PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/57, 9/50, C07K 5/10
C12Q 1/37, C07K 5/02, 13/00
C07K 7/02, 7/06

A1

(11) International Publication Number: **WO 93/01291**

(43) International Publication Date: 21 January 1993 (21.01.93)

(21) International Application Number: PCT/US92/05513

(22) International Filing Date: 2 July 1992 (02.07.92)

(30) Priority data:

725,308	5 July 1991 (05.07.91)	US
798,776	27 November 1991 (27.11.91)	US

(71) Applicant: THE JOHNS HOPKINS UNIVERSITY [US/US]; 720 Rutland Road, Baltimore, MD 21205 (US).

(72) Inventors: GIBSON, D., Wade ; 6414 Pinehurst Road, Baltimore, MD 21212 (US). WELCH, Anthony, R. ; 205 E. Joppa Road, Baltimore, MD 21204 (US).

(74) Agents: FOULKE, Cynthia, Lee et al.; Banner, Birch, McKie & Beckett, 1001 G Street, N.W., 11th Floor, Washington, DC 20001 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE).

Published
With international search report.

(54) Title: HERPES VIRUS PROTEINASE AND METHOD OF ASSAYING

(57) Abstract

A herpes virus proteinase has been found to be encoded by a member of a family of four nested genes in simian cytomegalovirus. Another member of the nested genes encodes the assembly protein precursor, which is a substrate for the proteinase. Homologous genes are found in other herpes viruses. Cleavage sites recognized by the proteinase are identified in cytomegalovirus and are found to be highly conserved in other herpes viruses. Substrates, inhibitors, assay kits, and methods of assaying are provided which rely on the proteinase and its activity.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

HERPES VIRUS PROTEINASE AND METHOD OF ASSAYING

This is a continuation-in-part of application Serial No. 725,308, filed July 5, 1991.

This invention was supported under NIH Research Grants R01 AI22711 and R01 AI13718. The United States Government retains certain rights in this invention.

TECHNICAL AREA OF THE INVENTION

This invention relates to the area of herpes virology. More particularly, it relates to a new enzyme and the use of that enzyme as a target for anti-viral therapy.

BACKGROUND OF THE INVENTION

Herpes viruses are large double stranded DNA viruses that are responsible for a number of human diseases including chicken pox, shingles, fever blisters, salivary gland virus disease, and infectious mononucleosis. The seven human herpes viruses that have been described thus far are HSV-1, HSV-2, cytomegalovirus (CMV), Epstein-Barr Virus (EBV), varicella zoster virus (VZV), HHV-6, and HHV-7.

Maturation of herpes virus particles is believed to occur through the formation of a procapsid structure, which acquires DNA and an envelope to become an infectious virion. A herpes virus group-common protein referred to as the assembly protein in CMV, and as p40, VP22a, NCP-3, and ICP35e in HSV-1, is an abundant constituent of the herpes virus procapsid. The assembly protein is phosphorylated and proteolytically processed from a precursor molecule. It is absent from the mature virion, although its fate is unknown. These characteristics of the assembly protein have suggested an analogy between it and the bacteriophage scaffolding protein, which

is an essential component for phage assembly but is not found in mature virus particles (Gibson et al. (1991) J. Virol. 64:1241-1249).

The proteolytic processing of the assembly protein has been implicated as a critical step in the maturation of the virus. A temperature sensitive (ts) mutant that is unable to process the HSV assembly protein homolog (p40) is incapable of producing DNA-containing capsids or virions (Preston et al. (1983) J. Virol. 45:1056-1064). Maturational processing of the simian CMV (SCMV) Colburn assembly protein results in loss of its carboxy terminus. (Gibson, 1991, supra.)

Up until the present time the enzyme responsible for the proteolytic maturation of the assembly protein has not been identified. Further, there is a need in the art for new agents for therapeutic treatment of herpes viruses.

SUMMARY OF THE INVENTION

It is an object of the invention to provide a preparation of a proteinase encoded by a herpes virus.

It is another object of the invention to provide a substrate for cleavage by a herpes virus proteinase.

It is yet another object of the invention to provide a kit for measuring activity of a herpes virus proteinase.

It is still another object of the invention to provide a method for measuring activity of a herpes virus proteinase.

It is another object of the invention to provide a recombinant DNA molecule which encodes a herpes virus proteinase.

It is yet another object of the invention to provide an inhibitor of a herpes virus proteinase.

These and other objects of the invention are provided by one or more of the embodiments described below.

In one embodiment of the invention a preparation of the proteinase encoded by a herpes virus is provided, said preparation being free of a intact infectious herpes

virus virion DNA.

In another embodiment of the invention substrates for cleavage by a herpes virus proteinase are provided. One substrate comprises a polypeptide containing the amino acid sequence:

$$aa_1\text{-}aa_2\text{-Ala-}aa_3,$$

wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, and aa_3 is Ser, Val, or Asn, wherein the proteinase cleaves the substrate on the carboxy terminal side of the Ala residue. Another substrate comprises a polypeptide containing the amino acid sequence:

$$\text{Tyr-}aa_4\text{-}aa_5\text{-Ala-}aa_6,$$

wherein aa_4 is Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn, and wherein the proteinase cleaves the substrate on the carboxy terminal side of the Ala residue.

In yet another embodiment of the invention a kit is provided for measuring activity of a herpes virus proteinase. The kit comprises a proteinase encoded by a herpes virus, and a substrate for cleavage by said proteinase. The substrate comprises a polypeptide containing the amino acid sequence:

$$aa_1\text{-}aa_2\text{-Ala-}aa_3 \text{ or } \text{Tyr-}aa_4\text{-}aa_5\text{-Ala-}aa_6,$$

wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, aa_3 is Ser, Val, or Asn, aa_4 is Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn, wherein the proteinase cleaves the substrate on the carboxy terminal side of the Ala residue, said kit being substantially free of intact infectious herpes virus.

In still another embodiment of the invention a method is provided for measuring activity of a herpes virus proteinase. The method comprises the steps of: contacting a proteinase encoded by a herpes virus with a substrate for cleavage by said proteinase, said substrate comprising a polypeptide containing the amino acid sequence;

$$aa_1\text{-}aa_2\text{-Ala-}aa_3 \text{ or } \text{Tyr-}aa_4\text{-}aa_5\text{-Ala-}aa_6,$$

wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, aa_3 is Ser, Val, or Asn, aa_4 is

Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn, wherein the proteinase cleaves the substrate on the carboxy terminal side of the Ala residue, said step of contacting occurring in the absence of an intact infectious herpes virus virion DNA; and monitoring cleavage of said substrate.

In another embodiment of the invention a recombinant DNA molecule is provided which encodes at least a portion of the herpes virus proteinase, said protein having the ability to cleave a herpes virus assembly protein precursor.

In yet another embodiment of the invention an inhibitor of a herpes virus proteinase is provided. The inhibitor, comprises a derivative of the substrate of the herpes virus proteinase. The inhibitor may differ from the substrate in the scissile peptide bond which is carboxyl to the Ala residue.

These, and other embodiments of the invention which will be obvious to one skilled in the art from the disclosure, are described in more detail below. These embodiments provide the art with a promising target for specific anti-viral therapeutic agents, which can be administered to humans and other animals without also impairing normal cellular functions.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide and predicted amino acid sequences of the CMV Colburn genomic region containing the assembly protein gene at the 3'end of a 1,770-base pair open reading frame. The open reading frame, designated APNG1, denotes the beginning of the coding sequence of the proteinase gene, and the open reading frame designating APNG.5 denotes the beginning of the coding sequence of the precursor assembly protein gene. Each of the four designated open reading frames are in frame and are 3'co-terminal.

Figure 2 shows a comparison of portions of the putative active site domains of the proteinase in CMV Colburn, located between amino acids 15 and 195 and those in other herpes viruses. Two highly conserved motifs within this region are also identified in human CMV (HCMV), herpes simplex virus-1 (HSV-1), varicella zoster

virus (VZV), Epstein-Barr virus (EBV), and infectious laryngotracheitis virus (ILTV). The absolutely conserved amino acids are shown in bold type.

Figure 3A shows the cleavage site in the assembly protein of SCMV located between amino acids Ala_{557} and Ser_{558}. This region is shown as compared to homologous and conserved regions in other herpes viruses. Absolutely conserved amino acids are shown in bold type. The arrow in the sequence denotes the cleavage site. Figure 3B shows the cleavage site for release of the herpesvirus proteinase from the primary translation product of the APNG1 gene is located in six herpes viruses between amino acids 234 and 262. Absolutely conserved amino acids are shown in bold type. The space following the alanine residue denotes the cleavage site:

Figure 4 shows products of an in vitro transcription and translation of the cloned CMV Colburn assembly protein gene (APNG.5) as well as the cloned proteinase gene (APNG1). Proteins are revealed by their reactivity with antibodies (i.e., Anti-C-1) reactive only with noncleaved assembly protein nested gene products.

Figure 5 shows that assembly protein cleavage occurs in cells cotransfected with the genes for the assembly protein precursor (APNG.5/AW1) and for the proteinase (APNG1/AW4).

DETAILED DESCRIPTION

It is a finding of the present invention that the assembly protein of herpes viruses is maturationally processed by a herpes virus-encoded proteinase. Fascinatingly, the proteinase has been found to be a member of a family of four nested 3'co-terminal genes which includes the assembly protein itself. Each of the genes appears to be transcribed into separate mRNAs.

It appears that proteolytic cleavages may occur in this family of gene products at a number of locations. One such location, which has been identified with certainty, is the cleavage site within the assembly protein precursor itself. This site occurs between the alanine at amino acid position 557 and the serine at amino acid position 558 in the CMV Colburn APNG1 gene product. [Amino acid numbering in this

application begins with the first putative initiation codon of APNG1 shown in Figure 1 as the first underlined methionine codon.] The cleavage site in herpes virus assembly protein precursors have the conserved motif of aa_1-aa_2-Ala-aa_3, wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, and aa_3 is Ser, Val, or Asn. This cleavage site is herein referred to as the assembly protein maturation cleavage site. Another putative cleavage site within this family of nested proteins occurs after the Ala residue of the sequence Tyr-Val-Lys-Ala, which occurs at amino acids 246 to 249 in the CMV Colburn APNG1 gene product. This site has been used as the carboxy terminus of a recombinant construct, and the construct has been found to have proteinase activity. This suggests that this site may be used in vivo for autoprocessing of the proteinase molecule. The cleavage site in the primary translation product of the gene encoding for herpes virus proteinase have the conserved motif of Tyr-aa_4-aa_5-Ala-aa_6, wherein aa_4 is Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn. This cleavage site is herein referred to as the enzyme release cleavage site. Both the maturation cleavage site and the enzyme release cleavage site in SCMV are highly conserved among herpes viruses as shown in Figures 3A and 3B.

SEQ ID NOS 3-9 show the maturational cleavage site, located between numbered amino acids 11 and 12 in the assembly protein of CMV Colburn, HCMV, HSV-1, VZV, EBV, ILTV and HSV-6, respectively.

SEQ ID NOS 10-15 show the enzyme release cleavage site located between numbered amino acid 10 and 11, for the release of the herpesvirus proteinase of CMV Colburn, HCMV, HSV-1, VZV, EBV and ILTV, respectively.

