



Office de la Propriété
Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2440563 A1 2002/09/19

(21) **2 440 563**

(12) **DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION**

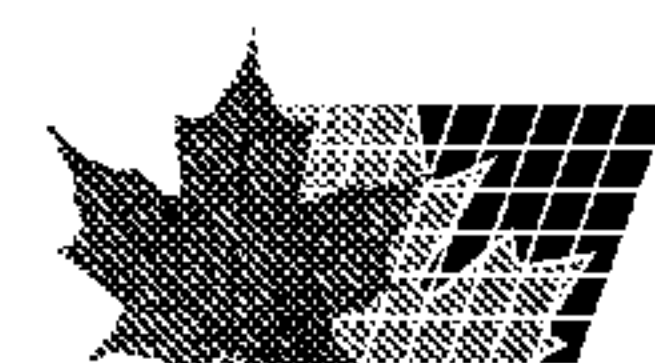
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2002/03/06	(51) Cl.Int. ⁷ /Int.Cl. ⁷ C12N 15/12, C12N 15/11, C07H 21/00
(87) Date publication PCT/PCT Publication Date: 2002/09/19	(71) Demandeur/Applicant: LEXICON GENETICS INCORPORATED, US
(85) Entrée phase nationale/National Entry: 2003/09/10	(72) Inventeurs/Inventors: YU, XUANCHUAN, US; MIRANDA, MARICAR, US
(86) N° demande PCT/PCT Application No.: US 2002/007477	(74) Agent: OGILVY RENAULT
(87) N° publication PCT/PCT Publication No.: 2002/072611	
(30) Priorité/Priority: 2001/03/12 (60/275,013) US	

(54) Titre : NOUVELLES PROTEINES HUMAINES DE LA FAMILLE EGF ET POLYNUCLEOTIDES CODANT POUR CES PROTEINES
(54) Title: NOVEL HUMAN EGF-FAMILY PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME

(57) **Abrégé/Abstract:**

Novel human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and pharmacogenomic applications.



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 September 2002 (19.09.2002)

PCT
(10) International Publication Number
WO 02/072611 A2

- (51) International Patent Classification⁷: **C07K**

(21) International Application Number: PCT/US02/07477

(22) International Filing Date: 6 March 2002 (06.03.2002)

(25) Filing Language: English

(26) Publication Language: English

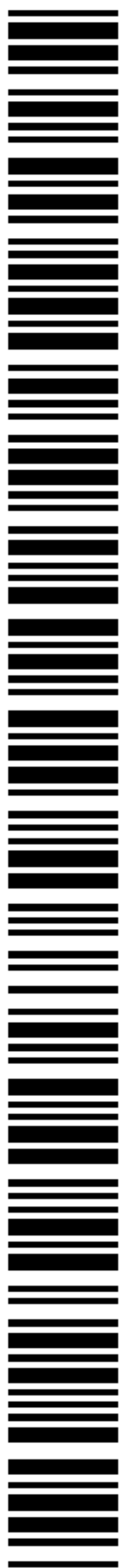
(30) Priority Data:
60/275,013 12 March 2001 (12.03.2001) US

(71) Applicant: **LEXICON GENETICS INCORPORATED**
[US/US]; 4000 Research Forest Drive, The Woodlands, TX 77381 (US).
- (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, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, 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, GQ, GW, ML, MR, NE, SN, TD, TG).
- (72) Inventors: **YU, Xuanchuan**; 6 Danville Crossing, Conroe, TX 77385 (US). **MIRANDA, Maricar**; 14814 Belterraza Drive, Houston, TX 77083 (US).

(74) Agent: **ISHIMOTO, Lance, K.**; Lexicon Genetics Incorporated, 4000 Research Forest Drive, The Woodlands, TX 77381 (US).
- 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.



WO 02/072611 A2

- (54) Title: NOVEL HUMAN EGF-FAMILY PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME
- (57) Abstract: Novel human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and pharmacogenomic applications.

NOVEL HUMAN EGF-FAMILY PROTEINS AND
POLYNUCLEOTIDES ENCODING THE SAME

The present application claims the benefit of U.S. Provisional Application Number 60/275,013 which was filed on March 12, 2001 and is herein incorporated by reference in its entirety.

1. INTRODUCTION

The present invention relates to the discovery, identification, and characterization of novel human polynucleotides encoding proteins that share sequence similarity with mammalian membrane proteins. The invention encompasses the described polynucleotides, host cell expression systems, the encoded proteins, fusion proteins, polypeptides and peptides, antibodies to the encoded proteins and peptides, and genetically engineered animals that either lack or overexpress the disclosed genes, antagonists and agonists of the proteins, and other compounds that modulate the expression or activity of the proteins encoded by the disclosed genes, which can be used for diagnosis, drug screening, clinical trial monitoring, the treatment of diseases and disorders, and cosmetic or nutraceutical applications.

