

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number
WO 02/00891 A2

(51) International Patent Classification⁷: **C12N 15/57**,
9/64, C12Q 1/68, G01N 33/50, A61K 38/48, C12N 15/62

(21) International Application Number: PCT/EP01/07000

(22) International Filing Date: 21 June 2001 (21.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/214,010 26 June 2000 (26.06.2000) US

(71) Applicant (for all designated States except US): **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **RAMAKRISHNAN, Shyam** [IN/US]; 76 Euston Road, Apt. #10, Brighton, MA 02135 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REGULATION OF HUMAN NAPSIN-LIKE ASPARTYL PROTEASE

(57) Abstract: Reagents which regulate human napsin-like aspartyl protease activity and reagents which bind to human napsin-like aspartyl protease gene products can be used to treat cancer and osteoporosis.



WO 02/00891 A2

REGULATION OF HUMAN NAPSIN-LIKE ASPARTYL PROTEASE

TECHNICAL FIELD OF THE INVENTION

5

The invention relates to the area of regulation human napsin-like aspartyl protease activity to provide therapeutic effects.

BACKGROUND OF THE INVENTION

10

Metastasizing cancer cells invade the extracellular matrix using plasma membrane protrusions that contact and dissolve the matrix with proteases. Agents which inhibit such protease activity can be used to suppress metastases. Proteases also are expressed during development, when degradation of the extracellular matrix is desired. In cases where appropriate extracellular matrix degradation does not occur, supplying a molecule with a protease activity can provide the necessary enzymatic activity. Thus, there is a need in the art for identifying new proteases and methods of regulating extracellular matrix degradation.

15

SUMMARY OF THE INVENTION

20

It is an object of the invention to provide reagents and methods of regulating human napsin-like aspartyl protease. These and other objects of the invention are provided by one or more of the embodiments described below.

25

One embodiment of the invention is a napsin-like aspartyl protease polypeptide comprising an amino acid sequence selected from the group consisting of: amino acid sequences which are at least about 60% identical to the amino acid sequence shown in SEQ ID NO: 2; and the amino acid sequence shown in SEQ ID NO: 2.

30

- 2 -

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a napsin-like aspartyl protease polypeptide comprising an amino acid sequence selected from the group consisting of:

- 5 amino acid sequences which are at least about 60% identical to the amino acid sequence shown in SEQ ID NO: 2; and
the amino acid sequence shown in SEQ ID NO: 2.

10 Binding between the test compound and the napsin-like aspartyl protease polypeptide is detected. A test compound which binds to the napsin-like aspartyl protease polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the napsin-like aspartyl protease.

15 Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a napsin-like aspartyl protease polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

- 20 nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and
the nucleotide sequence shown in SEQ ID NO: 1.

25 Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the napsin-like aspartyl protease through interacting with the napsin-like aspartyl protease mRNA.

30 Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a

- 3 -

napsin-like aspartyl protease polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 60% identical to the amino acid sequence shown in SEQ ID NO: 2; and

5 the amino acid sequence shown in SEQ ID NO: 2.

A napsin-like aspartyl protease activity of the polypeptide is detected. A test compound which increases napsin-like aspartyl protease activity of the polypeptide relative to napsin-like aspartyl protease activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases napsin-like aspartyl protease activity of the polypeptide relative to napsin-like aspartyl protease activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

15

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a napsin-like aspartyl protease product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

20 nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

Binding of the test compound to the napsin-like aspartyl protease product is detected. A test compound which binds to the napsin-like aspartyl protease product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

25

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a napsin-like aspartyl protease polypeptide or the product

30

- 4 -

encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

5 the nucleotide sequence shown in SEQ ID NO: 1.

Napsin-like aspartyl protease activity in the cell is thereby decreased.

10 The invention thus provides reagents and methods for regulating human napsin-like aspartyl protease activity which can be used *inter alia*, to suppress metastatic activity of malignant cells and to enhance extracellular matrix degradation during development.

BRIEF DESCRIPTION OF THE DRAWINGS

15

Fig. 1 shows the DNA-sequence encoding a napsin-like aspartyl protease polypeptide (SEQ ID NO:1).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO:2).

20 Fig. 3 shows the amino acid sequence of the protein identified by SwissProt Accession No. O96009 (SEQ ID NO:3).

Fig. 4 shows the BLASTX alignment of napsin-like aspartyl protease as shown in SEQ ID NO:2 with the protein identified by SwissProt Accession No. O96009 (SEQ ID NO:3).

25

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide encoding a napsin-like aspartyl protease polypeptide and being selected from the group consisting of:

30 a. a polynucleotide encoding a napsin-like aspartyl protease polypeptide comprising an amino acid sequence selected from the group consisting of:

- 5 -

- amino acid sequences which are at least about 60% identical to the amino acid sequence shown in SEQ ID NO: 2; and the amino acid sequence shown in SEQ ID NO: 2.
- b) a polynucleotide comprising the sequence of SEQ ID NO: 1;
- 5 c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);
- d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and
- 10 e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that regulators of a napsin-like aspartyl protease, particularly a human napsin-like aspartyl protease, can be used to regulate degradation of the extracellular matrix. Human napsin-like aspartyl protease as shown in SEQ ID NO:2 is at least 52% identical to the human protein identified by SwissProt Accession No. O96009 (SEQ ID NO:3) and annotated as a napsin 1 precursor (FIG. 1). Human napsin-like aspartyl protease is therefore expected to be useful for the same purposes as previously identified serine proteases.

15

20

Polypeptides

Napsin-like aspartyl protease polypeptides according to the invention comprise at least 10, 15, 25, 50, 75, 100, 125, or 140 contiguous amino acids selected from SEQ ID NO:2 or from a biologically active variant thereof, as defined below. An napsin-like aspartyl protease polypeptide of the invention therefore can be a portion of an napsin-like aspartyl protease molecule, a full-length napsin-like aspartyl protease molecule, or a fusion protein comprising all or a portion of an napsin-like aspartyl protease molecule.

25

30

Biologically Active Variants

Napsin-like aspartyl protease variants which are biologically active, *i.e.*, retain an napsin-like aspartyl protease activity, also are napsin-like aspartyl protease polypeptides. Preferably, naturally or non-naturally occurring napsin-like aspartyl protease variants have amino acid sequences which are at least about 50, 55, 60, 65, 70, more preferably about 75, 80, 85, 90, 96, or 98% identical to the amino acid sequence shown in SEQ ID NO:2. Percent identity between a putative napsin-like aspartyl protease variant and an amino acid sequence of SEQ ID NO:2 is determined using the Blast2 alignment program (Blosum62, Expect 10, standard genetic codes).

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active napsin-like aspartyl protease polypeptide can readily be determined by assaying for fibronectin binding or for napsin-like aspartyl protease activity, as is known in the art and described, for example, in Example 2.

Fusion Proteins

Fusion proteins are useful for generating antibodies against napsin-like aspartyl protease amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of an napsin-like aspartyl protease polypeptide, including its active site and fibronectin

domains. Methods such as protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

5

An napsin-like aspartyl protease fusion protein comprises two protein segments fused together by means of a peptide bond. For example, the first protein segment can comprise at least 10, 15, 25, 50, 75, 100, 125, or 140 contiguous amino acids selected from SEQ ID NO:2 or a biologically active variant thereof.

10

The second protein segment can be a full-length protein or a protein fragment or polypeptide. Proteins commonly used in fusion protein construction include β -galactosidase, β -glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the napsin-like aspartyl protease polypeptide-encoding sequence and the heterologous protein sequence, so that the napsin-like aspartyl protease polypeptide can be cleaved and purified away from the heterologous moiety.