According to the present invention a preparation of proteinase encoded by a herpes virus is provided. The preparation is substantially free of intact infectious herpes virus virion DNA. Virion DNA refers to the DNA which is present in viral particles. Preparations of the present invention can be provided which are totally free of virion DNA because they are produced in cells which have been transfected with a recombinant construct encoding the proteinase. Thus cells producing the proteinase

may not ever have been infected with herpes virus. The herpes virus proteinase from cytomegalovirus (CMV, simian strain Colburn) is encoded by a 1,770 base pair gene referred to APNG1 (assembly protein nested gene 1). The nucleotide and amino acid sequence of this gene is shown in SEQ ID NO 1. SEQ ID NO 2 shows only the amino acid sequence shown in SEQ ID NO 1. This gene has homologs in human CMV (HCMV, i.e., UL80a), herpes simplex virus-1 (HSV-1, i.e., UL26), varicella zoster virus (VZV, i.e., UL33), Epstein-Barr virus (EBV, i.e., BVRF2), infectious laryngotracheitis virus (ILTV, i.e., p40 gene), and probably in all herpes viruses. A proteinase according to the present invention may be all or an active portion of the APNG1 primary translation product, or its homologs on other herpes viruses. As previously alluded to, not all of the APNG1 primary translation product is necessary for proteinase activity. For example, constructs which have only the first 249 (LM8) or first 280 (LM7) amino acids beginning with the initial methionine codon on the APNG1 gene both demonstrate proteinase activity. Activity is defined as the ability to proteolytically process the assembly protein precursor of herpes virus to the mature assembly protein or to cleave site mimetic substances.

The preparation of proteinase of the present invention may be made in cells by recombinant DNA techniques, but need not be. The protein may be expressed in mammalian cells, as well as in bacterial, yeast, insect cells, and other cell types, as is convenient for a particular application or purpose. Alternatively, the protein can be chemically synthesized, or expressed in vitro using an in vitro transcription and/or translation system. In still another method of obtaining such a proteinase preparation, infected cells can be used as a source material and standard protein purification techniques can be used. Such purification techniques will typically include an affinity separation step (e.g., immunoaffinity; substrate affinity).

The active site domain of the proteinase enzyme has been tentatively identified as the region between and including amino acids 15 (Asp) and 95 (Ser) in the CMV Colburn APNG1 proteinase. This region contains two motifs that are highly

conserved among the homologous genes of HCMV, HSV-L, VZV, EBV, ILTV, and probably all herpes viruses. See Figure 2. These motifs are referred to as conserved motif 1 and conserved motif 2.

SEQ ID NOS 17, 19, 21, 23, 25 and 27, respectively, correspond to conserved motif 1 of CMV Colburn, HCMV, HSV-1, VZV, EBV and ILTV, respectively.

SEQ ID NOS 16, 18, 20, 22, 24 and 26, respectively, correspond to conserved motif 2 of CMV Colburn, HCMV, HSV-1, VZV, EBV and ILTV, respectively.

A recombinant construct of the proteinase gene was made having a 15 amino acid insertion between conserved motifs 1 and 2. This construct had greatly diminished (i.e., less than about 1% of the wild-type level) proteinase activity, which supports the assignment of the active site domain.

The cleavage site in the assembly protein precursor (i.e., the maturation cleavage site) which leads to formation of the mature assembly protein has been defined with particularity. In simian CMV (Colburn) the cleavage site has been defined as occurring between amino acids 557 and 558. The sequence immediately surrounding this site is Val-Asn-Ala-Ser-Cys. When the assembly protein sequences of other herpes viruses are compared it is found that this site is well conserved. (See Figure 3.) The consensus cleavage site appears to require aa_1-aa_2-Ala-aa_3, wherein aa_1, is Val or Leu, aa_2 is a polar amino acid, and aa_3 is Ser or Val. The amino acid represented as aa_2 is most often an asparagine residue.

While not wishing to be bound by any particular theory, there is evidence (Welch et al. (1991) Proc. Natl. Acad. Sci. USA, in press) that an additional cleavage site or sites for the proteinase occurs near the middle of the proteinase sequence. It is likely that the proteinase which is responsible for the maturational cleavage of the assembly protein is also involved in self-processing, possibly to create an active form of the proteinase. The carboxyl half of the APNG1 gene product has been identified in transfected cells, indicating that cleavage in the middle of the APNG1 primary translation product is biologically relevant. (See Figure 5, $APNG1_c$.) Cleavage at

this site (i.e., the enzyme release cleavage site) may be required for the life cycle of the herpes viruses. The consensus sequence for this site comprises Tyr-aa_4-aa_5-Ala-aa_6, wherein aa_4 is Leu or Val, aa_5 is Lys or Gln, and aa_6 is Ser or Asn.

Having defined the actual cleavage site in the assembly protein precursor and putative cleavage site in the proteinase, it is now possible to design smaller synthetic moieties which can be used as substrates for cleavage by the herpes virus proteinase. These substrates for cleavage typically comprise a polypeptide having an amino acid sequence which has been shown to be a recognized cleavage site by a herpes virus proteinase. The polypeptides will contain the amino acid sequence aa_1-aa_2-Ala-aa_3 or Tyr-aa_4-aa_5-Ala-aa_6 and most often will contain the amino acid sequence aa_1-aa_2-Ala-Ser or Tyr-aa_4-aa_5-Ala-Ser. The substrate is substantially free of the assembly protein precursor or the entire primary translation product of the gene encoding the herpes virus proteinase. This is possible because the entire assembly protein precursor or the entire primary translation product of the gene encoding the proteinase need not be used as a substrate. Synthetic or recombinant substrates can be made which are recognized and cleaved by a herpes virus proteinase. Substrates for the proteinase will typically comprise a polypeptide portion of between about 15 and 25 amino acids. A sufficient number is required for the proteinase to be able to recognize and bind to the cleavable site. Extraneous amino acids are not desirable because they may cause steric inhibition by formation of three-dimensional structures which block the cleavage site. Substrates which mimic the maturation cleavage site or the enzyme release cleavage site can also be made.

The substrate itself need not be a totally proteinaceous molecule. It may be linked to other moieties and polymers as is convenient. The substrate will typically be used for assaying proteinase activity in cellular, extracts or in synthetic proteinase preparations, as described above, as well as for, screening for inhibitory substances which block the proteinase cleavage reaction. In one embodiment of the present invention the polypeptide portion of the substrate is linked to a fluorescent moiety and

a quenching moiety. Typically these will be linked on opposite ends of the polypeptide. While linked to the polypeptide the fluorescent moiety will not fluoresce due to the proximity of the quenching moiety. However, upon cleavage of the polypeptide, the separation of the two moieties will lead to a loss of quenching and to detectable fluorescence. An example of a similar quenched fluorogenic substrate is taught by Matayoshi, et al. (Science (1990) 247:954-958). There the fluorogenic and quenching moieties employed are 4-(4-dimethylaminophenylazo)benzoic acid (DABCYL) and 5-[(2-aminoethyl)amino]naphthalane-1 sulfonic acid (EDANS). As another example of an indicator substrate, a substrate having the cleavage site engineered into a protein, such as β-galactosidase or luciferase, so that cleavage inactivates the activity of the indicator, is mentioned.

In another embodiment of the invention the substrate for cleavage of a herpes virus proteinase is labeled with a radioactive moiety. After exposure of the substrate to the proteinase, the chemical or physical properties of the radioactive species can be determined, specific changes in these properties can be used to monitor cleavage by the proteinase. One such property is size, a reduction in size of the radioactive species indicating cleavage by the proteinase. Alternatively, after exposure of the substrate to the proteinase, the substrate can be extracted into a solvent. A change in the extractability of the radioactive species can be used to indicate cleavage of the substrate. In yet another embodiment of the invention an enzyme is linked to the polypeptide comprising the cleavage site. The polypeptide sterically inhibits the activity of the enzyme. However, upon cleavage of the polypeptide moiety the steric inhibition is relieved and the enzyme activity is regained and can be assayed. Increase of enzyme activity therefore is an indication of cleavage. In an alternative embodiment, the substrate for the enzyme which is linked to the polypeptide for cleavage is also linked to the polypeptide for cleavage. Again, the enzyme is sterically inhibited by its linkage to the polypeptide. However, upon cleavage of the polypeptide the steric inhibition is released and the enzyme can interact with its

substrate.

Having discovered the proteinase of herpes virus and its particular sites for cleavage (i.e., the maturation cleavage site and the enzyme release cleavage site), a kit can be readily prepared for measuring the activity of a herpes virus proteinase. The kit comprises a proteinase, or portion thereof, encoded by herpes virus and a substrate for cleavage by said proteinase. The substrate for cleavage has the properties described above. Briefly, a substrate for cleavage contains a polypeptide having the amino acid sequence aa_1-aa_2-Ala or Tyr-aa_4-aa_5-Ala, and the proteinase cleaves the substrate on the carboxy terminal side of such sequences. The kit is substantially free of intact infectious herpes virus. This purity can be achieved in a number of ways. Preferably, it can be achieved by expressing the proteinase and the substrate for cleavage in a mammalian cell which is free of herpes virus infection. The cleavage of the substrate occurs within the mammalian cell and can be monitored by observation of a change in the size of the substrate, for example. Alternatively, the proteinase and the substrate can be expressed in an in vitro cell-free system, such as a rabbit reticulocyte system, or synthesized chemically. In such cases the two components of the kit can be contacted in vitro and the cleavage reaction observed. The proteinase and the substrate can also be expressed in separate cells of any suitable species. The cells may be either mammalian, bacterial, yeast, insect, or other cell type, as is convenient for the particular application involved. After separately expressing the proteinase and its substrate they can be contacted in vitro to determine an amount of herpes virus proteinase activity.

In another embodiment of the invention, the cleavage reaction can be used diagnostically to test for, the presence of a herpes virus. For example, putatively infected cells can be used as a source of proteinase and contacted with a substrate for cleavage. The cleavage of the substrate would indicate the presence in the source of a herpes virus proteinase and therefore of a herpes virus infection.

Also contemplated by the present invention is a method for measuring activity

of a herpes virus proteinase. According to the method, a proteinase encoded by a herpes virus is contacted with a substrate for cleavage by the proteinase. The substrate for cleavage has the properties described above. The contacting of the substrate with the proteinase occurs in the absence of intact infectious herpes virus virion DNA; this can be accomplished by using as sources of substrate and proteinase cells which are not infected with a herpes virus. The second step of the method involves monitoring cleavage of the substrate. Such monitoring can be accomplished by determining a change of size of said substrate, for example, by observing an altered mobility of the substrate on an electrophoretic gel matrix or on a chromatography medium. Alternatively, the monitoring can be accomplished by observing a change of fluorescence if the substrate has been labelled with a fluorescent moiety as described above. If the substrate has been labelled with a radiolabelled moiety then the cleavage reaction can be monitored by looking for a change in its physical properties, as described above. In another embodiment a substrate that has been labelled with an enzyme is used and the cleavage reaction is monitored by determining a colorimetric change of a chromogenic substrate for the enzyme. Suitable enzymes for such purposes are known in the art and include β-galactosidase, alkaline phosphatase, and luciferase.