2. BACKGROUND OF THE INVENTION

Membrane proteins can serve as recognition markers, mediate signal transduction, can serve as receptors for ligands, and can mediate or facilitate the passage of materials across the lipid bilayer. As such, membrane proteins, are proven drug targets.

3. SUMMARY OF THE INVENTION

The present invention relates to the discovery, identification, and characterization of nucleotides that encode novel human proteins and the corresponding amino acid sequences of these proteins. The novel human proteins (NHPs) described for the first time herein share structural similarity with mammalian proteins of the epidermal growth factor (EGF) family and notch proteins, among others.

The novel human nucleic acid sequences described herein encode alternative proteins/open reading frames (ORFs) of 1140 and 586 amino acids in length (SEQ ID NOS: 2 and 4).

The invention also encompasses agonists and antagonists of the described NHPs, including small molecules, large molecules, mutant NHPs, or portions thereof, that compete with native NHP, peptides, and antibodies, as well as nucleotide sequences that can be used to inhibit the expression of the described NHPs (e.g., antisense and ribozyme molecules, and open reading frame or regulatory sequence replacement constructs) or to enhance the expression of the described NHPs (e.g., expression constructs that place the described polynucleotide under the control of a strong promoter system), and transgenic animals that express a NHP sequence, or "knock-outs" (which can be conditional) that do not express a functional NHP. Knock-out mice can be produced in several ways, one of which involves the use of mouse embryonic stem cells ("ES cells") lines that contain gene trap mutations in a murine homolog of at least one of the described NHPs. When the unique NHP sequences described in SEQ ID NOS:1-4 are "knocked-out" they provide a method of identifying phenotypic expression of the particular gene as well as a method of assigning function to previously unknown genes. In addition, animals in which the unique NHP sequences described in SEQ ID NOS:1-4 are "knocked-

out" provide a unique source in which to elicit antibodies to homologous and orthologous proteins which would have been previously viewed by the immune system as "self" and therefore would have failed to elicit significant antibody responses.

5 Additionally, the unique NHP sequences described in SEQ ID NOS:1-4 are useful for the identification of protein coding sequence and mapping a unique gene to a particular chromosome. These sequences identify actual, biologically verified, and therefore relevant, exon splice junctions as opposed to those
10 that may have been bioinformatically predicted from genomic sequence alone. The sequences of the present invention are also useful as additional DNA markers for restriction fragment length polymorphism (RFLP) analysis, and in forensic biology.

 Further, the present invention also relates to processes
15 for identifying compounds that modulate, *i.e.*, act as agonists or antagonists, of NHP expression and/or NHP activity that utilize purified preparations of the described NHPs and/or NHP product, or cells expressing the same. Such compounds can be used as therapeutic agents for the treatment of any of a wide
20 variety of symptoms associated with biological disorders or imbalances.

4. DESCRIPTION OF THE SEQUENCE LISTING AND FIGURES

 The Sequence Listing provides the sequences of the ORFs
25 encoding the described NHP amino acid sequences.

5. DETAILED DESCRIPTION OF THE INVENTION

 The NHPs described for the first time herein are novel proteins that can be expressed in, *inter alia*, human cell
30 lines, prostate, pituitary, skeletal muscle, fetal brain, brain, adrenal gland, and osteosarcoma cells.

The present invention encompasses the nucleotides presented in the Sequence Listing, host cells expressing such nucleotides, the expression products of such nucleotides, and: (a) nucleotides that encode mammalian homologs of the described genes, including the specifically described NHPs, and the NHP products; (b) nucleotides that encode one or more portions of the NHPs that correspond to functional domains, and the polypeptide products specified by such nucleotide sequences including, but not limited to, the novel regions of any active domain(s); (c) isolated nucleotides that encode mutant versions, engineered or naturally occurring, of the described NHPs in which all or a part of at least one domain is deleted or altered, and the polypeptide products specified by such nucleotide sequences including, but not limited to, soluble proteins and peptides in which all or a portion of the signal (or hydrophobic transmembrane) sequence is deleted; (d) nucleotides that encode chimeric fusion proteins containing all or a portion of a coding region of an NHP, or one of its domains (e.g., a receptor or ligand binding domain, accessory protein/self-association domain, etc.) fused to another peptide or polypeptide; or (e) therapeutic or diagnostic derivatives of the described polynucleotides such as oligonucleotides, antisense polynucleotides, ribozymes, dsRNA, or gene therapy constructs comprising a sequence first disclosed in the Sequence Listing.