25

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two protein segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises napsin-like aspartyl protease coding sequences disclosed herein in proper reading frame with nucleotides encoding the second protein segment and expressing

30

the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human napsin-like aspartyl protease can be obtained using napsin-like aspartyl protease polynucleotides (described below) to make suitable probes or primers to screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of napsin-like aspartyl protease, and expressing the cDNAs as is known in the art.

Polynucleotides

An napsin-like aspartyl protease polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for an napsin-like aspartyl protease polypeptide. The complement of a coding sequence for the napsin-like aspartyl protease of SEQ ID NO:2 is shown in SEQ ID NO:1.

Degenerate nucleotide sequences encoding human napsin-like aspartyl protease polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, or 70, more preferably about 75, 90, 96, or 98% identical to the complement of the napsin-like aspartyl protease sequence shown in SEQ ID NO:1 also are napsin-like aspartyl protease polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of napsin-like aspartyl protease polynucleotides which encode biologically active napsin-like aspartyl protease polypeptides also are napsin-like aspartyl protease polynucleotides.

Identification of Variants and Homologs

Variants and homologs of the napsin-like aspartyl protease polynucleotides disclosed above also are napsin-like aspartyl protease polynucleotides. Typically, homologous napsin-like aspartyl protease polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known napsin-like aspartyl protease polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 °C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the napsin-like aspartyl protease polynucleotides disclosed herein can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of napsin-like aspartyl protease polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol.* 81, 123 (1973). Variants of human napsin-like aspartyl protease polynucleotides or napsin-like aspartyl protease polynucleotides of other species can therefore be identified, for example, by hybridizing a putative homologous napsin-like aspartyl protease polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or its complement to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising napsin-like aspartyl protease polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to napsin-like aspartyl protease polynucleotides or their complements following stringent hybridization and/or wash conditions are also napsin-like aspartyl protease polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20 °C below the calculated T_m of the hybrid under study. The T_m of a hybrid between an napsin-like aspartyl protease polynucleotide having a coding sequence disclosed herein and a polynucleotide sequence which is at least about 50, 55, 60, 65, 70, more preferably about 75, 90, 96, or 98% identical to that nucleotide sequence can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A.* 48, 1390 (1962):

$$T_m = 81.5\text{ }^{\circ}\text{C} - 16.6(\log_{10} [\text{Na}^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l,$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65 °C, or 50% formamide, 4X SSC at 42 °C, or 0.5X SSC, 0.1% SDS at 65 °C. Highly stringent wash conditions include, for example, 0.2X SSC at 65 °C.

Preparation of Polynucleotides

A naturally occurring napsin-like aspartyl protease polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or synthesized using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated napsin-like aspartyl protease polynucleotides. For example, restriction enzymes and probes can

be used to isolate polynucleotide fragments which comprise napsin-like aspartyl protease nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

5 Napsin-like aspartyl protease cDNA molecules can be made with standard molecular biology techniques, using napsin-like aspartyl protease mRNA as a template. Napsin-like aspartyl protease cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to
10 obtain additional copies of napsin-like aspartyl protease polynucleotides, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize napsin-like aspartyl protease polynucleotides. The degeneracy of the genetic code allows
15 alternate nucleotide sequences to be synthesized which will encode an napsin-like aspartyl protease polypeptide having, for example, the amino acid sequence shown in SEQ ID NO:2 or a biologically active variant of that sequence.

Extending Polynucleotides

20 Various PCR-based methods can be used to extend the nucleic acid sequences encoding the disclosed portions of human napsin-like aspartyl protease to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic.* 2, 318-322, 1993). Genomic DNA is
25 first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res.* 16, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68 - 72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic.* 1, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations are used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res.* 19, 3055-3060, 1991. Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA. This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For

example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Obtaining Polypeptides

Napsin-like aspartyl protease polypeptides can be obtained, for example, by purification from cells, by expression of napsin-like aspartyl protease polynucleotides, or by direct chemical synthesis.

Protein Purification

Napsin-like aspartyl protease polypeptides can be purified from cells, such as primary tumor cells, metastatic cells, or cancer cell lines (*e.g.*, colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, or the H392 glioblastoma cell line), as well as cells transfected with an napsin-like aspartyl protease expression construct. Squamous cell carcinomas are especially useful sources of napsin-like aspartyl protease polypeptides. A purified napsin-like aspartyl protease polypeptide is separated from other compounds which normally associate with the napsin-like aspartyl protease polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified napsin-like aspartyl protease polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known

in the art, such as SDS-polyacrylamide gel electrophoresis. Enzymatic activity of the purified preparations can be assayed, for example, as described in Example 2.

Expression of Polynucleotides

5 To express an napsin-like aspartyl protease polypeptide, an napsin-like aspartyl protease polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding napsin-like aspartyl
10 protease polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

15 A variety of expression vector/host systems can be utilized to contain and express sequences encoding an napsin-like aspartyl protease polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with
20 yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

25 The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including
30 constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the

BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO, and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding an napsin-like aspartyl protease polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the napsin-like aspartyl protease polypeptide. For example, when a large quantity of an napsin-like aspartyl protease polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the napsin-like aspartyl protease polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem.* 264, 5503-5509, 1989 or pGEX vectors (Promega, Madison, Wis.) can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or Factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol.* 153, 516-544, 1987.

5

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding napsin-like aspartyl protease polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be
10 used alone or in combination with the omega leader sequence from TMV (Takamatsu *EMBO J.* 6, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J.* 3, 1671-1680, 1984; Broglie *et al.*, *Science* 224, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ.* 17, 85-105, 1991). These constructs can be introduced into plant
15 cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

20 An insect system also can be used to express an napsin-like aspartyl protease polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding napsin-like aspartyl protease polypeptides can be cloned into a non-essential region of the
25 virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of napsin-like aspartyl protease polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which napsin-like aspartyl protease
30 polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci.* 91, 3224-3227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be utilized in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding napsin-like aspartyl protease polypeptides can be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing an napsin-like aspartyl protease polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci.* 81, 3655-3659, 1984). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding napsin-like aspartyl protease polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding an napsin-like aspartyl protease polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell

system which is used (see Scharf *et al.*, *Results Probl. Cell Differ.* 20, 125-162, 1994).

Host Cells

5 A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process an expressed napsin-like aspartyl protease polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro"
10 form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen
15 to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express napsin-like aspartyl protease polypeptides can be transformed using expression vectors which can contain viral
20 origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which
25 successfully express the introduced napsin-like aspartyl protease sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems can be used to recover transformed cell lines.
30 These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell* 11, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy

et al., *Cell* 22, 817-23, 1980). Genes which can be employed in *tk⁻* or *aprt⁻* cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci.* 77, 3567-70, 1980); *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol.* 150, 1-14, 1981); and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992 *supra*). Additional selectable genes have been described, for example *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci.* 85, 8047-51, 1988). Visible markers such as anthocyanins, β -glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol.* 55, 121-131, 1995).

15

Detecting Expression of Polypeptides

Although the presence of marker gene expression suggests that the napsin-like aspartyl protease polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding an napsin-like aspartyl protease polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode an napsin-like aspartyl protease polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding an napsin-like aspartyl protease polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the napsin-like aspartyl protease polynucleotide.

25

Alternatively, host cells which contain an napsin-like aspartyl protease polynucleotide and which express an napsin-like aspartyl protease polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations

30

and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein.

5 The presence of a polynucleotide sequence encoding an napsin-like aspartyl protease polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding an napsin-like aspartyl protease polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding an napsin-like
10 aspartyl protease polypeptide to detect transformants which contain an napsin-like aspartyl protease polynucleotide.