In one embodiment of the method of the present invention, a test substance is also added to the proteinase (or active portion thereof) and substrate to determine the level of inhibition caused by the test substrate. This method can be used as a screen for potential therapeutic molecules. The level of inhibition can be readily determined by measuring the activity of the proteinase in the presence and absence of the test substrate. A significant diminution of the activity of the proteinase in the presence of the test substance indicates a potential anti-herpetic agent.

Inhibitors of the herpes virus proteinase are also provided by the present invention. Typically, these are non-cleavable derivatives of substrates of the proteinase. The inhibitors may comprise a polypeptide portion of about 6 to 12 amino

acids and often will mimic the structure of the appropriate substrate for the proteinase. However, the inhibitor may differ from the substrates for the enzyme in having a modification of the scissile peptide bond which is carboxyl to the sequence aa_1-aa_2-Ala or Tyr-aa_4-aa_5-Ala. Any modification of this bond can be used which partially inhibits or totally blocks the proteinase cleavage. Such modifications of the scissile peptide bond include replacement by a hydroxyethylamine linkage, a phosphonamide linkage, a carbon fluoride aldehyde, and a dialcohol linkage. Such inhibitors will bind to the proteinase active site domain but will be either totally non-cleavable or cleavable at a much lower rate than a proper substrate. As the cleavage reaction is known to be essential for the formation of herpes virus particles, inhibition of the cleavage reaction can be used as an anti-herpetic therapeutic treatment.

Certain modifications to the inhibitors of the present invention may be desired in order to render them more resistant to proteolysis in the human body or to render them more easily taken up by infected cells. One such modification is to place an amide moiety on the carboxy terminal end of the polypeptide. This reduces the charge of the molecule rendering it more accessible to cells. Another possible modification involves placing a D-tyrosine moiety on the amino terminal end of the inhibitor. This renders the inhibitor less susceptible to proteolysis.

Other inhibitors may now be designed based on the 3-dimensional structure of the proteinase. Typically, X-ray crystallography is used to determine a structure for the enzyme and inhibitors are designed to conform to the determined structure. Since it has been shown that proteinase activity resides within the first 249 amino acids of the CMV Colburn APNG1 protein, the use of X-ray crystallography to determine the 3-dimensional structure of the amino terminal 249 residues can be used to design inhibitors of this proteolytically active sequence.

Recombinant DNA molecules are also provided by the present invention. These molecules encode at least a portion of the herpes virus proteinase. The proteinase portion retains the ability to cleave a herpes virus assembly protein.

Applicants have found that the entire proteinase gene which is transcribed in vivo as a 1.8 kb RNA molecule, is not necessary for expression of proteinase activity. It has been determined that the portion of the APNG1 gene encoding the assembly protein precursor is not needed for proteolytic activity. Portions of the proteinase which comprise only amino acids 1 through 249 have been found to retain proteolytic activity. Further, as discussed above, it is possible that further, shortening of the proteinase molecule is possible without loss of proteolytic activity.

EXAMPLES

Example 1

This example provides the sequence of the simian CMV proteinase gene and compares portions of it to other herpes virus sequences.

The XbaI R fragment of strain Colburn CMV DNA was cloned into the plasmid pUC18, and the nucleotide sequence of both strands was determined by the dideoxy nucleotide chain termination method (Sanger et al., Proc. Natl. Acad. Sci. USA (1977) 74:5463-5467) with appropriate DNA oligonucleotide primers and the Sequenase kit (USB, Cleveland, Ohio).

Nucleotide sequence analysis of the CMV (Colburn) genomic XbaI R fragment confirmed the cDNA sequence previously determined for the assembly protein-coding region and revealed that the 930-bp coding sequence for the assembly protein precursor (nucleotides 1072 to 2001) is the 3′ end of a 1,770-bp open reading frame (ORF) (nucleotides 232 to 2001) that begins with a methionine and, together with its upstream regulatory region, was designated assembly protein nested gene 1 (APGNG1) (Figure 1). APNG1 includes an upstream potential TATA promoter element, contains three internal potential TATA promoters and three corresponding ATG translational start codons in addition to its own, and is followed by a single downstream polyadenylation signal. This organization indicated that the APNG1 region could give rise to four 3′co-terminal mRNAs able to encode four corresponding in-frame, overlapping proteins. These nested coding sequences are numbered

according to their fractional length relative to that of the longest, APNG1. Figure 1 presents the nucleotide and amino acid sequences of the APNG1 region and shows the positions of (1) proposed TATA promoter elements (italicized and dot underlined), (2) proposed translational start methionines for the coding sequence in each of the nested genes (capitalized and doubly underlined, and the designation of the corresponding assembly protein nested gene (APNG) is indicated above each), (3) the single polyadenylation signal at the 3′ end (underlined). The APNG1 (proteinase) gene has homologs in human CMV (HCMV, i.e., UL80a), herpes simplex virus type-1 (HSV-1, i.e., UL27), varicella zoster virus (VZV, i.e., UL33), Epstein-Barr virus (EBV, i.e., BVRF2), and infections laryngotracheitis virus (ILTV, i.e., p40).

At least a portion of the active site domain of the proteinase has now been tentatively identified as the region between amino acids 15 and 195 in the CMV Colburn APNG1 protein. This region contains two motifs that are highly conserved among the homologous genes of HCMV, HSV-L, VZV, EBV, and ILTV, and probably all herpes viruses (Figure 2). These motifs are referred to as "conserved motif 1" (CM1) and "conserved motif 2". Striking similarities in the spacing of possible active site residues resembling both cysteine (i.e., His_{47}, Cys_{146}, His_{142}) and serine (His_{47}, Asp_{104}, Ser_{195}) proteinases are detected among all six herpes viruses, suggesting that the herpes virus proteinases may have two separate proteolytic activities.

It has been found that an altered form of APNG1 (LM3) which contains a 15 amino acid sequence (the C3 epitope of poliovirus VP2) inserted between CM1 and CM2, has only a trace amount of proteinase activity (i.e., $\leq$ 1%) (see Figure 5, lane g). Insertion of the same sequence into the carboxyl end of APNG1 did not reduce proteinase activity (Figure 5, lane f). This suggests that the CM1/CM2 region does contain at least a portion of the active site domain of this proteinase.

Furthermore, two subclones of APNG1 were made which expressed portions of the proteinase gene comprising amino acids 1-249 (LM8) and 1-280 (LM7). Both

are proteolytically active using assembly protein precursor as a substrate. This, too, supports the active site domain assignment.

Example 2

This example demonstrates the precise cleavage site involved in the maturational processing of assembly protein precursor to assembly protein, as well as the conservation of the site among herpes viruses.

The mature assembly protein was treated with endoproteinase Lys-C or endoproteinase Glu-C (V8 proteinase). Specific peptide products were isolated and subjected to analysis by mass spectrometry. The diagnostic molecular ions identified from HPLC-purified peptides of the Colburn CMV assembly protein were mass 902.5 (Endo-Lys-C fragment, SAERGVVNA) and mass 616.4 (Endo-Glu-C fragment, RGVVNA). Thus the cleavage site is between Ala_{557} and Ser_{558} in SCMV Colburn. This cleavage site is well conserved in HCMV, HSV-I, VZV, EBV, ILTV, and probably in all herpes group viruses (Figure 3).

Example 3

This example provides proteinase substrate derivatives with altered chemistry at the scissile peptide bond.

Based on the cleavage site sequence, several classes of anti-herpes virus peptide mimetics can be synthesized. These include hydroxyethylamine-, dialcohol-, phosphonamide-, and carbon fluoride aldehyde-derivatives of the scissile peptide bond (i.e., carboxyl to the alanine in the general sequences aa_1-aa_2-Ala-aa_3 and Tyr-aa_4-aa_5-Ala-aa_6, such as:

aa_1-aa_2-Ala-CHOH-CH_2-NH-aa_3 (hydroxyethylamine derivative)

aa_1-aa_2-Ala-PO_2-NH-aa_3 (phosphonamide derivative)

aa_1-aa_2-Ala-CF_2CHO (carbon fluoride aldehyde)

aa_1-aa_2-Ala-$C(OH)_2$-NH-aa_3 (dialcohol derivative)

Tyr-aa_4-aa_5-Ala-CHOH-CH_2-NH-aa_6 (hydroxyethylamine derivative)

Tyr-aa_4-aa_5-Ala-PO_2-NH-aa_6 (phosphonamide derivative)

Tyr-aa_4-aa_5-Ala-CF_2-CHO (carbon fluoride aldehyde)

Tyr-aa_4-aa_5-Ala-C(OH_2)-NH-aa_6 (dialcohol derivative)

Example 4

This example demonstrates the in vitro transcription and translation of the cloned CMV Colburn assembly protein precursor gene.

The assembly protein precursor gene (APNG.5, see Figure 1) and the overlapping APNG1 gene were cloned from the simian strain Colburn CMV DNA, using PCR amplification, into a pGEM4Z plasmid to produce plasmids AW2 and AW3, respectively. T7-promoted run-off transcripts of both genes were prepared, and translated in rabbit reticulocyte lysates containing ^{35}S-methionine. The proteins were separated by electrophoresis in an SDS-containing polyacrylamide gel (10%), electrotransferred onto an Immobilon-P™ membrane and probed with an antiserum to the carboxyl 21 amino acids of the assembly protein precursor (i.e., Anti-C1, see Schenk, et al. (1991) J. Virol. 65: 1525-1529). The resulting protein-antibody complexes were visualized by using ^{125}I-Protein A. A fluorogram of the blot is shown in Figure 4. The exposure technique used recorded only ^{125}I-radioactivity (i.e., Kodak DEF film and black paper between blot and recording Kodak XAR film).

The in vitro translated assembly protein (lane 3, APNG.5/AW2) comigrated with the infected cell assembly protein precursor (i.e., 40-kDa band in lane 5) and was not proteolytically processed in the reticulocyte lysate. The protein product of the APNG1 gene (lane 4) comigrated with the 85-kDa protein present in the Colburn CMV-infected cell cytoplasm (i.e., 85-kDa band in lane 5). Mock infected nuclear and cytoplasmic fractions (lanes 1 and 6) show no evidence of proteins reactive with the Anti-C1 antibody.

Example 5

This example demonstrates that assembly protein cleavage occurs in cells cotransfected with the genes for the assembly protein precursor and for the APNG1 protein.

Human cells were transfected with an expression plasmid containing the gene for the assembly protein precursor (AW1), or with an expression plasmid containing the gene for the APNG1 protein (AW4), or with both plasmids (AW1+AW4). Parallel cotransfections were done using the AW1 plasmid in combination with altered versions of AW4 that contain (1) a 13 amino acid sequence inserted into the carboxyl end of APNG1 (LM2), or (2) a 15 amino acid sequence inserted into the amino end of APNG1 (LM3).

Following transfection the cells were solubilized and the proteins were separated by electrophoresis in an SDS-containing polyacrylamide gel (SDS-PAGE). The resolved proteins were electrotransferred to an Immobilon-P™ membrane and visualized by probing the membrane with an antiserum (Anti-N1) that reacts with the amino end of the assembly protein and with other assembly protein nested gene products (e.g., 85-kDa APNG1 protein). Colburn CMV-infected cell proteins were run as markers. The results of these experiments are shown in Figure 5.