As discussed above, the present invention includes: (a) the human DNA sequences presented in the Sequence Listing (and vectors comprising the same) and additionally contemplates any nucleotide sequence encoding a contiguous NHP open reading frame (ORF) that hybridizes to a complement of a DNA sequence presented in the Sequence Listing under highly stringent conditions, e.g., hybridization to filter-bound DNA in 0.5 M

NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., NY, at p. 2.10.3) and encodes a functionally equivalent expression product.

Additionally contemplated are any nucleotide sequences that hybridize to the complement of a DNA sequence that encodes and expresses an amino acid sequence presented in the Sequence Listing under moderately stringent conditions, e.g., washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989, *supra*), yet still encodes a functionally equivalent NHP product. Functional equivalents of a NHP include naturally occurring NHPs present in other species and mutant NHPs whether naturally occurring or engineered (by site directed mutagenesis, gene shuffling, directed evolution as described in, for example, U.S. Patent No. 5,837,458). The invention also includes degenerate nucleic acid variants of the disclosed NHP polynucleotide sequences.

Additionally contemplated are polynucleotides encoding NHP ORFs, or their functional equivalents, encoded by polynucleotide sequences that are about 99, 95, 90, or about 85 percent similar or identical to corresponding regions of the nucleotide sequences of the Sequence Listing (as measured by BLAST sequence comparison analysis using, for example, the GCG sequence analysis package using standard default settings).

The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the described NHP gene nucleotide sequences. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances where the nucleic acid molecules are

deoxyoligonucleotides ("DNA oligos"), such molecules are generally about 16 to about 100 bases long, or about 20 to about 80, or about 34 to about 45 bases long, or any variation or combination of sizes represented therein that incorporate a contiguous region of sequence first disclosed in the Sequence Listing. Such oligonucleotides can be used in conjunction with the polymerase chain reaction (PCR) to screen libraries, isolate clones, and prepare cloning and sequencing templates, etc.

10 Alternatively, such NHP oligonucleotides can be used as hybridization probes for screening libraries, and assessing gene expression patterns (particularly using a micro array or high-throughput "chip" format). Additionally, a series of the described NHP oligonucleotide sequences, or the complements
15 thereof, can be used to represent all or a portion of the described NHP sequences. An oligonucleotide or polynucleotide sequence first disclosed in at least a portion of one or more of the sequences of SEQ ID NOS: 1-4 can be used as a hybridization probe in conjunction with a solid support
20 matrix/substrate (resins, beads, membranes, plastics, polymers, metal or metallized substrates, crystalline or polycrystalline substrates, etc.). Of particular note are spatially addressable arrays (*i.e.*, gene chips, microtiter plates, etc.) of oligonucleotides and polynucleotides, or corresponding
25 oligopeptides and polypeptides, wherein at least one of the biopolymers present on the spatially addressable array comprises an oligonucleotide or polynucleotide sequence first disclosed in at least one of the sequences of SEQ ID NOS: 1-4, or an amino acid sequence encoded thereby. Methods for
30 attaching biopolymers to, or synthesizing biopolymers on, solid support matrices, and conducting binding studies thereon are disclosed in, *inter alia*, U.S. Patent Nos. 5,700,637,

5,556,752, 5,744,305, 4,631,211, 5,445,934, 5,252,743, 4,713,326, 5,424,186, and 4,689,405 the disclosures of which are herein incorporated by reference in their entirety.

Addressable arrays comprising sequences first disclosed
5 in SEQ ID NOS:1-4 can be used to identify and characterize the temporal and tissue specific expression of a gene. These addressable arrays incorporate oligonucleotide sequences of sufficient length to confer the required specificity, yet be within the limitations of the production technology. The
10 length of these probes is within a range of between about 8 to about 2000 nucleotides. Preferably the probes consist of 60 nucleotides and more preferably 25 nucleotides from the sequences first disclosed in SEQ ID NOS:1-4.

For example, a series of the described oligonucleotide
15 sequences, or the complements thereof, can be used in chip format to represent all or a portion of the described sequences. The oligonucleotides, typically between about 16 to about 40 (or any whole number within the stated range) nucleotides in length can partially overlap each other and/or
20 the sequence may be represented using oligonucleotides that do not overlap. Accordingly, the described polynucleotide sequences shall typically comprise at least about two or three distinct oligonucleotide sequences of at least about 8 nucleotides in length that are each first disclosed in the
25 described Sequence Listing. Such oligonucleotide sequences can begin at any nucleotide present within a sequence in the Sequence Listing and proceed in either a sense (5'-to-3') orientation vis-a-vis the described sequence or in an antisense orientation.