A variety of protocols for detecting and measuring the expression of an napsin-like aspartyl protease polypeptide, using either polyclonal or monoclonal antibodies
15 specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on an napsin-like aspartyl protease polypeptide can be used, or a competitive binding assay can be employed.
20 These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med.* 158, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the
25 art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding napsin-like aspartyl protease polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding an napsin-like aspartyl protease
30 polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to

synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase, such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding an napsin-like aspartyl protease polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode napsin-like aspartyl protease polypeptides can be designed to contain signal sequences which direct secretion of napsin-like aspartyl protease polypeptides through a prokaryotic or eukaryotic cell membrane.

Other constructions can be used to join a sequence encoding an napsin-like aspartyl protease polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the napsin-like aspartyl protease polypeptide can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing an napsin-like aspartyl protease polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate

purification on IMAC (immobilized metal ion affinity chromatography as described in Porath *et al.*, *Prot. Exp. Purif.* 3, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the napsin-like aspartyl protease polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol.* 12, 441-453, 1993).

Chemical Synthesis

Sequences encoding an napsin-like aspartyl protease polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al.* *Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, an napsin-like aspartyl protease polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence. For example, napsin-like aspartyl protease polypeptides can be produced by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc.* 85, 2149-2154, 1963; Roberge *et al.*, *Science* 269, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of napsin-like aspartyl protease polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, *PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES*, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic napsin-like aspartyl protease polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the napsin-like aspartyl protease polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce napsin-like aspartyl protease polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter napsin-like aspartyl protease polypeptide-encoding sequences for a variety of reasons, including modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of an napsin-like aspartyl protease polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, F(ab')₂, and Fv, which are capable of binding an epitope of an napsin-like aspartyl protease polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of an napsin-like aspartyl protease polypeptide can be used therapeutically, as well as in immunochemical assays, including but not limited to Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical

assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to an napsin-like aspartyl protease polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to napsin-like aspartyl protease polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate an napsin-like aspartyl protease polypeptide from solution.

Napsin-like aspartyl protease polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, an napsin-like aspartyl protease polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies which specifically bind to an napsin-like aspartyl protease polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature* 256, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods* 81, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci.* 80, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol.* 62, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*,
5 *Proc. Natl. Acad. Sci.* 81, 6851-6855, 1984; Neuberger *et al.*, *Nature* 312, 604-608, 1984; Takeda *et al.*, *Nature* 314, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or
10 may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, one can produce humanized antibodies using recombinant methods, as described in
15 GB2188638B. Antibodies which specifically bind to an napsin-like aspartyl protease polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can
20 be adapted using methods known in the art to produce single chain antibodies which specifically bind to napsin-like aspartyl protease polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, *Proc. Natl. Acad. Sci.* 88, 11120-23, 1991).

25 Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev.* 5, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain
30 antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol.* 15,

159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem.* 269, 199-206.

5 A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology. Verhaar *et al.*, 1995, *Int. J. Cancer* 61, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth.* 165, 81-10 91.

Antibodies which specifically bind to napsin-like aspartyl protease polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci.* 86, 3833-3837, 1989; 15 Winter *et al.*, *Nature* 349, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in 20 WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies of the invention can be purified by methods well known in the art. For 25 example, antibodies can be affinity purified by passage over a column to which an napsin-like aspartyl protease polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

30 Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary

nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of napsin-like aspartyl protease gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such as alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamides, carboxymethyl esters, carbonates, and phosphate triesters. See Brown, *Meth. Mol. Biol.* 20, 1-8, 1994; Sonveaux, *Meth. Mol. Biol.* 26, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev.* 90, 543-583, 1990.

Modifications of napsin-like aspartyl protease gene expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the napsin-like aspartyl protease gene. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (e.g., Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful duplex formation between an antisense oligonucleotide and the complementary sequence of an napsin-like aspartyl protease polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to an napsin-like aspartyl protease polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent napsin-like aspartyl protease nucleotides, can provide targeting specificity for napsin-like aspartyl protease mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular napsin-like aspartyl protease polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to an napsin-like aspartyl protease polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. See, e.g., Agrawal *et al.*, *Trends Biotechnol.* 10, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev.* 90, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett.* 215, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. See, e.g., Cech, *Science* 236, 1532-1539; 1987; Cech, *Ann. Rev. Biochem.* 59, 543-568; 1990, Cech, *Curr. Opin.*

Struct. Biol. 2, 605-609; 1992, Couture & Stinchcomb, *Trends Genet.* 12, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of an napsin-like aspartyl protease polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the napsin-like aspartyl protease polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al.* *Nature* 334, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within an napsin-like aspartyl protease RNA target are initially identified by scanning the RNA molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the napsin-like aspartyl protease target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. The suitability of candidate targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related; thus, upon hybridizing to

the napsin-like aspartyl protease target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

5 Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease napsin-like aspartyl protease expression. Alternatively, if it is desired that the cells stably retain the DNA construct, it can be supplied on a plasmid and maintained as a separate element or
10 integrated into the genome of the cells, as is known in the art. The DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

15 As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of napsin-like aspartyl protease mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

20

Screening Methods

The invention provides methods for identifying modulators, *i.e.*, candidate or test compounds which bind to napsin-like aspartyl protease polypeptides or polynucleotides and/or have a stimulatory or inhibitory effect on, for example,
25 expression or activity of the napsin-like aspartyl protease polypeptide or polynucleotide, so as to regulate degradation of the extracellular matrix. Decreased extracellular matrix degradation is useful for preventing or suppressing malignant cells from metastasizing. Increased extracellular matrix degradation may be desired, for example, in developmental disorders characterized by inappropriately low levels
30 of extracellular matrix degradation or in regeneration.

The invention provides assays for screening test compounds which bind to or modulate the activity of an napsin-like aspartyl protease polypeptide or an napsin-like aspartyl protease polynucleotide. A test compound preferably binds to an napsin-like aspartyl protease polypeptide or polynucleotide. More preferably, a test compound decreases an napsin-like aspartyl protease activity of an napsin-like aspartyl protease polypeptide or expression of an napsin-like aspartyl protease polynucleotide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

10 Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. See Lam, *Anticancer Drug Des.* 12, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (see, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90, 6909, 1993; Erb *et al.* *Proc. Natl. Acad. Sci. U.S.A.* 91, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem.* 37, 2678, 1994; Cho *et al.*, *Science* 261, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2061; Gallop *et al.*, *J. Med. Chem.* 37, 1233, 1994). Libraries of compounds can be presented in solution (see, e.g., Houghten, *BioTechniques* 13, 412-421, 1992), or on beads (Lam, *Nature*

354, 82-84, 1991), chips (Fodor, *Nature* 364, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89, 1865-1869, 1992), or phage (Scott & Smith, *Science* 249, 386-390, 1990; Devlin, *Science* 249, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci.* 97, 6378-6382, 1990; Felici, *J. Mol. Biol.* 222, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to napsin-like aspartyl protease polypeptides or polynucleotides or to affect napsin-like aspartyl protease activity or napsin-like aspartyl protease gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 μ l. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 19, 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme

assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity* 2, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupies the active site or a fibronectin domain of the napsin-like aspartyl protease polypeptide, thereby making the active site or fibronectin domain inaccessible to substrate such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules. In binding assays, either the test compound or the napsin-like aspartyl protease polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to the napsin-like aspartyl protease polypeptide can then be accomplished, for example, by direct counting of radioemission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to an napsin-like aspartyl protease polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a target polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and an napsin-like aspartyl protease polypeptide. (McConnell *et al.*, *Science* 257, 1906-1912, 1992).