The assembly protein (AP) and its precursor (preAP) can be seen in the cytoplasmic and nuclear fractions of Colburn CMV-infected cells (lanes a and b, respectively). The precursor form is more abundant in the cytoplasm (lane a) and the mature form is more abundant in the nucleus (lane b).

Cells transfected with only the gene coding for the assembly protein precursor (i.e., plasmid AW1) expressed the precursor form of the assembly protein but no product (lane d). The much less abundant, slightly larger protein is believed to correspond to a protein also detected in infected cells, but present there in vanishingly small amounts.

Cells transfected with only the gene coding for the larger protein of the "assembly protein nested gene family" (i.e., 85-kDa protein product of APNG1 encoded by plasmid AW4) expressed the 85k-Da protein (APNG1), and a second doublet band believed to represent the carboxyl end of the APNG1 protein (i.e., $APNG1_c$, 45 to 50 kDa in size, lane e).

Cells cotransfected with the assembly protein precursor gene and the APNG1 gene (i.e., AW1+AW4) contained: (1) the APNG1 and APNG1, proteins encoded by AW4, and (2) predominantly the mature (i.e., cleaved) form of the assembly protein (AP) (lane c). Essentially no precursor form of the assembly protein (preAP) was detected with this serum (Anti-N1) or with an even more sensitive antiserum for the precursor form (i.e., Anti-C1). This finding indicates that cleavage of the precursor in transfected cells is highly efficient.

Cells cotransfected with the assembly protein precursor gene (AW1) and an altered form of the APNG1 gene (LM2, altered at the carboxyl end) that contains a 45 bp insert at the single APNG1 DraI site, encoding the polio virus VP2 C3 epitope (Charbit, et al., 1986, EMBO J., 5:5029-3038) contained: (1) a higher molecular weight form of APNG1 (ΔAPNG1, indicated by dot to left of lane f) resulting from the inserted 15 amino acids of VP2 C3, (2) a correspondingly larger form of the carboxyl end of the APNG1 protein (i.e., ΔAPNG1,), and (3) only the mature (i.e., cleaved) form of the assembly protein (AP)(lane f). The band close to the position of the assembly protein precursor (i.e., preAP) in lane 4 is not reactive with Anti-C1 and is believed to correspond to the processed form of the slightly larger, low abundance protein expressed by AW1 and mentioned above.

Cells cotransfected with the assembly protein precursor gene (AW1) and an altered form of the APNG1 gene (LM3, altered at amino, N-end) that contains a 45 bp insert (VP2 C3) at the single APNG1 EcoRV site, located between the highly conserved motifs CM1 and CM2, contained: (1) a higher molecular weight form of APNG1 (ΔAPNG1, indicated by arrow to left of lane g) resulting from the inserted

15 amino acids of VP2 C3, (2) a weak band at the position of ΔAPNG1$_c$, but corresponding in size to the band in lane f because the carboxyl cleavage at Val-Asn-Ser has not occurred (i.e., band in lane g reacted with Anti-C1, in contrast to band in lane f), and (3) only the noncleaved, precursor form of the assembly protein (preAP)(lane g). Again, the lower abundance band above the assembly protein precursor in lane g is thought to correspond to a scarce species also detected in infected cells.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: GIBSON, D. WADE
WELCH, ANTHONY R.

(ii) TITLE OF INVENTION: HERPES VIRUS PROTEINASE AND METHOD OF ASSAYING

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Banner, Birch, McKie & Beckett
(B) STREET: 1001 G Street, N.W., Eleventh Floor
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: U.S.A.
(F) ZIP: 20001-4597

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/798,776
(B) FILING DATE: 27-NOV-1991
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: FOULKE, CYNTHIA L.
(B) REGISTRATION NUMBER: 32,364
(C) REFERENCE/DOCKET NUMBER: 1107.037080

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (202) 508-9100
(B) TELEFAX: (202) 508-9299
(C) TELEX: 197430 BBMB UT

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Simian cytomegalovirus
(B) STRAIN: Colburn

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 175..2001

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 175..231

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 232..2001

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
TTGTCCGACA CCCCCAGGTT ATTGGTGGTC TCGCGGGGGG GGAACAGGGG GGTTTGCAGG
60

CCTCGGTTAA AGAGCAGCAC GCAGATGAGT CTCAAGATCT TGAGTTCTTC CAGCCGCAGG
120

GTGTTGAGCG GCTGTCCCCG CGACATCTTT TCGCTGATCT GTAATATTAG ATGA TTG
177
                                                              Leu
                                                              -19

GCA CAA GTA AAG GAG AAT TTG CCG GTT CGA ACC CGG GCC TCC TCC GTG
225
Ala Gln Val Lys Glu Asn Leu Pro Val Arg Thr Arg Ala Ser Ser Val
          -15                 -10                 -5

TTG GAC ATG GCC GAT CCC GTC TAC GTC GGG GGT TTT TTG GTG CGC TAC
273
Leu Asp Met Ala Asp Pro Val Tyr Val Gly Gly Phe Leu Val Arg Tyr
        1               5                  10

GAC GAG CCT CCC GGA GAA GCT GAG CTG TTT CTG CCC TCG GGG GTG GTA
321
Asp Glu Pro Pro Gly Glu Ala Glu Leu Phe Leu Pro Ser Gly Val Val
 15                  20                  25                  30

GAC CGC TGG TTG CGC GAT TGC CGA GGC CCG CTG CCC CTG AAT GTC AAT
369
Asp Arg Trp Leu Arg Asp Cys Arg Gly Pro Leu Pro Leu Asn Val Asn
                35                  40                  45

CAC GAC GAG TCG GCG ACC GTG GGC TAT GTG GCT GGG CTC CAG AAT GTC
417
His Asp Glu Ser Ala Thr Val Gly Tyr Val Ala Gly Leu Gln Asn Val
            50                  55                  60

CGG GCC GGC TTG TTC TGT TTG GGA CGT GTT ACG TCC CCC AAG TTT CTG
465
Arg Ala Gly Leu Phe Cys Leu Gly Arg Val Thr Ser Pro Lys Phe Leu
        65                  70                  75
```

```
GAT ATC GTT CAA AAA GCC TCG GAA AAA TCC GAG TTG GTG TCC CGG GGA
513
Asp Ile Val Gln Lys Ala Ser Glu Lys Ser Glu Leu Val Ser Arg Gly
 80                  85                  90

CCT CCG TCC GAG TCC TCG TTG CGG CCG GAC GGC GTG TTG GAG TTT CTC
561
Pro Pro Ser Glu Ser Ser Leu Arg Pro Asp Gly Val Leu Glu Phe Leu
 95                 100                 105                 110

AGC GGC AGT TAT TCG GGC CTG TCG CTC TCC AGC CGC CGA GAT ATA AAC
609
Ser Gly Ser Tyr Ser Gly Leu Ser Leu Ser Ser Arg Arg Asp Ile Asn
                115                 120                 125

GCG GCC GAT GGC GCC GCG GGC GAT GCA GAA ACA GCG TGC TTC AAA CAT
657
Ala Ala Asp Gly Ala Ala Gly Asp Ala Glu Thr Ala Cys Phe Lys His
            130                 135                 140

GTG GCT CTG TGC AGC GTG GGC CGC CGC CGG GGC ACG TTG GCG GTG TAT
705
Val Ala Leu Cys Ser Val Gly Arg Arg Arg Gly Thr Leu Ala Val Tyr
        145                 150                 155

GGC AGG CAG CCA GAT TGG GTG ATG GAA CGT TTC CCG GAT CTC ACC GAG
753
Gly Arg Gln Pro Asp Trp Val Met Glu Arg Phe Pro Asp Leu Thr Glu
    160                 165                 170

GCC GAC CGG GAA GCG CTG CGA AAT CAG CTA TCG GGA AGT GGG GAA GTT
801
Ala Asp Arg Glu Ala Leu Arg Asn Gln Leu Ser Gly Ser Gly Glu Val
175                 180                 185                 190

GCC GCG AAG GAA AGT GCG GAA TCG TCT GCC GCC GCC GCC GTC GAT CCC
849
Ala Ala Lys Glu Ser Ala Glu Ser Ser Ala Ala Ala Ala Val Asp Pro
                195                 200                 205

TTT CAG TCG GAT TCG TAC GGG CTG TTG GGG AAC AGT GTG GAC GCG CTG
897
Phe Gln Ser Asp Ser Tyr Gly Leu Leu Gly Asn Ser Val Asp Ala Leu
            210                 215                 220

TAC ATT CAA GAG CGT CTC CCT AAG CTG CGC TAT GAC AAG CGG CTG GTC
945
Tyr Ile Gln Glu Arg Leu Pro Lys Leu Arg Tyr Asp Lys Arg Leu Val
        225                 230                 235

GGG GTC ACG GCT CGG GAG TCG TAC GTG AAA GCC AGT GTT TCG CCC GCC
993
Gly Val Thr Ala Arg Glu Ser Tyr Val Lys Ala Ser Val Ser Pro Ala
    240                 245                 250

GAG CAG GAG ACG TGC GAT ATT AAA GTA GAA AAA GAG CGG CCG AAG GAG
1041
Glu Gln Glu Thr Cys Asp Ile Lys Val Glu Lys Glu Arg Pro Lys Glu
255                 260                 265                 270
```

```
CCA GAG CAG AGC CAC GTA CCG ACC GAG TCA ATG TCT CAC CCT ATG AGC
1089
Pro Glu Gln Ser His Val Pro Thr Glu Ser Met Ser His Pro Met Ser
                275                 280                 285

GCC GTG GCT ACT CCG GCG GCC TCG ACC GTC GCG CCT TCT CAG GCG CCG
1137
Ala Val Ala Thr Pro Ala Ala Ser Thr Val Ala Pro Ser Gln Ala Pro
            290                 295                 300

CTG GCG CTG GCC CAT GAC GGT GTT TAT TTA CCT AAA GAC GCT TTT TTC
1185
Leu Ala Leu Ala His Asp Gly Val Tyr Leu Pro Lys Asp Ala Phe Phe
        305                 310                 315

TCG CTC ATC GGG GCC AGT CGT CCC CTG GCC GAG GCG GCG GGA GCG CGC
1233
Ser Leu Ile Gly Ala Ser Arg Pro Leu Ala Glu Ala Ala Gly Ala Arg
    320                 325                 330

GCC GCG TAT CCG GCT GTC CCG CCG CCA CCC GCG TAT CCG GTA ATG AAT
1281
Ala Ala Tyr Pro Ala Val Pro Pro Pro Pro Ala Tyr Pro Val Met Asn
335                 340                 345                 350

TAT GAG GAC CCC TCC TCA CGT CAC TTT GAC TAC AGT GCC TGG CTG CGG
1329
Tyr Glu Asp Pro Ser Ser Arg His Phe Asp Tyr Ser Ala Trp Leu Arg
                355                 360                 365

CGG CCA GCT TAT GAC GCC GTG CCT CCC CTG CCT CCT CCC CCC GTC ATG
1377
Arg Pro Ala Tyr Asp Ala Val Pro Pro Leu Pro Pro Pro Pro Val Met
            370                 375                 380

CCC ATG CCG TAT CGC AGA CGC GAC CCC ATG ATG GAG GAG GCC GAG CGC
1425
Pro Met Pro Tyr Arg Arg Arg Asp Pro Met Met Glu Glu Ala Glu Arg
        385                 390                 395