30 Microarray-based analysis allows the discovery of broad patterns of genetic activity, providing new understanding of gene functions and generating novel and unexpected insight into

transcriptional processes and biological mechanisms. The use of addressable arrays comprising sequences first disclosed in SEQ ID NOS:1-4 provides detailed information about transcriptional changes involved in a specific pathway, potentially leading to the identification of novel components or gene functions that manifest themselves as novel phenotypes.

Probes consisting of sequences first disclosed in SEQ ID NOS:1-4 can also be used in the identification, selection and validation of novel molecular targets for drug discovery. The use of these unique sequences permits the direct confirmation of drug targets and recognition of drug dependent changes in gene expression that are modulated through pathways distinct from the drugs intended target. These unique sequences therefore also have utility in defining and monitoring both drug action and toxicity.

As an example of utility, the sequences first disclosed in SEQ ID NOS:1-4 can be utilized in microarrays or other assay formats, to screen collections of genetic material from patients who have a particular medical condition. These investigations can also be carried out using the sequences first disclosed in SEQ ID NOS:1-4 *in silico* and by comparing previously collected genetic databases and the disclosed sequences using computer software known to those in the art.

Thus the sequences first disclosed in SEQ ID NOS:1-4 can be used to identify mutations associated with a particular disease and also as a diagnostic or prognostic assay.

Although the presently described sequences have been specifically described using nucleotide sequence, it should be appreciated that each of the sequences can uniquely be described using any of a wide variety of additional structural attributes, or combinations thereof. For example, a given sequence can be described by the net composition of the

nucleotides present within a given region of the sequence in conjunction with the presence of one or more specific oligonucleotide sequence(s) first disclosed in the SEQ ID NOS: 1-4. Alternatively, a restriction map specifying the relative positions of restriction endonuclease digestion sites, or various palindromic or other specific oligonucleotide sequences can be used to structurally describe a given sequence. Such restriction maps, which are typically generated by widely available computer programs (e.g., the University of Wisconsin GCG sequence analysis package, SEQUENCHER 3.0, Gene Codes Corp., Ann Arbor, MI, etc.), can optionally be used in conjunction with one or more discrete nucleotide sequence(s) present in the sequence that can be described by the relative position of the sequence relative to one or more additional sequence(s) or one or more restriction sites present in the disclosed sequence.

For oligonucleotide probes, highly stringent conditions may refer, e.g., to washing in 6x SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as NHP gene antisense molecules, useful, for example, in NHP gene regulation and/or as antisense primers in amplification reactions of NHP gene nucleic acid sequences. With respect to NHP gene regulation, such techniques can be used to regulate biological functions. Further, such sequences may be used as part of ribozyme and/or triple helix sequences that are also useful for NHP gene regulation.

Inhibitory antisense or double stranded oligonucleotides can additionally comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil,

hypoxanthine, xanthine, 4-acetylcytosine,
 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-
 2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil,
 beta-D-galactosylqueosine, inosine, N6-isopentenyladenine,
 5 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine,
 2-methyladenine, 2-methylguanine, 3-methylcytosine,
 5-methylcytosine, N6-adenine, 7-methylguanine,
 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil,
 beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil,
 10 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-
 oxyacetic acid (v), wybutoxosine, pseudouracil, queosine,
 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil,
 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid
 methylester, uracil-5-oxyacetic acid (v), 5-methyl-
 15 2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w,
 and 2,6-diaminopurine.

The antisense oligonucleotide can also comprise at least
 one modified sugar moiety selected from the group including,
 but not limited to, arabinose, 2-fluoroarabinose, xylulose, and
 20 hexose.

In yet another embodiment, the antisense oligonucleotide
 will comprise at least one modified phosphate backbone selected
 from the group including, but not limited to, a
 phosphorothioate, a phosphorodithioate, a phosphoramidothioate,
 25 a phosphoramidate, a phosphordiamidate, a methylphosphonate, an
 alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide
 is an α -anomeric oligonucleotide. An α -anomeric
 oligonucleotide forms specific double-stranded hybrids with
 30 complementary RNA in which, contrary to the usual β -units, the
 strands run parallel to each other (Gautier et al., 1987, Nucl.
 Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-

methylnucleotide (Inoue *et al.*, 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, FEBS Lett. 215:327-330). Alternatively, double stranded RNA can be used to disrupt the expression and function of a
5 targeted NHP.