10

Determining the ability of a test compound to bind to an napsin-like aspartyl protease polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA). Sjolander & Urbaniczky, *Anal. Chem.* 63, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5, 699-705, 1995. BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

15

In yet another aspect of the invention, an napsin-like aspartyl protease polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell* 72, 223-232, 1993; Madura *et al.*, *J. Biol. Chem.* 268, 12046-12054, 1993; Bartel *et al.*, *BioTechniques* 14, 920-924, 1993; Iwabuchi *et al.*, *Oncogene* 8, 1693-1696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the napsin-like aspartyl protease polypeptide and modulate its activity.

20

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct a polynucleotide encoding an napsin-like aspartyl protease polypeptide is fused to a poly-

25

30

nucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence that encodes an unidentified protein ("prey" or "sample") is fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the napsin-like aspartyl protease polypeptide.

It may be desirable to immobilize either the napsin-like aspartyl protease polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the napsin-like aspartyl protease polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the napsin-like aspartyl protease polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to an napsin-like aspartyl protease polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, an napsin-like aspartyl protease polypeptide is a fusion protein comprising a domain that allows the napsin-like aspartyl protease polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed napsin-like aspartyl protease polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing polypeptides or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either an napsin-like aspartyl protease polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated napsin-like aspartyl protease polypeptides or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to an napsin-like aspartyl protease polypeptide polynucleotides, or a test compound, but which do not interfere with a desired binding site, such as the active site or a fibronectin domain of the napsin-like aspartyl protease polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the napsin-like aspartyl protease polypeptide (or polynucleotides) or test compound, enzyme-linked assays which rely on detecting an

napsin-like aspartyl protease activity of the napsin-like aspartyl protease polypeptide, and SDS gel electrophoresis under non-reducing conditions.

5 Screening for test compounds which bind to an napsin-like aspartyl protease polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises an napsin-like aspartyl protease polynucleotide or polypeptide can be used in a cell-based assay system. An napsin-like aspartyl protease polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established
10 cell line, including neoplastic cell lines such as the colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, and the H392 glioblastoma cell line, can be used. An intact cell is contacted with a test compound. Binding of the test compound to an napsin-like aspartyl protease polypeptide or polynucleotide is determined as described above, after lysing the cell
15 to release the napsin-like aspartyl protease polypeptide-test compound complex.

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease an napsin-like aspartyl protease activity of an napsin-like aspartyl protease polypeptide. Napsin-like aspartyl protease activity can be measured, for example, using the method described in Schauer-Vukasinovic *et al.*, *Eur. J. Biochem.* 267, 2573-80 (2000). Napsin-like aspartyl protease activity can be measured after contacting either a purified napsin-like aspartyl protease polypeptide, a cell extract, or an intact cell with
25 a test compound. A test compound which decreases napsin-like aspartyl protease activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing extracellular matrix degradation. A test compound which increases napsin-like aspartyl protease activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is
30 identified as a potential therapeutic agent for increasing extracellular matrix degradation.

Gene Expression

In another embodiment, test compounds which increase or decrease napsin-like aspartyl protease gene expression are identified. An napsin-like aspartyl protease polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the napsin-like aspartyl protease polynucleotide is determined. The level of expression of napsin-like aspartyl protease mRNA or polypeptide in the presence of the test compound is compared to the level of expression of napsin-like aspartyl protease mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of napsin-like aspartyl protease mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of napsin-like aspartyl protease mRNA or polypeptide expression. Alternatively, when expression of the mRNA or protein is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of napsin-like aspartyl protease mRNA or polypeptide expression.

The level of napsin-like aspartyl protease mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or protein. Either qualitative or quantitative methods can be used. The presence of polypeptide products of an napsin-like aspartyl protease polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into an napsin-like aspartyl protease polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses an napsin-like aspartyl protease polynucleotide can be used in a cell-based assay system. The napsin-like aspartyl protease

polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, including neoplastic cell lines such as the colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, and the H392 glioblastoma cell line, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise an napsin-like aspartyl protease polypeptide, napsin-like aspartyl protease polynucleotide, antibodies which specifically bind to an napsin-like aspartyl protease polypeptide, or mimetics, agonists, antagonists, or inhibitors of an napsin-like aspartyl protease polypeptide. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be

formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

5 Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose,
10 hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

15 Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity
20 of active compound, *i.e.*, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a
25 filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

30 Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks'

solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection
5 suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated
10 solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a
15 manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in
20 aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

25 Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of
30 administration.

Therapeutic Indications and Methods

Cancer. Cancer is a disease fundamentally caused by oncogenic cellular transformation. There are several hallmarks of transformed cells that distinguish them from their normal counterparts and underlie the pathophysiology of cancer. These include uncontrolled cellular proliferation, unresponsiveness to normal death-inducing signals (immortalization), increased cellular motility and invasiveness, increased ability to recruit blood supply through induction of new blood vessel formation (angiogenesis), genetic instability, and dysregulated gene expression. Various combinations of these aberrant physiologies, along with the acquisition of drug-resistance frequently lead to an intractable disease state in which organ failure and patient death ultimately ensue.

Most standard cancer therapies target cellular proliferation and rely on the differential proliferative capacities between transformed and normal cells for their efficacy. This approach is hindered by the facts that several important normal cell types are also highly proliferative and that cancer cells frequently become resistant to these agents. Thus, the therapeutic indices for traditional anti-cancer therapies rarely exceed 2.0.

The advent of genomics-driven molecular target identification has opened up the possibility of identifying new cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for cancer patients. Thus, newly discovered tumor-associated genes and their products can be tested for their role(s) in disease and used as tools to discover and develop innovative therapies. Genes playing important roles in any of the physiological processes outlined above can be characterized as cancer targets.

Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized *in vitro* for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical

modulators of their biochemical activities. Agonists and/or antagonists of target protein activity can be identified in this manner and subsequently tested in cellular and *in vivo* disease models for anti-cancer activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

The human napsin-like aspartyl protease gene provides a therapeutic target for decreasing extracellular matrix degradation, in particular for treating or preventing metastatic cancer. For example, blocking a fibronectin domain of human ephrin-like serine protease can suppress or prevent migration or metastasis of tumor cells in response to fibronectin (9, 10). Cancers whose metastasis can be suppressed according to the invention include adenocarcinoma, melanoma, cancers of the adrenal gland, bladder, bone, breast, cervix, gall bladder, liver, lung, ovary, pancreas, prostate, testis, and uterus. Circulating tumor cells arrested in the capillary beds of different organs must invade the endothelial cell lining and degrade its underlying basement membrane (BM) in order to invade into the extravascular tissue(s) where they establish metastasis (1, 2). Metastatic tumor cells often attach at or near the intercellular junctions between adjacent endothelial cells. Such attachment of the metastatic cells is followed by rupture of the junctions, retraction of the endothelial cell borders and migration through the breach in the endothelium toward the exposed underlying BM (1, 11).

Once located between endothelial cells and the BM, the invading cells must degrade the subendothelial glycoproteins and proteoglycans of the BM in order to migrate out of the vascular compartment. Several cellular enzymes (*e.g.*, collagenase IV, plasminogen activator, cathepsin B, elastase) are thought to be involved in degradation of BM (2, 11). Suppression of human napsin-like aspartyl protease activity therefore can be used to suppress tumor cell invasion and metastasis.

- 44 -

Osteoporosis. Osteoporosis is a disease characterized by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk . It is the most common human metabolic bone disorder. Established osteoporosis includes the presence of fractures.