GCC GCC TGG GAG CGC GGG TAC GCG CCT TCT GCT TAT GAC CAC TAC GTG
1473
Ala Ala Trp Glu Arg Gly Tyr Ala Pro Ser Ala Tyr Asp His Tyr Val
    400                 405                 410

AAC AAC GGC TCC TGG TCG CGG AGC CGC AGC GGC GCG CTC AAG AGG CGA
1521
Asn Asn Gly Ser Trp Ser Arg Ser Arg Ser Gly Ala Leu Lys Arg Arg
415                 420                 425                 430

AGG GAG CGC GAC GCG TCC TCG GAT GAG GAA GAG GAC ATG AGT TTT CCC
1569
Arg Glu Arg Asp Ala Ser Ser Asp Glu Glu Glu Asp Met Ser Phe Pro
                435                 440                 445

GGG GAA GCC GAC CAC GGC AAG GCT CGG AAA AGA CTC AAA GCT CAT CAC
1617
Gly Glu Ala Asp His Gly Lys Ala Arg Lys Arg Leu Lys Ala His His
            450                 455                 460
```

```
GGG CGT GAT AAT AAC AAC TCT GGG AGC GAT GCC AAG GGC GAT CGG TAC
1665
Gly Arg Asp Asn Asn Asn Ser Gly Ser Asp Ala Lys Gly Asp Arg Tyr
        465                 470                 475

GAC GAC ATT CGG GAA GCG TTA CAG GAG CTG AAG CGC GAG ATG CTG GCC
1713
Asp Asp Ile Arg Glu Ala Leu Gln Glu Leu Lys Arg Glu Met Leu Ala
    480                 485                 490

GTG CGG CAG ATC GCG CCA CGT GCG CTC TTG GCC CCC GCA CAG CTA GCG
1761
Val Arg Gln Ile Ala Pro Arg Ala Leu Leu Ala Pro Ala Gln Leu Ala
495                 500                 505                 510

ACG CCC GTG GCT TCT CCG ACA ACG ACC ACG TCG CAT CAA GCC GAG GCT
1809
Thr Pro Val Ala Ser Pro Thr Thr Thr Thr Ser His Gln Ala Glu Ala
                515                 520                 525

AGC GAA CCT CAG GCA TCG ACT GCC GCT GCC GCG TCG CCG TCA ACC GCT
1857
Ser Glu Pro Gln Ala Ser Thr Ala Ala Ala Ala Ser Pro Ser Thr Ala
            530                 535                 540

TCG TCG CAC GGC AGC AAG TCG GCC GAA CGC GGG GTG GTG AAC GCC TCG
1905
Ser Ser His Gly Ser Lys Ser Ala Glu Arg Gly Val Val Asn Ala Ser
        545                 550                 555

TGT CGC GTT GCG CCT CCG TTG GAG GCT GTG AAC CCC CCT AAG GAC ATG
1953
Cys Arg Val Ala Pro Pro Leu Glu Ala Val Asn Pro Pro Lys Asp Met
    560                 565                 570

GTG GAC TTG AAT CGT CGC CTG TTT GTG GCG GCG TTG AAT AAA ATG GAA
2001
Val Asp Leu Asn Arg Arg Leu Phe Val Ala Ala Leu Asn Lys Met Glu
575                 580                 585                 590

TAAAAACTCG TAC
2014
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 609 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Leu Ala Gln Val Lys Glu Asn Leu Pro Val Arg Thr Arg Ala Ser Ser
-19              -15                 -10                  -5
```

Val Leu Asp Met Ala Asp Pro Val Tyr Val Gly Gly Phe Leu Val Arg
1 5 10

Tyr Asp Glu Pro Pro Gly Glu Ala Glu Leu Phe Leu Pro Ser Gly Val
15 20 25

Val Asp Arg Trp Leu Arg Asp Cys Arg Gly Pro Leu Pro Leu Asn Val
30 35 40 45

Asn His Asp Glu Ser Ala Thr Val Gly Tyr Val Ala Gly Leu Gln Asn
50 55 60

Val Arg Ala Gly Leu Phe Cys Leu Gly Arg Val Thr Ser Pro Lys Phe
65 70 75

Leu Asp Ile Val Gln Lys Ala Ser Glu Lys Ser Glu Leu Val Ser Arg
80 85 90

Gly Pro Pro Ser Glu Ser Ser Leu Arg Pro Asp Gly Val Leu Glu Phe
95 100 105

Leu Ser Gly Ser Tyr Ser Gly Leu Ser Leu Ser Ser Arg Arg Asp Ile
110 115 120 125

Asn Ala Ala Asp Gly Ala Ala Gly Asp Ala Glu Thr Ala Cys Phe Lys
130 135 140

His Val Ala Leu Cys Ser Val Gly Arg Arg Arg Gly Thr Leu Ala Val
145 150 155

Tyr Gly Arg Gln Pro Asp Trp Val Met Glu Arg Phe Pro Asp Leu Thr
160 165 170

Glu Ala Asp Arg Glu Ala Leu Arg Asn Gln Leu Ser Gly Ser Gly Glu
175 180 185

Val Ala Ala Lys Glu Ser Ala Glu Ser Ser Ala Ala Ala Ala Val Asp
190 195 200 205

Pro Phe Gln Ser Asp Ser Tyr Gly Leu Leu Gly Asn Ser Val Asp Ala
210 215 220

Leu Tyr Ile Gln Glu Arg Leu Pro Lys Leu Arg Tyr Asp Lys Arg Leu
225 230 235

Val Gly Val Thr Ala Arg Glu Ser Tyr Val Lys Ala Ser Val Ser Pro
240 245 250

Ala Glu Gln Glu Thr Cys Asp Ile Lys Val Glu Lys Glu Arg Pro Lys
255 260 265

Glu Pro Glu Gln Ser His Val Pro Thr Glu Ser Met Ser His Pro Met
270 275 280 285

Ser Ala Val Ala Thr Pro Ala Ala Ser Thr Val Ala Pro Ser Gln Ala
290 295 300

Pro Leu Ala Leu Ala His Asp Gly Val Tyr Leu Pro Lys Asp Ala Phe
305 310 315

```
Phe Ser Leu Ile Gly Ala Ser Arg Pro Leu Ala Glu Ala Ala Gly Ala
        320                 325                 330

Arg Ala Ala Tyr Pro Ala Val Pro Pro Pro Pro Ala Tyr Pro Val Met
    335                 340                 345

Asn Tyr Glu Asp Pro Ser Ser Arg His Phe Asp Tyr Ser Ala Trp Leu
350                 355                 360                 365

Arg Arg Pro Ala Tyr Asp Ala Val Pro Pro Leu Pro Pro Pro Pro Val
                370                 375                 380

Met Pro Met Pro Tyr Arg Arg Arg Asp Pro Met Met Glu Glu Ala Glu
            385                 390                 395

Arg Ala Ala Trp Glu Arg Gly Tyr Ala Pro Ser Ala Tyr Asp His Tyr
        400                 405                 410

Val Asn Asn Gly Ser Trp Ser Arg Ser Arg Ser Gly Ala Leu Lys Arg
    415                 420                 425

Arg Arg Glu Arg Asp Ala Ser Ser Asp Glu Glu Glu Asp Met Ser Phe
430                 435                 440                 445

Pro Gly Glu Ala Asp His Gly Lys Ala Arg Lys Arg Leu Lys Ala His
                450                 455                 460

His Gly Arg Asp Asn Asn Asn Ser Gly Ser Asp Ala Lys Gly Asp Arg
            465                 470                 475

Tyr Asp Asp Ile Arg Glu Ala Leu Gln Glu Leu Lys Arg Glu Met Leu
        480                 485                 490

Ala Val Arg Gln Ile Ala Pro Arg Ala Leu Leu Ala Pro Ala Gln Leu
    495                 500                 505

Ala Thr Pro Val Ala Ser Pro Thr Thr Thr Thr Ser His Gln Ala Glu
510                 515                 520                 525

Ala Ser Glu Pro Gln Ala Ser Thr Ala Ala Ala Ala Ser Pro Ser Thr
                530                 535                 540

Ala Ser Ser His Gly Ser Lys Ser Ala Glu Arg Gly Val Val Asn Ala
            545                 550                 555

Ser Cys Arg Val Ala Pro Pro Leu Glu Ala Val Asn Pro Pro Lys Asp
        560                 565                 570

Met Val Asp Leu Asn Arg Arg Leu Phe Val Ala Ala Leu Asn Lys Met
    575                 580                 585

Glu
590
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Simian cytomegalovirus
(B) STRAIN: Colburn

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ser Lys Ser Ala Glu Arg Gly Val Val Asn Ala Ser Cys Arg Val Ala
1 5 10 15

Pro Pro

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human cytomegalovirus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ala Glu Arg Ala Gln Ala Gly Val Val Asn Ala Ser Cys Arg Leu Ala
1 5 10 15

Thr Ala

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Herpes simplex virus type 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ser Asn Ala Glu Ala Gly Ala Leu Val Asn Ala Ser Ser Ala Ala His
1 5 10 15

Val Asp

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Varicella-zoster virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

His Thr Asp Thr Val Gly Gln Asp Val Asn Ala Val Glu Ala Ser Ser
1 5 10 15

Lys Ala

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly His His Arg Gly Lys Lys Leu Val Gln Ala Ser Ala Ser Gly Val
1 5 10 15

Ala Gln

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Infectious Laryngotracheitis Virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Asn Gln Glu Ser Ala Arg Glu Thr Val Asp Ala Ser Met Pro Lys Arg
1 5 10 15

Leu Lys

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human Herpes Virus 6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ala Ala Ser Pro Lys Pro Ser Ile Leu Asn Ala Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Simian cytomegalovirus
(B) STRAIN: Colburn

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Thr Ala Arg Glu Ser Tyr Val Lys Ala Ser Val Ser Pro Ala Glu
1 5 10 15

Gln Glu Thr Cys
20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human cytomegalovirus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Thr Glu Arg Glu Ser Tyr Val Lys Ala Ser Val Ser Pro Glu Ala
1 5 10 15

Arg Ala Ile Leu
20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Herpes simplex virus type 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met
1 5 10 15

Trp Gly Ala Glu
20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Varicella-zoster virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Ile Met Gly His Val Tyr Leu Gln Ala Ser Thr Gly Tyr Gly Leu
1 5 10 15

Ala Arg Ile Thr
20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Asn Ile Pro Ala Glu Ser Tyr Leu Lys Ala Ser Asp Ala Pro Asp Leu
1 5 10 15

Gln Lys Pro Asp
20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Infectious Laryngotracheitis Virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Val Tyr Asn Pro Lys Tyr Leu Gln Ala Asn Glu Val Ile Thr Ile
1 5 10 15

Gly Ile Lys Glu
20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Simian cytomegalovirus
(B) STRAIN: Colburn

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Pro Leu Pro Leu Asn Val Asn His Asp Glu Ser Ala Thr Val Gly Tyr
1 5 10 15

Val

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Simian cytomegalovirus
(B) STRAIN: Colburn

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Phe Lys His Val Ala Leu Cys Ser Val Gly Arg Arg Arg Gly Thr Leu
1 5 10 15