Oligonucleotides of the invention can be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples,
10 phosphorothioate oligonucleotides can be synthesized by the method of Stein *et al.* (1988, Nucl. Acids Res. 16:3209), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

15 Low stringency conditions are well-known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook *et al.*, 1989, Molecular Cloning, A
20 Laboratory Manual (and periodic updates thereof), Cold Springs Harbor Press, N.Y.; and Ausubel *et al.*, 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Alternatively, suitably labeled NHP nucleotide probes can
25 be used to screen a human genomic library using appropriately stringent conditions or by PCR. The identification and characterization of human genomic clones is helpful for identifying polymorphisms (including, but not limited to, nucleotide repeats, microsatellite alleles, single nucleotide
30 polymorphisms, or coding single nucleotide polymorphisms), determining the genomic structure of a given locus/allele, and designing diagnostic tests. For example, sequences derived

from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (e.g., splice acceptor and/or donor sites), etc., that can be used in diagnostics and pharmacogenomics.

In another example, the present sequences can be used in restriction fragment length polymorphism (RFLP) analysis to identify specific individuals. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification (as generally described in U.S. Patent No. 5,272,057, incorporated herein by reference). In addition, the sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e., another DNA sequence that is unique to a particular individual). Actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments.

Further, a NHP gene homolog can be isolated from nucleic acid from an organism of interest by performing PCR using two degenerate or "wobble" oligonucleotide primer pools designed on the basis of amino acid sequences within the NHP products disclosed herein. The template for the reaction may be total RNA, mRNA, and/or cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express an allele of a NHP gene. The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequence of the desired NHP

gene. The PCR fragment can then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment can be labeled and used to screen a cDNA library, such as a bacteriophage cDNA library. Alternatively, 5 the labeled fragment can be used to isolate genomic clones via the screening of a genomic library.

PCR technology can also be used to isolate full length cDNA sequences. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue 10 source (*i.e.*, one known, or suspected, to express a NHP gene). A reverse transcription (RT) reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" 15 using a standard terminal transferase reaction, the hybrid may be digested with RNase H, and second strand synthesis may then be primed with a complementary primer. Thus, cDNA sequences upstream of the amplified fragment can be isolated. For a review of cloning strategies that can be used, see *e.g.*, 20 Sambrook *et al.*, 1989, *supra*.

A cDNA encoding a mutant NHP sequence can be isolated, for example, by using PCR. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed 25 in an individual putatively carrying a mutant NHP allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal sequence. Using these two primers, the product is 30 then amplified via PCR, optionally cloned into a suitable vector, and subjected to DNA sequence analysis through methods well-known to those of skill in the art. By comparing the DNA

sequence of the mutant NHP allele to that of a corresponding normal NHP allele, the mutation(s) responsible for the loss or alteration of function of the mutant NHP gene product can be ascertained.

5 Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant NHP allele (e.g., a person manifesting a NHP-associated phenotype such as, for example, osteoporosis, obesity, high blood pressure, connective tissue disorders,
10 infertility, etc.), or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express a mutant NHP allele. A normal NHP gene, or any suitable fragment thereof, can then be labeled and used as a probe to identify the corresponding mutant NHP allele in such libraries. Clones
15 containing mutant NHP sequences can then be purified and subjected to sequence analysis according to methods well-known to those skilled in the art.

 Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from
20 a tissue known, or suspected, to express a mutant NHP allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue can be expressed and screened using standard antibody screening techniques in conjunction with antibodies
25 raised against a normal NHP product, as described below. For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor, NY.

 Additionally, screening can be accomplished by screening
30 with labeled NHP fusion proteins, such as, for example, alkaline phosphatase-NHP or NHP-alkaline phosphatase fusion proteins. In cases where a NHP mutation results in an

expression product with altered function (e.g., as a result of a missense or a frameshift mutation), polyclonal antibodies to NHP are likely to cross-react with a corresponding mutant NHP expression product. Library clones detected via their reaction
5 with such labeled antibodies can be purified and subjected to sequence analysis according to methods well-known in the art.

The invention also encompasses (a) DNA vectors that contain any of the foregoing NHP coding sequences and/or their complements (*i.e.*, antisense); (b) DNA expression vectors that
10 contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences (for example, baculovirus as described in U.S. Patent No. 5,869,336 herein incorporated by reference); (c) genetically engineered host cells that contain
15 any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell; and (d) genetically engineered host cells that express an endogenous NHP sequence under the control of an exogenously introduced
20 regulatory element (*i.e.*, gene activation). As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Such regulatory elements include, but are
25 not limited to, the cytomegalovirus (hCMV) immediate early gene, regulatable, viral elements (particularly retroviral LTR promoters), the early or late promoters of SV40 adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage lambda, the
30 control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase (PGK), the promoters of acid phosphatase, and the promoters of the yeast α -mating factors.

