virol. **67:**7360–7372.
46. **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.