Ala Val Tyr Gly
20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human cytomegalovirus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Leu Pro Leu Asn Ile Asn His Asp Asp Thr Ala Val Val Gly His
1 5 10 15

Val

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human cytomegalovirus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Phe Lys His Val Ala Leu Cys Ser Val Gly Arg Arg Arg Gly Thr Leu
1 5 10 15

Ala Val Tyr Gly
20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Herpes simplex virus type 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Pro Leu Pro Ile Asn Val Asp His Arg Ala Gly Cys Glu Val Gly Arg
1 5 10 15

Val

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Herpes simplex virus type 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Phe Ala His Val Ala Leu Cys Ala Ile Gly Arg Arg Leu Gly Thr Ile
1 5 10 15

Val Thr Tyr Asp
20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Varicella-zoster virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Lys Ile Pro Ile Asn Ile Asp His Arg Lys Asp Cys Val Val Gly Glu
1 5 10 15

Val

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Varicella-zoster virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Phe Thr His Val Ala Leu Cys Val Val Gly Arg Arg Val Gly Thr Val
1 5 10 15

Val Asn Tyr Asp
20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Pro Leu Pro Leu Thr Val Glu His Leu Pro Asp Ala Pro Val Gly Ser
1 5 10 15

Val

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Phe Asp His Val Ser Ile Cys Ala Leu Gly Arg Arg Arg Gly Thr Thr
1 5 10 15

Ala Val Tyr Gly
20

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Infectious Laryngotracheitis Virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Thr Ile Pro Ile Asn Ile Asp His Glu Ser Ser Cys Val Val Gly Thr
1 5 10 15

Val

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Infectious Laryngotracheitis Virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Phe Ala His Val Ala Leu Cys Glu Leu Gly Arg Arg Glu Gly Thr Val
1 5 10 15

Ala Ile Tyr Gly
20

CLAIMS

1. A preparation of a proteinase encoded by a herpes virus, said proteinase able to cleave herpes virus assembly protein precursor to form assembly protein, said preparation being free of intact infectious virion DNA of the herpes virus.

2. A substrate for cleavage by a herpes virus proteinase, said substrate comprising a polypeptide containing an amino acid sequence:

$$aa_1\text{-}aa_2\text{-Ala-}aa_3,$$

wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, and aa_3 is Ser, Val, or Asn, wherein said proteinase cleaves said substrate on the carboxy terminal side of the Ala residue, and wherein said substrate is substantially free of assembly protein precursor of the herpes virus.

3. The substrate of claim 2 wherein said polypeptide contains the amino acid sequence Val-Asn-Ala-Ser.

4. The substrate of claim 2 wherein said polypeptide is linked to a fluorescent moiety and a quenching moiety, said moieties being separated by said polypeptide.

5. The substrate of claim 2 wherein said polypeptide is labeled with a radioactive moiety.

6. The substrate of claim 2 wherein an enzyme is linked to said polypeptide, said polypeptide sterically inhibiting said enzyme.

7. The substrate of claim 6 wherein a substrate for said enzyme is linked to the polypeptide.

8. The substrate of claim 2 wherein the amino acid sequence contained in said peptide is within an indicator molecule.

9. The substrate of claim 8 wherein said indicator molecule is β-galactosidase or luciferase.

10. A substrate for cleavage by a herpes virus proteinase, said substrate comprising a polypeptide containing an amino acid sequence

$$Tyr\text{-}aa_4\text{-}aa_5\text{-}Ala\text{-}aa_6$$

wherein aa_4 is Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn, wherein said proteinase cleaves said substrate on the carboxy terminal side of the Ala residue, and wherein said substrate is substantially free of primary translation product of the gene encoding said proteinase.

11. The substrate of claim 10 wherein aa_4 is Leu, aa_5 is Lys or Gln and aa_6 is Ser.

12. The substrate of claim 10 wherein said polypeptide contains the amino acid sequence Tyr-Val-Lys-Ala-Ser.

13. The substrate of claim 10 wherein said polypeptide is linked to a fluorescent moiety and a quenching moiety, said moieties being separated by said polypeptide.

14. The substrate of claim 10 wherein said polypeptide is labeled with a radioactive moiety.

15. The substrate of claim 10 wherein an enzyme is linked to said polypeptide, said polypeptide sterically inhibiting said enzyme.

16. The substrate of claim 15 wherein a substrate for said enzyme is linked to the polypeptide.

17. The substrate of claim 10 wherein the amino acid sequence contained in said peptide is within an indicator molecule.

18. The substrate of claim 17 wherein said indicator molecule is β-galactosidase or luciferase.

19. A kit for measuring activity of a herpes virus proteinase, comprising:

a proteinase encoded by a herpes virus; and

a substrate for cleavage by said proteinase, said substrate comprising a polypeptide containing an amino acid sequence selected from the group consisting of aa_1-aa_2-Ala-aa_3 and Tyr-aa_4-aa_5-Ala-aa_6, wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, aa_3 is Ser, Val, or Asn, aa_4 is Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn, wherein said proteinase cleaves said substrate on the carboxy terminal side of the Ala residue, said kit being substantially free of intact infectious herpes virus.

20. The kit of claim 20 wherein said substrate comprises a polypeptide containing the sequence aa_1-aa_2-Ala-aa_3.

21. The kit of claim 20 wherein said polypeptide contains the amino acid sequence Val-Asn-Ala-Ser.

22. The kit of claim 20 wherein said substrate comprises a polypeptide containing the sequence Tyr-aa_4-aa_5-Ala-aa_6.

23. The kit of claim 22 wherein aa_4 is Leu, aa_5 is Lys or Gln and aa_6 is Ser.

24. The kit of claim 22 wherein said polypeptide contains the amino acid sequence Tyr-Val-Lys-Ala-Ser.

25. The kit of claim 19 wherein said proteinase and said substrate for cleavage are co-expressed in a mammalian cell which is free of a herpes virus infection, and cleavage of said substrate occurs in the mammalian cell.

26. The kit of claim 19 wherein said proteinase and said substrate contacted in vitro.

27. The kit of claim 19 wherein said proteinase and said substrate are expressed in vitro.

28. The kit of claim 19 wherein said proteinase and said substrate are expressed in separate cells.

29. The kit of claim 19 wherein said proteinase and said substrate are expressed in mammalian cells.

30. The kit of claim 19 wherein said proteinase and said substrate are expressed in bacterial cells.

31. The kit of claim 19 wherein said proteinase and said substrate are expressed in yeast cells.

32. The kit of claim 19 wherein said proteinase and said substrate are expressed in insect cells.

33. The kit of claim 19 wherein said proteinase and said substrate are expressed chemically, or by in vitro transcription and translation.

34. A method for measuring activity of a herpes virus proteinase, comprising the steps of:

contacting a proteinase encoded by a herpes virus with a substrate for cleavage by said proteinase, said substrate comprising a polypeptide containing an amino acid sequence selected from the group consisting of aa_1-aa_2-Ala-aa_3 and Tyr-aa_4-aa_5-Ala-aa_6, wherein aa_1 is Val or Leu, aa_2 is a polar amino acid, aa_3 is Ser, Val, or Asn, aa_4 is Val or Leu, aa_5 is Lys or Gln and aa_6 is Ser or Asn, wherein said proteinase cleaves said substrate on the carboxy terminal side of the Ala residue, said step of contacting occurring in the absence of a intact infectious herpes virus virion DNA; and

monitoring cleavage of said substrate.

35. The method of claim 34 wherein the step of monitoring comprises determining a change in size of said substrate.

36. The method of claim 35 wherein said determining is done by observing an altered migration of said substrate on an electrophoretic gel matrix.

37. The method of claim 35 wherein said determining is done by observing an altered mobility of said substrate on a chromatography medium.

38. The method of claim 34 wherein the step of monitoring comprises monitoring a change in fluorescence of said substrate.

39. The method of claim 34 wherein the substrate is attached to a radiolabelled moiety.

40. The method of claim 39 wherein the step of monitoring comprises determining a change in solvent extractability of the radiolabelled moiety.

41. The method of claim 34 wherein the substrate is attached to an enzyme and the step of monitoring comprises determining a colorimetric change.

42. The method of claim 34 further comprising the step of adding a test substance to said proteinase and substrate to determine the level of inhibition of the activity of said proteinase caused by said test substance.

43. A recombinant DNA molecule which encodes at least a portion of a herpes virus proteinase, said portion having the ability to cleave a herpes virus assembly protein precursor.

44. The DNA molecule of claim 43 wherein the portion of the proteinase comprises amino acids numbered 1 through 249 of simian cytomegalovirus Colburn APNG1 protein or the homologous amino acids of another herpes virus.

45. The DNA molecule of claim 43 wherein the portion of the proteinase comprises amino acids numbered 1 through 280 simian cytomegalovirus Colburn

APNG1 protein or the homologous amino acids of another herpes virus.

46. The DNA molecule of claim 43 wherein the portion of the proteinase comprises the first 234 to 262 amino acids of the primary translation product of a homolog of the APNG1 gene of similar cytomegalovirus Colburn.

47. The DNA molecule of claim 43 wherein the portion of the proteinase lacks amino acids which comprise a herpes virus assembly protein precursor.

48. The DNA molecule of claim 43 which is capable of expressing said portion of the herpes virus proteinase in a host cell.

49. An inhibitor of a herpes virus proteinase comprising a derivative of the substrate of claim 2, said inhibitor differing from said substrate in the scissile peptide bond which is carboxyl to the Ala residue.

50. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a hydroxyethylamine linkage.

51. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a phosphonamide linkage.

52. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a carbon fluoride aldehyde.

53. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a dialcohol linkage.

54. An inhibitor of a herpes virus proteinase comprising a derivative of the substrate of claim 10, said inhibitor differing from said substrate in the scissile peptide

bond which is carboxyl to the Ala residue.

55. The inhibitor of claim 54 wherein the scissile peptide bond is replaced by a hydroxyethylamine linkage.

56. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a phosphonamide linkage.

57. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a carbon fluoride aldehyde.