The present invention also encompasses antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists and agonists of a NHP, as well as compounds or nucleotide constructs that inhibit expression of a NHP sequence
5 (transcription factor inhibitors, antisense and ribozyme molecules, or open reading frame sequence or regulatory sequence replacement constructs), or promote the expression of a NHP (e.g., expression constructs in which NHP coding sequences are operatively associated with expression control
10 elements such as promoters, promoter/enhancers, etc.).

The NHPs or NHP peptides, NHP fusion proteins, NHP nucleotide sequences, antibodies, antagonists and agonists can be useful for the detection of mutant NHPs or inappropriately expressed NHPs for the diagnosis of disease. The NHP proteins
15 or peptides, NHP fusion proteins, NHP nucleotide sequences, host cell expression systems, antibodies, antagonists, agonists and genetically engineered cells and animals can be used for screening for drugs (or high throughput screening of combinatorial libraries) effective in the treatment of the
20 symptomatic or phenotypic manifestations of perturbing the normal function of NHP in the body. The use of engineered host cells and/or animals may offer an advantage in that such systems allow not only for the identification of compounds that bind to the endogenous receptor for an NHP, but can also
25 identify compounds that trigger NHP-mediated activities or pathways.

Finally, the NHP products can be used as therapeutics. For example, soluble derivatives such as NHP peptides/domains corresponding to NHPs, NHP fusion protein products (especially
30 NHP-Ig fusion proteins, *i.e.*, fusions of a NHP, or a domain of a NHP, to an IgFc), NHP antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists or agonists

(including compounds that modulate or act on downstream targets in a NHP-mediated pathway) can be used to directly treat diseases, including infectious diseases, or disorders. For instance, the administration of an effective amount of soluble NHP, or a NHP-IgFc fusion protein or an anti-idiotypic antibody (or its Fab) that mimics the NHP could activate or effectively antagonize the endogenous NHP receptor. Nucleotide constructs encoding such NHP products can be used to genetically engineer host cells to express such products *in vivo*; these genetically engineered cells function as "bioreactors" in the body delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion protein to the body. Nucleotide constructs encoding functional NHPs, mutant NHPs, as well as antisense and ribozyme molecules can also be used in "gene therapy" approaches for the modulation of NHP expression. Thus, the invention also encompasses pharmaceutical formulations and methods for treating biological disorders.

Various aspects of the invention are described in greater detail in the subsections below.

20

5.1 THE NHP SEQUENCES

The cDNA sequences and the corresponding deduced amino acid sequences of the described NHPs are presented in the Sequence Listing. The NHP nucleotides were obtained from clustered genomic sequence, ESTs, and cDNAs produced using human prostate, pituitary, skeletal muscle, fetal brain, brain, adrenal gland mRNAs (Edge Biosystems, Gaithersburg, MD). Sequences corresponding to portions of the presently described NHP sequences have been described, and it is possible that such sequences were incomplete sequences that had been erroneously characterized as "full length."

30

The gene encoding the described NHPs is apparently encoded on human chromosome 5 (see GENBANK accession number AC026800). As such the described sequences can also be used to map the corresponding coding regions of the human genome.

5 Several polymorphisms were identified including an A/G polymorphism at the nucleotide position represented by, for example, position 1601 of SEQ ID NO:1 (which can result in a tyr or cys at the region corresponding to amino acid (aa) position 534 of, for example, SEQ ID NO:2), a T/G polymorphism
10 at nucleotide position 1704 (which can result in a cys or trp at corresponding aa position 568), a G/T polymorphism at the nucleotide position represented by, for example, position 1709 of SEQ ID NO:1 (which can result in a ser or ile at the region corresponding to aa position 600 of, for example, SEQ ID NO:2),
15 and a T/G polymorphism at nucleotide position 1712 (which can result in a val or gly at aa position 601). The present invention contemplates sequences comprising any and all combinations and permutations of the above polymorphisms.

 The described novel human polynucleotide sequences can be
20 used, among other things, in the molecular mutagenesis/evolution of proteins that are at least partially encoded by the described novel sequences using, for example, polynucleotide shuffling or related methodologies. Such approaches are described in U.S. Patent Nos. 5,830,721 and
25 5,837,458 which are herein incorporated by reference in their entirety.

 NHP gene products can also be expressed in transgenic animals. Animals of any species including, but not limited to, worms, mice, rats, rabbits, guinea pigs, pigs, micro-pigs,
30 birds, goats, and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate NHP transgenic animals.