5

Bone turnover occurs by the action of two major effector cell types within bone: the osteoclast, which is responsible for bone resorption, and the osteoblast, which synthesizes and mineralizes bone matrix. The actions of osteoclasts and osteoblasts are highly co-ordinated. Osteoclast precursors are recruited to the site of turnover; they differentiate and fuse to form mature osteoclasts which then resorb bone. Attached to the bone surface, osteoclasts produce an acidic microenvironment in a tightly defined junction between the specialized osteoclast border membrane and the bone matrix, thus allowing the localized solubilization of bone matrix. This in turn facilitates the proteolysis of demineralized bone collagen. Matrix degradation is thought to release matrix-associated growth factor and cytokines, which recruit osteoblasts in a temporally and spatially controlled fashion. Osteoblasts synthesise and secrete new bone matrix proteins, and subsequently mineralise this new matrix. In the normal skeleton this is a physiological process which does not result in a net change in bone mass. In pathological states, such as osteoporosis, the balance between resorption and formation is altered such that bone loss occurs. See WO 99/45923.

The osteoclast itself is the direct or indirect target of all currently available osteoporosis agents with the possible exception of fluoride. Antiresorptive therapy prevents further bone loss in treated individuals. Osteoblasts are derived from multipotent stem cells which reside in bone marrow and also gives rise to adipocytes, chondrocytes, fibroblasts and muscle cells. Selective enhancement of osteoblast activity is a highly desirable goal for osteoporosis therapy since it would result in an increase in bone mass, rather than a prevention of further bone loss. An effective anabolic therapy would be expected to lead to a significantly greater reduction in fracture risk than currently available treatments.

30

- 45 -

5 The agonists or antagonists to the newly discovered polypeptides may act as antiresorptive by directly altering the osteoclast differentiation, osteoclast adhesion to the bone matrix or osteoclast function of degrading the bone matrix. The agonists or antagonists could indirectly alter the osteoclast function by interfering in the synthesis and/or modification of effector molecules of osteoclast differentiation or function such as cytokines, peptide or steroid hormones, proteases, etc.

10 The agonists or antagonists to the newly discovered polypeptides may act as anabolics by directly enhancing the osteoblast differentiation and /or its bone matrix forming function. The agonists or antagonists could also indirectly alter the osteoblast function by enhancing the synthesis of growth factors, peptide or steroid hormones or decreasing the synthesis of inhibitory molecules.

15 The agonists and antagonists may be used to mimic, augment or inhibit the action of the newly discovered polypeptides which may be useful to treat osteoporosis, Paget's disease, degradation of bone implants particularly dental implants.

20 The invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a polypeptide-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as
25 described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

30 A reagent which affects napsin-like aspartyl protease activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce napsin-like aspartyl protease

activity. The reagent preferably binds to an expression product of a human napsin-like aspartyl protease gene. If the expression product is a polypeptide, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The
5 cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about
10 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung or liver.

15 A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably
20 about 1.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

25 Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to
30 polyethylene glycol. Optionally, a liposome comprises a compound capable of

targeting the liposome to a tumor cell, such as a tumor cell ligand exposed on the outer surface of the liposome.

5 Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μ g to about 10 μ g of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μ g to about 5 μ g of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μ g of polynucleotides is combined
10 with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al.* *Trends in Biotechnol.* 11, 202-05 (1993);
15 Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem.* 263, 621-24 (1988); Wu *et al.*, *J. Biol. Chem.* 269, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem.* 266, 338-42 (1991).

20 If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads,
25 protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of
30 those skilled in the art. A therapeutically effective dose refers to that amount of

active ingredient which increases or decreases extracellular matrix degradation relative to that which occurs in the absence of the therapeutically effective dose.

5 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

10 Therapeutic efficacy and toxicity, *e.g.*, ED₅₀ (the dose therapeutically effective in 50% of the population) and LD₅₀ (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD₅₀/ED₅₀.

15 Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED₅₀ with
20 little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to
25 provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be
30 administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

Effective *in vivo* dosages of an antibody are in the range of about 5 µg to about 50 µg/kg, about 50 µg to about 5 mg/kg, about 100 µg to about 500 µg/kg of patient body weight, and about 200 to about 250 µg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 µg to about 2 mg, about 5 µg to about 500 µg, and about 20 µg to about 100 µg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of an napsin-like aspartyl protease polynucleotide or activity of an napsin-like aspartyl protease polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of an napsin-like aspartyl protease polynucleotide or the activity of an napsin-like aspartyl protease polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to napsin-like aspartyl protease-specific mRNA, quantitative RT-PCR, immunologic detection of an napsin-

like aspartyl protease polypeptide, or measurement of napsin-like aspartyl protease activity.

5 In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described
10 above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows,
15 horses, rabbits, monkeys, and most preferably, humans.

The above disclosure generally describes the present invention, and all patents and patent applications cited in this disclosure are expressly incorporated herein. A more complete understanding can be obtained by reference to the following specific
20 examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1*Detection of napsin-like aspartyl protease activity*

The polynucleotide of SEQ ID NO: 1 is inserted into the expression vector pCEV4 and the expression vector pCEV4-napsin-like aspartyl protease polypeptide obtained is transfected into human embryonic kidney 293 cells. Lysates from these cells are prepared by sonication and subsequent treatment of the cellular pellet with 1 ml PBS buffer containing 20 mM EDTA, 1 mM PMSF, and 1 % /; 1 Elugent for 5 min. Following centrifugation at 14000 rpm for 15 min at 4°C. Enzymatic activity of the resultant lysate is monitored by the increase in fluorescence per minute using a Fluostar (BMG LabTechnologies, Offenburg, Germany) and 96-well plates (Dynex Technologies, Chantilly, V A. USA) at room temperature. The reaction is performed in assay buffer containing 100 mM sodium acetate and j20 mM EDTA, pH 4,7. Typically, 35 micro-l of assay buffer is mixed with 5 micro-l of lysated (50 times prediluted in assay buffer) and reaction is initiated by the addition Of 10 micro-l of 100 micro-M fluorogenic substrate DS3. Initial velocities (data are derived from one transfection experiment with triplicate measurements) are determined using gamma-ex = 390 nm and gamme-em = 538 nm. For inhibition experiments, pepstatin is diluted into the assay buffer just before rate measurements are made with final concentration 1x10-7 M. The velocity corresponding to the maximum activity of napsin-like aspartyl protease cells is set to 100%. It is shown that the polypeptide of SEQ ID NO: 2 has napsin-like aspartyl protease activity.

EXAMPLE 2*Expression of recombinant napsin-like aspartyl protease*

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of a recombinant human napsin-like aspartyl protease in yeast. The encoding DNA sequence is derived from the complement of the nucleotide sequence shown in SEQ ID NO:1. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains

at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag, and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added.

5 After digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes, the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

10 The yeast is cultivated under usual conditions in 5 liter shake flasks, and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San
15 Diego, CA) according to manufacturer's instructions. Purified human napsin-like aspartyl protease is obtained.

EXAMPLE 3

Identification of a test compound which binds to an napsin-like aspartyl protease polypeptide
20

Purified napsin-like aspartyl protease polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at
25 pH 7.0 in a physiological buffer solution. Napsin-like aspartyl protease polypeptides comprise the amino acid sequence shown in SEQ ID NO:2. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

30 The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to an napsin-like aspartyl protease polypeptide is

detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound was not incubated is identified as a compound which binds to an napsin-like aspartyl protease polypeptide.

5

EXAMPLE 4

Identification of a test compound which decreases napsin-like aspartyl protease activity

10 Cellular extracts from the human colon cancer cell line HCT116 are contacted with test compounds from a small molecule library and assayed for napsin-like aspartyl protease activity. Control extracts, in the absence of a test compound, also are assayed. Aspartyl protease activity can be measured as described in Schauer-Vukasinovic *et al.*, *Eur. J. Biochem.* 267, 2573-80 (2000).

15

A test compound which decreases napsin-like aspartyl protease activity of the extract relative to the control extract by at least 20% is identified as an napsin-like aspartyl protease inhibitor.