58. The inhibitor of claim 49 wherein the scissile peptide bond is replaced by a dialcohol linkage.

FIG. 1A

```
                   20                  40                  60
5'-TTGTCCGACACCCCCAGGTTATTGGTGGTCTCGCGGGGGGGAACAGGGGGGTTTGCAGG

                   80                 100                 120
CCTCGGTTAAAGAGCAGCACGCAGATGAGTCTCAAGATCTTGAGTTCTTCCAGCCGCAGG

                  140                 160                 180
GTGTTGAGCGGCTGTCCCCGCGACATCTTTTCGCTGATCTGTAATATTAGATGATTGGCA
                                                   EndLeuAla>
                  200                 220       APNG1→ 240
CAAGTAAAGGAGAATTTGCCGGTTCGAACCCGGCCTCCTCCGTGTTGGACATGGCCGAT
GlnValLysGluAsnLeuProValArgThrArgAlaSerSerValLeuAspMETAlaAsp>
                  260                 280                 300
CCCGTCTACGTCGGGGGTTTTTTGGTGCGCTACGACGAGCCTCCCGGAGAAGCTGAGCTG
ProValTyrValGlyGlyPheLeuValArgTyrAspGluProProGlyGluAlaGluLeu>
                  320                 340         5'End   360
TTTCTGCCCTCGGGGGTGGTAGACCGCTGGTTGCGCGATTGCCGAGGCCCGCTGCCCCTG
PheLeuProSerGlyValValAspArgTrpLeuArgAspCysArgGlyProLeuProLeu>
                  380                 400                 420
AATGTCAATCACGACGAGTCGGCGACCGTGGGCTATGTGGCTGGGCTCCAGAATGTCCGG
AsnValAsnHisAspGluSerAlaThrValGlyTyrValAlaGlyLeuGlnAsnValArg>
                  440                 460                 480
GCCGGCTTGTTCTGTTTGGGACGTGTTACGTCCCCCAAGTTTCTGGATATCGTTCAAAAA
AlaGlyLeuPheCysLeuGlyArgValThrSerProLysPheLeuAspIleValGlnLys>
```

FIG. 1B

```
                500                 520                 540
GCCTCGGAAAAATCCGAGTTGGTGTCCCGGGGACCTCCGTCCGAGTCCTCGTTGCGGCCG
AlaSerGluLysSerGluLeuValSerArgGlyProProSerGluSerSerLeuArgPro>
                560                 580                 600
GACGGCGTGTTGGAGTTCTCAGCGGCAGTTATTCGGGCCTGTCGCTCTCCAGCCGCCGA
AspGlyValLeuGluPheLeuSerGlyTyrSerGlyLeuSerLeuSerSerArgArg>
                620                 640                 660
GATATAAACGCGGCCGCCGCGGGCGATGCAGAAACAGCGTGCTTCAAACATGTG
AspIleAsnAlaAlaAlaGlyAlaAspAlaGluThrAlaCysPheLysHisVal>
                680                 700                 720
GCTCTGTGCAGCGTGGGCCGCCGCGGGGCACGTTGGCGGTGTATGGCAGGCCAGCCAGAT
AlaLeuCysSerValGlyArgArgArgGlyThrLeuAlaValTyrGlyArgGlnProAsp>
                740                 760                 780
TGGGTGATGGAACGTTTCCCGGATCTCACCGAGGCCGACCGGGAAGCGCGCTGCGAAATCAG
TrpValMETGluArgPheProAspLeuThrGluAlaAspArgGluAlaLeuArgAsnGln>
                800                 820                 840
CTATCGGGAAGTGGGGAAGTTGCCGCGAAGGAAAGTGCGGAATCGTCTGCCGCCGCCGCC
LeuSerGlySerGlyGluValAlaAlaLysGluSerAlaGluSerSerAlaAlaAlaAla>
      5' Mid        860                 880                 900
GTCGATCCCTTTCAGTCGGATTCGTACGGGCTGTTGGGGAACAGTGTGGACGCGCTGTAC
ValAspProPheGlnSerAspSerTyrGlyLeuLeuGlyAsnSerValAspAlaLeuTyr>
                920                 940                 960
ATTCAAGAGCGTCTCCCTAAGCTGCGCTATGACAAGCGGCTGGTCGGGGTCACGGCTCGG
IleGlnGluArgLeuProLysLeuArgTyrAspLysArgLeuValGlyValThrAlaArg>
                980                 1000                1020
GAGTCGTACGTGAAAGCCAGTGTTTCGCCCCGCCGAGGAGACGTGCGATATAAAGTA
GluSerTyrValLysAlaSerValSerProAlaGluGlnGluThrCysAspIleLysVal>
```

APNG.7→

FIG. IC

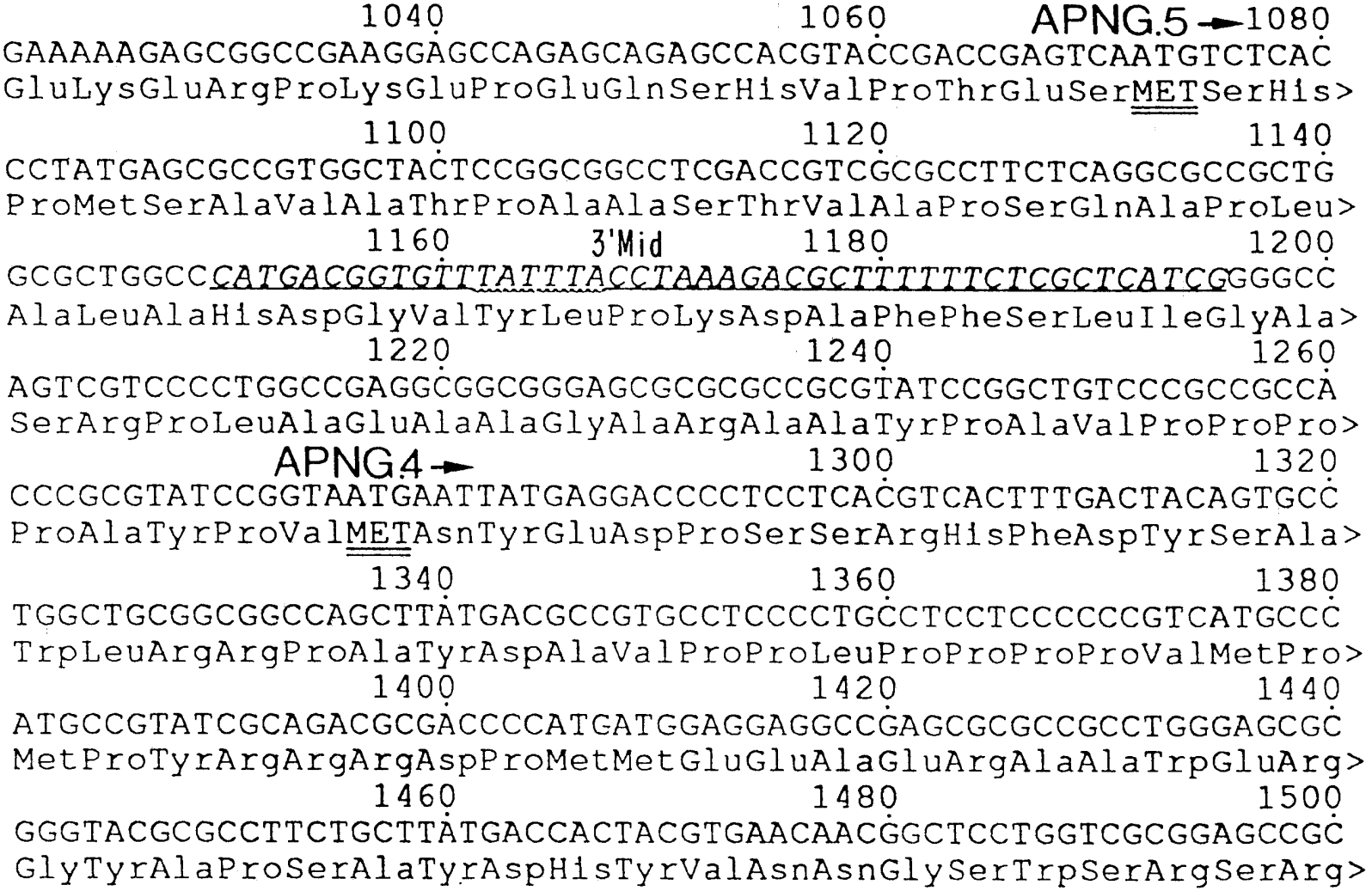

1040
1060
APNG.5
1080
GAAAAAGAGCGGCCGAAGGAGCCAGAGCAGAGCCACGTACCGACCGAGTCAATGTCTCAC
GluLysGluArgProLysGluProGluGlnSerHisValProThrGluSerMETSerHis>
1100
1120
1140
CCTATGAGCGCCGTGGCTACTCCGGCGGCCTCGACCGTCGCGCCTTCTCAGGCGCCGCTG
ProMetSerAlaValAlaThrProAlaAlaSerThrValAlaProSerGlnAlaProLeu>
1160
3'Mid
1180
1200
GCGCTGGCCCATGACGGTGTTTATTTACCTAAAGACGCTTTTTTCTCGCTCATCGGGGCC
AlaLeuAlaHisAspGlyValTyrLeuProLysAspAlaPhePheSerLeuIleGlyAla>
1220
1240
1260
AGTCGTCCCCTGGCCGAGGCGGCGGGAGCGCGCGCCGCGTATCCGGCTGTCCCGCCGCCA
SerArgProLeuAlaGluAlaAlaGlyAlaArgAlaAlaTyrProAlaValProProPro>
APNG4
1300
1320
CCCGCGTATCCGGTAATGAATTATGAGGACCCCTCCTCACGTCACTTTGACTACAGTGCC
ProAlaTyrProValMETAsnTyrGluAspProSerSerArgHisPheAspTyrSerAla>
1340
1360
1380
TGGCTGCGGCGGCCAGCTTATGACGCCGTGCCTCCCCTGCCTCCTCCCCCCGTCATGCCC
TrpLeuArgArgProAlaTyrAspAlaValProProLeuProProProProValMetPro>
1400
1420
1440
ATGCCGTATCGCAGACGCGACCCCATGATGGAGGAGGCCGAGCGCGCCGCCTGGGAGCGC
MetProTyrArgArgArgAspProMetMetGluGluAlaGluArgAlaAlaTrpGluArg>
1460
1480
1500
GGGTACGCGCCTTCTGCTTATGACCACTACGTGAACAACGGCTCCTGGTCGCGGAGCCGC
GlyTyrAlaProSerAlaTyrAspHisTyrValAsnAsnGlySerTrpSerArgSerArg>

FIG. ID

```
                1520                    1540                    1560
AGCGGCGCGCTCAAGAGGCGAAGGGAGCGCGACGCGTCCTCGGATGAGGAAGAGGACATG
SerGlyAlaLeuLysArgArgArgGluArgAspAlaSerSerAspGluGluGluAspMet>
                1580                    1600                    1620
AGTTTTCCCGGGGAAGCCGACCACGGCAAGGCTCGGAAAAGACTCAAAGCTCATCACGGG
SerPheProGlyGluAlaAspHisGlyLysAlaArgLysArgLeuLysAlaHisHisGly>
                1640                    1660                    1680
CGTGATAATAACAACTCTGGGAGCGATGCCAAGGGCGATCGGTACGACGACATTCGGGAA
ArgAspAsnAsnAsnSerGlySerAspAlaLysGlyAspArgTyrAspAspIleArgGlu>
                1700                    1720                    1740
GCGTTACAGGAGCTGAAGCGCGAGATGCTGGCCGTGCGGCAGATCGCGCCACGTGCGCTC
AlaLeuGlnGluLeuLysArgGluMetLeuAlaValArgGlnIleAlaProArgAlaLeu>
                1760                    1780                    1800
TTGGCCCCCGCACAGCTAGCGACGCCCGTGGCTTCTCCGACAACGACCACGTCGCATCAA
LeuAlaProAlaGlnLeuAlaThrProValAlaSerProThrThrThrThrSerHisGln>
                1820                    1840                    1860
GCCGAGGCTAGCGAACCTCAGGCATCGACTGCCGCTGCCGCGTCGCCGTCAACCGCTTCG
AlaGluAlaSerGluProGlnAlaSerThrAlaAlaAlaAlaSerProSerThrAlaSer>
                1880                    1900                    1920
TCGCACGGCAGCAAGTCGGCCGAACGCGGGGTGGTGAACGCCTCGTGTCGCGTTGCGCCT
SerHisGlySerLysSerAlaGluArgGlyValValAsnAlaSerCysArgValAlaPro>
                1940                    1960  3'End             1980
CCGTTGGAGGCTGTGAACCCCCCTAAG GACATGGTGGACTTGAATCGTCGCCTGTTT GTG
ProLeuGluAlaValAsnProProLysAspMetValAspLeuAsnArgArgLeuPheVal>
                2000
GCGGCGTTGAATAAAATGGAATAAAAACTCGTAC-3'
AlaAlaLeuAsnLysMetGluEnd
```