Any technique known in the art may be used to introduce a NHP transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Patent No. 4,873,191); retrovirus-mediated gene transfer into germ lines (Van der Putten *et al.*, 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (Thompson *et al.*, 1989, Cell 56:313-321); electroporation of embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814); and sperm-mediated gene transfer (Lavitrano *et al.*, 1989, Cell 57:717-723); *etc.* For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229, which is incorporated by reference herein in its entirety.

The present invention provides for transgenic animals that carry the NHP transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.*, mosaic animals or somatic cell transgenic animals. The transgene may be integrated as a single transgene or in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell-type by following, for example, the teaching of Lasko *et al.*, 1992, Proc. Natl. Acad. Sci. USA 89:6232-6236. The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell-type of interest, and will be apparent to those of skill in the art.

When it is desired that a NHP transgene be integrated into the chromosomal site of the endogenous NHP gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous NHP gene are designed for the

purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous NHP gene (*i.e.*, "knockout" animals).

5 The transgene can also be selectively introduced into a particular cell-type, thus inactivating the endogenous NHP gene in only that cell-type, by following, for example, the teaching of Gu *et al.*, 1994, *Science*, 265:103-106. The regulatory sequences required for such a cell-type specific inactivation
10 will depend upon the particular cell-type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant NHP gene may be assayed utilizing standard techniques. Initial screening may be accomplished by
15 Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot
20 analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR. Samples of NHP gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the NHP transgene product.

The present invention also provides for "knock-in"
25 animals. Knock-in animals are those in which a polynucleotide sequence (*i.e.*, a gene or a cDNA) that the animal does not naturally have in its genome is inserted in such a way that the sequence is expressed. Examples include, but are not limited to, a human gene or cDNA used to replace its murine ortholog in
30 the mouse, a murine cDNA used to replace the murine gene in the mouse, and a human gene or cDNA or murine cDNA that is tagged with a reporter construct used to replace the murine ortholog

or gene in the mouse. Such replacements can occur at the locus of the murine ortholog or gene, or at another specific site. Such knock-in animals are useful for the *in vivo* study, testing and validation of, *intra alia*, human drug targets, as well as
5 for compounds that are directed at the same and therapeutic proteins.

5.2 NHPS AND NHP POLYPEPTIDES

NHPs, NHP polypeptides, peptide fragments, mutated,
10 truncated, or deleted forms of the NHPs, and/or NHP fusion proteins can be prepared for a variety of uses. These uses include, but are not limited to, therapeutic products, the generation of antibodies, as reagents in diagnostic assays, the identification of other cellular gene products related to a
15 NHP, as reagents in assays for screening for compounds that can be used as pharmaceutical reagents useful in the therapeutic treatment of mental, biological, or medical disorders and diseases. Given the similarity information and expression data, the described NHPs can be targeted (by drugs, oligos,
20 antibodies, etc,) in order to diagnose or treat disease, or to therapeutically supplant or augment the efficacy of, for example, chemotherapeutic agents used in the treatment of cancer, inflammatory disorders, arthritis, or to promote healing or treat infectious disease.

25 The Sequence Listing discloses the amino acid sequences encoded by the described NHP sequences. The NHPs display initiator methionines in DNA sequence contexts consistent with translation initiation sites, and prototypical signal sequences often associated with secreted or membrane-associated proteins.

30 The NHP amino acid sequences of the invention include the amino acid sequence presented in the Sequence Listing as well as analogues and derivatives thereof. Further, corresponding

NHP homologues from other species are encompassed by the invention. In fact, any NHP protein encoded by the NHP nucleotide sequences described above are within the scope of the invention as are any novel polynucleotide sequences
5 encoding all or any novel portion of an amino acid sequence presented in the Sequence Listing. The degenerate nature of the genetic code is well-known, and, accordingly, each amino acid presented in the Sequence Listing, is generically representative of the well-known nucleic acid "triplet" codon,
10 or in many cases codons, that can encode the amino acid. As such, as contemplated herein, the amino acid sequences presented in the Sequence Listing, when taken together with the genetic code (see, for example, Table 4-1 at page 109 of "Molecular Cell Biology", 1986, J. Darnell et al. eds.,
15 Scientific American Books, New York, NY, herein incorporated by reference) are generically representative of all the various permutations and combinations of nucleic acid sequences that can encode such amino acid sequences.