20

EXAMPLE 5

Identification of a test compound which decreases napsin-like aspartyl protease gene expression

25 A test compound is administered to a culture of the breast tumor cell line MDA-468 and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells incubated for the same time without the test compound provides a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem.* 18, 5294-99, 1979). Northern blots are prepared using 20 to 30 µg total RNA and hybridized with a ³²P-labeled napsin-like aspartyl protease-specific probe at 65 °C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides

30

selected from SEQ ID NO:1. A test compound which decreases the napsin-like aspartyl protease -specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of napsin-like aspartyl protease gene expression.

5

EXAMPLE 6

Treatment of a breast tumor with a reagent which specifically binds to an napsin-like aspartyl protease gene product

- 10 Synthesis of antisense napsin-like aspartyl protease oligonucleotides comprising at least 11 contiguous nucleotides selected from SEQ ID NO:1 is performed on a Pharmacia Gene Assembler series synthesizer using the phosphoramidite procedure (Uhlmann *et al.*, *Chem. Rev.* 90, 534-83, 1990). Following assembly and deprotection, oligonucleotides are ethanol-precipitated twice, dried, and suspended in
- 15 phosphate-buffered saline (PBS) at the desired concentration. Purity of these oligonucleotides is tested by capillary gel electrophoreses and ion exchange HPLC. Endotoxin levels in the oligonucleotide preparation are determined using the *Limulus* Amebocyte Assay (Bang, *Biol. Bull. (Woods Hole, Mass.)* 105, 361-362, 1953).
- 20 An aqueous composition containing the antisense oligonucleotides at a concentration of 0.1-100 μ M is injected directly into a breast tumor with a needle. The needle is placed in the tumors and withdrawn while expressing the aqueous composition within the tumor.
- 25 The breast tumor is monitored over a period of days or weeks. Additional injections of the antisense oligonucleotides can be given during that time. Metastasis of the breast tumor is suppressed due to decreased napsin-like aspartyl protease activity of the breast tumor cells.

REFERENCES

1. Nicolson (1988) Organ specificity of tumor metastasis: Role of preferential adhesion, invasion and growth of malignant cells at specific secondary sites. *Cancer Met. Rev.* 7, 143-188.
5
2. Liotta *et al.* (1983) Tumor invasion and the extracellular matrix. *Lab. Invest.* 49, 639-649.
- 10 3. Vlodavsky *et al.* (1987) Endothelial cell-derived basic fibroblast growth factor: Synthesis and deposition into subendothelial extracellular matrix. *Proc. Natl. Acad. Sci. USA* 84, 2292-2296.
- 15 4. Folkman *et al.* (1980) A heparin-binding angiogenic protein--basic fibroblast growth factor--is stored within basement membrane. *Am. J. Pathol.* 130, 393-400.
5. Cardon-Cardo *et al.* (1990) Expression of basic fibroblast growth factor in normal human tissues. *Lab. Invest.* 63, 832-840.
- 20 6. Vlodavsky *et al.* (1991) Extracellular sequestration and release of fibroblast growth factor: a regulatory mechanism? *Trends Biochem. Sci.* 16, 268-271.
- 25 7. Vlodavsky *et al.* (1993) Extracellular matrix-bound growth factors, enzymes and plasma proteins. In BASEMENT MEMBRANES: CELLULAR AND MOLECULAR ASPECTS Rohrbach & Timpl, eds., pp327-343. Academic Press Inc., Orlando, Fla.
8. Ross (1993) The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature (Lond.)* 362, 801-809.

9. McCarthy *et al.* (1986). Human fibronectin contains distinct adhesion- and motility-promoting domains for metastatic melanoma cells. *J. Cell Biol.* 102, 179-88.
- 5 10. Van Muijen *et al.* (1995) Properties of metastasizing and non-metastasizing human melanoma cells. *Recent Results in Cancer Research* 139, 104-22.
11. Price *et al.* (1997) The Biochemistry of Cancer Dissemination, in *Critical Reviews in Biochemistry and Mol. Biol.* 32, 175-253.
- 10 12. Leytus SP, Loeb KR, Hagen FS, Kurachi K, Davie EW. A novel trypsin-like serine protease (hepsin) with a putative epidermis-specific domain expressed by human liver and hepatoma cells. *Biochemistry*. 1988 Feb 9;27(3):1067-74.
- 15 13. Tanimoto H, Yan Y, Clarke J, Korourian S, Shigemasa K, Parmley TH, Parham GP, O'Brien TJ. Hepsin, a cell surface serine protease identified in hepatoma cells, is overexpressed in ovarian cancer. *Cancer Res.* 1997 Jul 15;57(14):2884-7.
- 20 14. Zacharski LR, Ornstein DL, Memoli VA, Rousseau SM, Kisiel W. Expression of the factor VII activating protease, hepsin, in situ in renal cell carcinoma. *Thromb Haemost.* 1998 Apr;79(4):876-7.

CLAIMS

1. An isolated polynucleotide encoding a napsin-like aspartyl protease polypeptide and being selected from the group consisting of:
 - 5 a) a polynucleotide encoding a napsin-like aspartyl protease polypeptide comprising an amino acid sequence selected from the group consisting of:
amino acid sequences which are at least about 60% identical to
the amino acid sequence shown in SEQ ID NO: 2; and
10 the amino acid sequence shown in SEQ ID NO: 2.
 - b) a polynucleotide comprising the sequence of SEQ ID NO: 1;
 - c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);
 - d) a polynucleotide the sequence of which deviates from the
15 polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and
 - e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a to (d).
- 20 2. An expression vector containing any polynucleotide of claim 1.
3. A host cell containing the expression vector of claim 2.
4. A substantially purified napsin-like aspartyl protease polypeptide encoded by
25 a polynucleotide of claim 1.
5. A method for producing a napsin-like aspartyl protease polypeptide, wherein the method comprises the following steps:
 - a) culturing the host cell of claim 3 under conditions suitable for the
30 expression of the napsin-like aspartyl protease polypeptide; and

- b) recovering the napsin-like aspartyl protease polypeptide from the host cell culture.
- 5 6. A method for detection of a polynucleotide encoding a napsin-like aspartyl protease polypeptide in a biological sample comprising the following steps:
- a) hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and
- b) detecting said hybridization complex.
- 10 7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.
8. A method for the detection of a polynucleotide of claim 1 or a napsin-like aspartyl protease polypeptide of claim 4 comprising the steps of:
- 15 contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the napsin-like aspartyl protease polypeptide.
9. A diagnostic kit for conducting the method of any one of claims 6 to 8.
- 20 10. A method of screening for agents which decrease the activity of a napsin-like aspartyl protease, comprising the steps of:
- contacting a test compound with any napsin-like aspartyl protease polypeptide encoded by any polynucleotide of claim 1;
- detecting binding of the test compound to the napsin-like aspartyl protease polypeptide, wherein a test compound which binds to the polypeptide is
- 25 identified as a potential therapeutic agent for decreasing the activity of a napsin-like aspartyl protease.
11. A method of screening for agents which regulate the activity of a napsin-like aspartyl protease, comprising the steps of:
- 30

contacting a test compound with a napsin-like aspartyl protease polypeptide encoded by any polynucleotide of claim 1; and

detecting a napsin-like aspartyl protease activity of the polypeptide, wherein a test compound which increases the napsin-like aspartyl protease activity is identified as a potential therapeutic agent for increasing the activity of the napsin-like aspartyl protease, and wherein a test compound which decreases the napsin-like aspartyl protease activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the napsin-like aspartyl protease.

10

12. A method of screening for agents which decrease the activity of a napsin-like aspartyl protease, comprising the steps of:

contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of napsin-like aspartyl protease.

15

13. A method of reducing the activity of napsin-like aspartyl protease, comprising the steps of:

contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any napsin-like aspartyl protease polypeptide of claim 4, whereby the activity of napsin-like aspartyl protease is reduced.