FIG. 2

HERPESVIRUS	CONSERVED REGION	
SIMIAN CMV (COLBURN)	PLPLNVNHDESATVGYV	FKHVALCSVGRRRGTLAVYG
HUMAN CMV (AD169)	ALPLNI NHDDTAVVGHV	FKHVALCSVGRRRGTLAVYG
HSV-1	PLPINVDHRAGCEVGRV	FAHVALCA I GRR LGT IVTYD
VZV	KI PI NI DHRKDCVVGEV	FTHVALCVVGRRVGTVVNYD
EBV	PLPLTVEH LPDAPVGSV	FDHVSI CALGRRRGTTAVYG
ILTV	TI PI NI DH ESSCVVGTV	FAHVALCELGRREGTVAI YG
	CONSERVED MOTIF 2	CONSERVED MOTIF 1

FIG. 3A

VIRUS	RECOGNITION/CLEAVAGE DOMAIN
SCMV	SKSAERGVVNA ↓ SCRVAPP
HCMV	AERAQAGVVNA ↓ SCRLATA
HSV-1	SNAEAG ALVNA ↓ SSAAHVD
VZV	HTDTVGQDVNA ↓ VEASSKA
EBV	GHHRGKKLVQA ↓ SASGVAQ
ILTV	NQESARETVDA ↓ SMPKRLK
HHV-6	AA SPKPS I LNA ↓ S - - - - - -

FIG. 3B

COLBURN:	VTARESYVKA ↓ SVSPAEQETC
HCMV AD169:	VTERESYVKA ↓ SVSPEARAI L
HSV-1:	GIAGHTYLQA ↓ SEKFKMWGAE
VZV:	GIMGHVYLQA ↓ STGYGLAR I T
EBV:	NI PAESYLK A ↓ SDAPDLQKPD
ILTV:	AVYNPKYLQA ↓ NEVI TI GI K E

FIG. 4

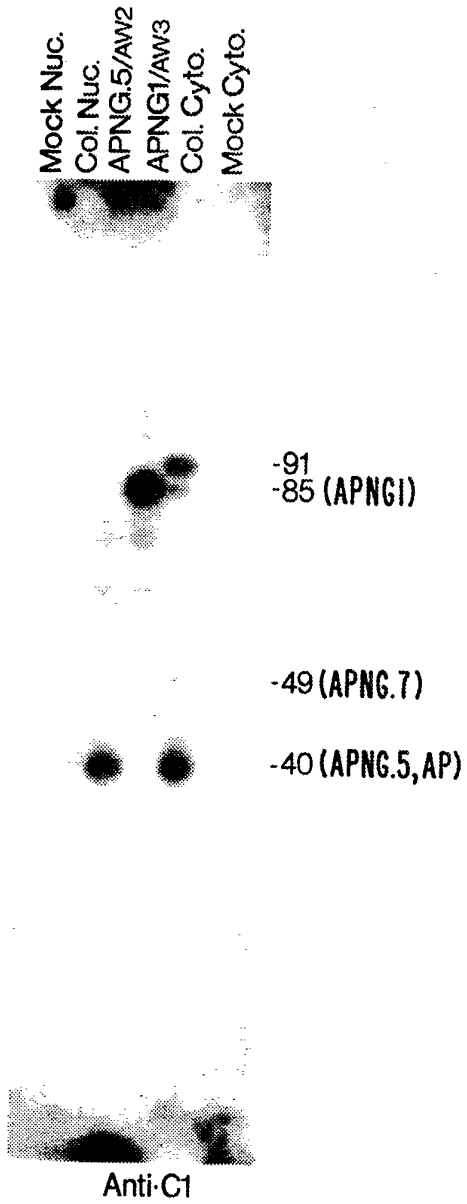
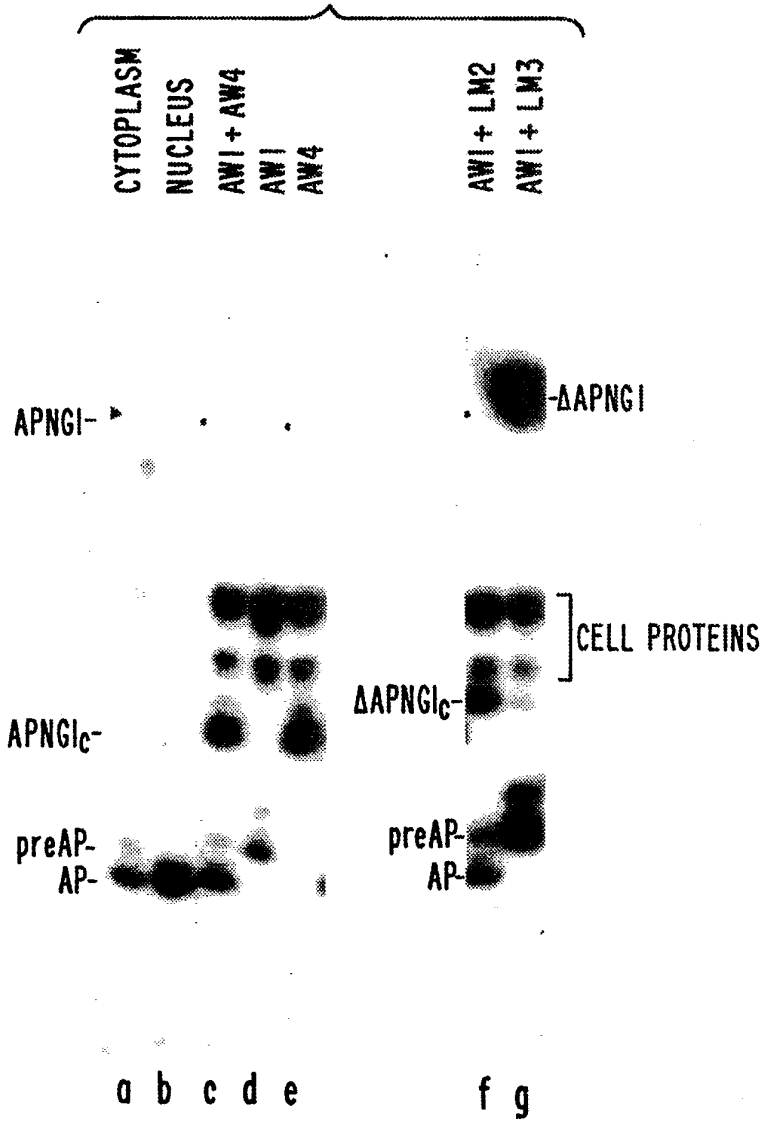

Mock Nuc.
Col. Nuc.
APNG.5/AW2
APNG1/AW3
Col. Cyto.
Mock Cyto.
-91
-85 (APNG1)
-49 (APNG.7)
-40 (APNG.5, AP)
Anti-C1

FIG. 5


CYTOPLASM
NUCLEUS
AW1 + AW4
AW1
AW4
AW1 + LM2
AW1 + LM3
APNGI-
-ΔAPNGI
APNGIc-
ΔAPNGIc-
CELL PROTEINS
preAP-
AP-
preAP-
AP-
a b c d e
f g

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 92/05513

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/57; C12N9/50; C07K5/10; C12Q1/37
C07K5/02; C07K13/00; C07K7/02; C07K7/06

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C12N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category °	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	JOURNAL OF VIROLOGY vol. 65, no. 1, January 1991, pages 206 - 212 F.LIU AND B. ROIZMAN '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' see the whole document	1,43
Y	---	44-48
O,X	J. CELL BIOCHEM SUPPL O vol. 15, no. G, April 1991, page 138 A. WELCH AND W. GIBSON 'Proteolytic processing of the cytomegalovirus assembly protein: Is a viral gene responsible ?' see abstract	1,43-48

	-/--	

° Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
26 OCTOBER 1992	02. 11. 92
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	VAN DER SCHAAL C.A.

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
X	JOURNAL OF VIROLOGY vol. 64, no. 3, March 1990, pages 1241 - 1249 W. GIBSON ET AL 'Identification of precursor to cytomegalovirus capsid assembly protein and evidence that processing results in loss of its carboxy-terminal end'	1-3,43
Y	see abstract see page 1247, right column, paragraph 2 see page 1248, right column, last paragraph	4-9, 44-51

O,X	JOURNAL OF VIROLOGY vol. 65, no. 8, August 1991, pages 4091 - 4100 A. WELCH ET AL 'Cytomegalovirus assembly protein nested gene family' see the whole document Presented at the 9th meeting of the American Society for Virology, 8 to 12 July 1990, the Federation of American Societies for Experimental Biology Meeting on Viral Assembly, 15 to 20 July 1990 and the 15th International Herpesvirus Workshop, 2 to 8 August 1990. see page 4091, right column, paragraph 3	1,43-48

P,X	PROCEEDINGS OF THE NATIONAL ACADAMY OF SCIENCE vol. 88, December 1991, pages 10792 - 10796 A.WELCH ET AL 'A Herpesvirus maturational proteinase, assemblin: Identification of its gene, putative active site domain,and cleavage site'	1-9, 43-48
Y	see the whole document	19-42

P,X	JOURNAL OF VIROLOGY vol. 65, no. 10, October 1991, pages 5149 - 5156 F.LIU AND B. ROIZMAN 'The Herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate'	1,43
Y	see the whole document	44-48

-/--

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	H.U. BERGMEYER 'Methods in enzymatic analysis. Volume V, Enzymes 3: Peptidases, proteinases and their inhibitors' 1984 , VERLAG CHEMIE , WEINHEIM see page 84, paragraph 2 ---	4-9, 19-42
X	NUCLEIC ACIDS RESEARCH. vol. 18, no. 23, 1990, ARLINGTON, VIRGINIA US page 7159 S. OHAGI ET AL 'Sequence of cDNA encoding human LPR (leukocyte common antigen-related peptide)' see figure 1 ---	2-9
X	MOLECULAR AND GENERAL GENETICS vol. 227, June 1991, BERLIN DE pages 318 - 329 M. FLING ET AL 'Analysis of a Candida albicans gene that encodes as novel mechanism for resistance to benomyl and methotrexate' see figure 3 ---	10-18
Y	JOURNAL OF MEDICINAL CHEMISTRY vol. 33, 1990, WASHINGTON US pages 1285 - 1288 D.RICH ET AL 'Hydroxyethylamine analogues of the p17/p24 substrate cleavage site are tight-binding inhibitors of HIV protease' see the whole document ---	49,50
Y	BIOCHEMISTRY. vol. 28, 1989, EASTON, PA US pages 4949 - 4951 D. GROBELNY ET AL 'Binding energetics of phosphorus-containing inhibitors of thermolysin' see abstract see page 4949, left column, paragraph 1 -----	49,51