20

14. A reagent that modulates the activity of a napsin-like aspartyl protease polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

25

15. A pharmaceutical composition, comprising:
the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

30

16. Use of the pharmaceutical composition of claim 15 for modulating the activity of a napsin-like aspartyl protease in a disease.
17. Use of claim 16 wherein the disease is cancer and osteoporosis.
- 5 18. A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
19. The cDNA of claim 18 which comprises SEQ ID NO:1.
- 10 20. The cDNA of claim 18 which consists of SEQ ID NO:1.
21. An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
- 15 22. The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO:1.
23. A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
- 20 24. The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO:1.
- 25 25. A purified polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
26. The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NO:2.

27. A fusion protein comprising a polypeptide having the amino acid sequence shown in SEQ ID NO:2.
28. A method of producing a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, comprising the steps of:
culturing a host cell comprising an expression vector which encodes the polypeptide under conditions whereby the polypeptide is expressed; and isolating the polypeptide.
29. The method of claim 28 wherein the expression vector comprises SEQ ID NO:1.
30. A method of detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, comprising the steps of:
hybridizing a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and detecting the hybridization complex.
31. The method of claim 30 further comprising the step of amplifying the nucleic acid material before the step of hybridizing.
32. A kit for detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, comprising:
a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1; and instructions for the method of claim 30.
33. A method of detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, comprising the steps of:
contacting a biological sample with a reagent that specifically binds to the polypeptide to form a reagent-polypeptide complex; and

detecting the reagent-polypeptide complex.

34. The method of claim 33 wherein the reagent is an antibody.
- 5 35. A kit for detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, comprising:
an antibody which specifically binds to the polypeptide; and
instructions for the method of claim 33.
- 10 36. A method of screening for agents which can modulate the activity of a human napsin-like aspartyl protease, comprising the steps of:
contacting a test compound with a polypeptide comprising an amino acid
sequence selected from the group consisting of: (1) amino acid sequences
which are at least about 60% identical to the amino acid sequence shown in
15 SEQ ID NO:2 and (2) the amino acid sequence shown in SEQ ID NO:2; and
detecting binding of the test compound to the polypeptide, wherein a test
compound which binds to the polypeptide is identified as a potential agent for
regulating activity of the human napsin-like aspartyl protease.
- 20 37. The method of claim 36 wherein the step of contacting is in a cell.
38. The method of claim 36 wherein the cell is in vitro.
39. The method of claim 36 wherein the step of contacting is in a cell-free
25 system.
40. The method of claim 36 wherein the polypeptide comprises a detectable label.
41. The method of claim 36 wherein the test compound comprises a detectable
30 label.

42. The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.
43. The method of claim 36 wherein the polypeptide is bound to a solid support.
- 5 44. The method of claim 36 wherein the test compound is bound to a solid support.
- 10 45. A method of screening for agents which modulate an activity of a human napsin-like aspartyl protease, comprising the steps of:
contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 60% identical to the amino acid sequence shown in SEQ ID NO:2 and (2) the amino acid sequence shown in SEQ ID NO:2; and
15 detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human napsin-like aspartyl protease, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential agent for decreasing the activity of the human napsin-
20 like aspartyl protease.
46. The method of claim 45 wherein the step of contacting is in a cell.
47. The method of claim 45 wherein the cell is in vitro.
- 25 48. The method of claim 45 wherein the step of contacting is in a cell-free system.
49. A method of screening for agents which modulate an activity of a human napsin-like aspartyl protease, comprising the steps of:
- 30

- 64 -

5 contacting a test compound with a product encoded by a polynucleotide
 which comprises the nucleotide sequence shown in SEQ ID NO:1; and
 detecting binding of the test compound to the product, wherein a test
 compound which binds to the product is identified as a potential agent for
 regulating the activity of the human napsin-like aspartyl protease.

50. The method of claim 49 wherein the product is a polypeptide.

51. The method of claim 49 wherein the product is RNA.

10

52. A method of reducing activity of a human napsin-like aspartyl protease,
 comprising the step of:

 contacting a cell with a reagent which specifically binds to a product encoded
 by a polynucleotide comprising the nucleotide sequence shown in SEQ ID
15 NO:1, whereby the activity of a human napsin-like aspartyl protease is
 reduced.

53. The method of claim 52 wherein the product is a polypeptide.

20 54. The method of claim 53 wherein the reagent is an antibody.

55. The method of claim 52 wherein the product is RNA.

56. The method of claim 55 wherein the reagent is an antisense oligonucleotide.

25

57. The method of claim 56 wherein the reagent is a ribozyme.

58. The method of claim 52 wherein the cell is in vitro.

30 59. The method of claim 52 wherein the cell is in vivo.

- 65 -

- 5 60. A pharmaceutical composition, comprising:
a reagent which specifically binds to a polypeptide comprising the amino acid
sequence shown in SEQ ID NO:2; and
a pharmaceutically acceptable carrier.
61. The pharmaceutical composition of claim 60 wherein the reagent is an
antibody.
- 10 62. A pharmaceutical composition, comprising:
a reagent which specifically binds to a product of a polynucleotide
comprising the nucleotide sequence shown in SEQ ID NO:1; and
a pharmaceutically acceptable carrier.
- 15 63. The pharmaceutical composition of claim 62 wherein the reagent is a
ribozyme.
64. The pharmaceutical composition of claim 62 wherein the reagent is an
antisense oligonucleotide.
- 20 65. The pharmaceutical composition of claim 62 wherein the reagent is an
antibody.
- 25 66. A pharmaceutical composition, comprising:
an expression vector encoding a polypeptide comprising the amino acid
sequence shown in SEQ ID NO:2; and
a pharmaceutically acceptable carrier.
- 30 67. The pharmaceutical composition of claim 66 wherein the expression vector
comprises SEQ ID NO:1.

- 66 -

68. A method of treating a napsin-like aspartyl protease disfunction related disease, wherein the disease is cancer and osteoporosis, comprising the step of:
administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human napsin-like aspartyl protease, whereby symptoms of the napsin-like aspartyl protease disfunction related disease are ameliorated.
69. The method of claim 68 wherein the reagent is identified by the method of claim 36.
70. The method of claim 68 wherein the reagent is identified by the method of claim 45.
71. The method of claim 68 wherein the reagent is identified by the method of claim 49.

FIG. 1

cgggaaacgg cccctgccaa gtcttcccc acccagggcc cctccccgga gggcaacccc
ccccagccc ccccgggcc ctgttttaa tgtccccggg gggaaaaacg ccccatccc
ccccaaaaa aacgtccccg gggaccctgc gggcggggga acaccaggc
ccggaaccc gaccaaccaa agggggcccc ctttcaagt atccgggata acgaatccat
ggcccgggg gtaaacccac ccccccaa aagaaaagg aactggggg accttggaa
ttcccaacc caggatgatg tcctcccac ccacagggg aatccccca aggttggaat
gcaggccccg aatccctca gggggccct ggataag

FIG. 2

LIQGPPEGIP GLHSTLGGIP LWGGEDIILG LGIPKVPPVS FLFGGGWVYP PGHGFVIPDT
XKGGPLWLVG FPGLGVPPPA GALWVPGDVF LGGDGGVFPP GDIKNRARGG LGGCPCPGRG
PGWGKTWQGP FP

Fig. 3

MSPPPLLQPL LLLPLLNVE PSGATLIRIP LHRVQPGRR I LNLLRGWREP AELPKLGAPS
 PGDKPIFVPL SNYRDVQYFG EIGLGTPPQN FTVAFDTGSS NLWVPSRRCH FFSVPCWLHH
 RFDPKASSF QANGTKFAIQ YGTGRVDGIL SEDKLTIGGI KGASVIFGEA LWEPSLVFAF
 AHFDGILGLG FPILSVEGVR PPMDVLEQG LLDKPVFSFY LNRDPEEPDG GELVLGSDP
 AHYIPPLTFV PVTVPAYWQI HMERVKVGP LTLCAKCAA ILDTGTSLIT GPTEEIRALH
 AAIGIPLLA GEYIILCSEI PKLPAVSFLL GGWFNLTAH DYVIQTRNG VRLCLSGFQA
 LDVPPAGPF WILGDVFLGT YVAVFDRGDM KSSARVGLAR ARTRGADLGW GETAQAFPFG

FIG. 4

BLASTX - Query = AW276392; Hit = swissnew|O96009|NAP1_HUMAN

Score = 126 bits (313), Expect = 9e-29
Identities = 69/132 (52%), Positives = 77/132 (58%)
Frame = -1

Q: 397 LIQPPGIPGLHSTLGGIPLWGGEDIILGLIPKVPVVSFLFGGWWVPPGHHGVIPDT 218
LI GP E I LH+ +GGIPL GE IIL IPK+P VSFL GG W H +VI T
H: 288 LITGPTTEIRALHAAIGGIPLLAGIYIILCSEIPKLPAVSFLLGGVWFNLTADHYVIQTT 347

Q: 217 *KGGPLWLVGFPGLGVPPPPAGALWVPDGVFLGGDGGVFPDGIKNRARGGLGGCPPGRG 38
G L L GF L VPPPPAG W+ GDVFLG VF GD+K+ AR GL G
H: 348 RNVRLCLSGFQALDVPPPPAGPFWILGDVFLGTYVAVFDRGDMKSSARVGLARARTRGAD 407

Q: 37 PGWGKTWQGPFP 2
GWG+T Q FP
H: 408 LGWGETAQAQFP 419

SEQUENCE LISTING

<110> Bayer AG

Bull, Christof

<120> REGULATION OF HUMAN NAPSIN-LIKE ASPARTYL PROTEASE

<130> Lio093 foreign countries

<150> 60/214,010

<151> 2000-06-26

<160> 3

<170> PatentIn version 3.0

<210> 1

<211> 397

<212> DNA

<213> Homo sapiens

<400> 1

```
cgggaaacgg cccctgccaa gtcttcccc acccagggcc cctcccggga gggcaacccc      60
ccccagccc ccccggggcc ctgtttttaa tgtccccggg gggaaaaacg ccccatccc      120
ccccaaaaa aacgtccccg gggacccaaa gggcccctgc gggcggggga acaccaggc      180
ccggaaaccc gaccaacca agggggcccc cttttcaagt atccgggata acgaatccat      240
ggcccggggg gtaaaccac cccccccaa aaagaaagga aactgggggg acctttggaa      300
ttccaaccc caggatgatg tcctccccac ccacagggg aatcccccca aggggtggaat      360
gcaggcccg aattccctca gggggccct ggataag                                397
```

<210> 2

<211> 132

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<222> ()..()

<223> Xaa = any amino acid

<400> 2

```

Leu Ile Gln Gly Pro Pro Glu Gly Ile Pro Gly Leu His Ser Thr Leu
1           5           10           15
Gly Gly Ile Pro Leu Trp Gly Gly Glu Asp Ile Ile Leu Gly Leu Gly
20           25           30
Ile Pro Lys Val Pro Pro Val Ser Phe Leu Phe Gly Gly Gly Trp Val
35           40           45
Tyr Pro Pro Gly His Gly Phe Val Ile Pro Asp Thr Xaa Lys Gly Gly
50           55           60
Pro Leu Trp Leu Val Gly Phe Pro Gly Leu Gly Val Pro Pro Pro Ala
65           70           75           80
Gly Ala Leu Trp Val Pro Gly Asp Val Phe Leu Gly Gly Asp Gly Gly
85           90           95
Val Phe Pro Pro Gly Asp Ile Lys Asn Arg Ala Arg Gly Gly Leu Gly
100          105          110
Gly Gly Cys Pro Pro Gly Arg Gly Pro Gly Trp Gly Lys Thr Trp Gln
115          120          125
Gly Pro Phe Pro
130

```

<210> 3

<211> 420

<212> PRT

<213> Homo sapiens

<400> 3

```

Met Ser Pro Pro Pro Leu Leu Gln Pro Leu Leu Leu Leu Leu Pro Leu
1           5           10           15

```

Leu Asn Val Glu Pro Ser Gly Ala Thr Leu Ile Arg Ile Pro Leu His
 20 25 30
 Arg Val Gln Pro Gly Arg Arg Ile Leu Asn Leu Leu Arg Gly Trp Arg
 35 40 45
 Glu Pro Ala Glu Leu Pro Lys Leu Gly Ala Pro Ser Pro Gly Asp Lys
 50 55 60
 Pro Ile Phe Val Pro Leu Ser Asn Tyr Arg Asp Val Gln Tyr Phe Gly
 65 70 75 80
 Glu Ile Gly Leu Gly Thr Pro Pro Gln Asn Phe Thr Val Ala Phe Asp
 85 90 95
 Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Arg Arg Cys His Phe Phe
 100 105 110
 Ser Val Pro Cys Trp Leu His His Arg Phe Asp Pro Lys Ala Ser Ser
 115 120 125
 Ser Phe Gln Ala Asn Gly Thr Lys Phe Ala Ile Gln Tyr Gly Thr Gly
 130 135 140
 Arg Val Asp Gly Ile Leu Ser Glu Asp Lys Leu Thr Ile Gly Gly Ile
 145 150 155 160
 Lys Gly Ala Ser Val Ile Phe Gly Glu Ala Leu Trp Glu Pro Ser Leu
 165 170 175
 Val Phe Ala Phe Ala His Phe Asp Gly Ile Leu Gly Leu Gly Phe Pro
 180 185 190
 Ile Leu Ser Val Glu Gly Val Arg Pro Pro Met Asp Val Leu Val Glu
 195 200 205
 Gln Gly Leu Leu Asp Lys Pro Val Phe Ser Phe Tyr Leu Asn Arg Asp
 210 215 220
 Pro Glu Glu Pro Asp Gly Gly Glu Leu Val Leu Gly Gly Ser Asp Pro
 225 230 235 240
 Ala His Tyr Ile Pro Pro Leu Thr Phe Val Pro Val Thr Val Pro Ala
 245 250 255
 Tyr Trp Gln Ile His Met Glu Arg Val Lys Val Gly Pro Gly Leu Thr
 260 265 270
 Leu Cys Ala Lys Gly Cys Ala Ala Ile Leu Asp Thr Gly Thr Ser Leu
 275 280 285
 Ile Thr Gly Pro Thr Glu Glu Ile Arg Ala Leu His Ala Ala Ile Gly
 290 295 300
 Gly Ile Pro Leu Leu Ala Gly Glu Tyr Ile Ile Leu Cys Ser Glu Ile
 305 310 315 320
 Pro Lys Leu Pro Ala Val Ser Phe Leu Leu Gly Gly Val Trp Phe Asn
 325 330 335
 Leu Thr Ala His Asp Tyr Val Ile Gln Thr Thr Arg Asn Gly Val Arg
 340 345 350

Leu Cys Leu Ser Gly Phe Gln Ala Leu Asp Val Pro Pro Pro Ala Gly
355 360 365

Pro Phe Trp Ile Leu Gly Asp Val Phe Leu Gly Thr Tyr Val Ala Val
370 375 380

Phe Asp Arg Gly Asp Met Lys Ser Ser Ala Arg Val Gly Leu Ala Arg
385 390 395 400

Ala Arg Thr Arg Gly Ala Asp Leu Gly Trp Gly Glu Thr Ala Gln Ala
405 410 415

Gln Phe Pro Gly
420