The invention also encompasses proteins that are
20 functionally equivalent to the NHPs encoded by the presently described nucleotide sequences as judged by any of a number of criteria, including, but not limited to, the ability to bind and cleave a substrate of a NHP, or the ability to effect an identical or complementary downstream pathway, or a change in
25 cellular metabolism (e.g., proteolytic activity, ion flux, tyrosine phosphorylation, etc.). Such functionally equivalent NHP proteins include, but are not limited to, additions or substitutions of amino acid residues within the amino acid sequence encoded by the NHP nucleotide sequences described
30 above, but which result in a silent change, thus producing a functionally equivalent expression product. Amino acid substitutions may be made on the basis of similarity in

polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

10 A variety of host-expression vector systems can be used to express the NHP nucleotide sequences of the invention. Where, as in the present instance, the NHP peptide or polypeptide is thought to be membrane protein, the hydrophobic regions of the protein can be excised and the resulting soluble peptide or polypeptide can be recovered from the culture media. 15 Such expression systems also encompass engineered host cells that express a NHP, or functional equivalent, *in situ*. Purification or enrichment of a NHP from such expression systems can be accomplished using appropriate detergents and 20 lipid micelles and methods well-known to those skilled in the art. However, such engineered host cells themselves may be used in situations where it is important not only to retain the structural and functional characteristics of the NHP, but to assess biological activity, e.g., in certain drug screening 25 assays.

The expression systems that may be used for purposes of the invention include, but are not limited to, microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing NHP nucleotide sequences; yeast 30 (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing NHP nucleotide sequences;

insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing NHP nucleotide sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing NHP nucleotide sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing NHP nucleotide sequences and promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the NHP product being expressed. For example, when a large quantity of such a protein is to be produced for the generation of pharmaceutical compositions of or containing NHP, or for raising antibodies to a NHP, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited to, the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which a NHP coding sequence may be ligated individually into the vector in frame with the *lacZ* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Pharmacia or American Type Culture Collection) can also 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.

The PGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target expression product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign polynucleotide sequences. The virus grows in *Spodoptera frugiperda* cells. A NHP coding sequence can be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of NHP coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted sequence is expressed (*e.g.*, see Smith *et al.*, 1983, J. Virol. 46: 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the NHP nucleotide sequence of interest may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric sequence may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a NHP product in infected hosts (*e.g.*, See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted NHP nucleotide sequences. These

signals include the ATG initiation codon and adjacent sequences. In cases where an entire NHP gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional
5 translational control signals may be needed. However, in cases where only a portion of a NHP coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the
10 desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements,
15 transcription terminators, etc. (See Bitter et al., 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the expression product in the specific fashion
20 desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and
25 expression products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and
30 phosphorylation of the expression product may be used. Such mammalian host cells include, but are not limited to, CHO,

VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in particular, human cell lines.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell
5 lines which stably express the NHP sequences described above can be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer sequences,
10 transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers
15 resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the NHP product. Such engineered cell lines may
20 be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the NHP product.

A number of selection systems may be used including, but not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, 1977, *Cell* 11:223), hypoxanthine-guanine
25 phosphoribosyltransferase (Szybalska and Szybalski, 1962, *Proc. Natl. Acad. Sci. USA* 48:2026), and adenine phosphoribosyltransferase (Lowy *et al.*, 1980, *Cell* 22:817) genes, which can be employed in tk⁻, hgprt⁻ or aprt⁻ cells, respectively. Also, antimetabolite resistance can be used as
30 the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler *et al.*, 1980, *Proc. Natl. Acad. Sci. USA* 77:3567; O'Hare *et al.*, 1981, *Proc. Natl.*

Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin et al., 1981, J. Mol. Biol. 150:1); and hygro, which confers resistance to hygromycin (Santerre et al., 1984, Gene 30:147).

Alternatively, any fusion protein can be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-8976). In this system, the sequence of interest is subcloned into a vaccinia recombination plasmid such that the sequence's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni²⁺-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

Also encompassed by the present invention are fusion proteins that direct the NHP to a target organ and/or facilitate transport across the membrane into the cytosol. Conjugation of NHPs to antibody molecules or their Fab fragments could be used to target cells bearing a particular epitope. Attaching the appropriate signal sequence to the NHP would also transport the NHP to the desired location within the cell. Alternatively targeting of NHP or its nucleic acid sequence might be achieved using liposome or lipid complex based delivery systems. Such technologies are described in "Liposomes: A Practical Approach", New, R.R.C., ed., Oxford University Press, New York and in U.S. Patent Nos. 4,594,595,

5,459,127, 5,948,767 and 6,110,490 and their respective disclosures which are herein incorporated by reference in their entirety. Additionally embodied are novel protein constructs engineered in such a way that they facilitate transport of the NHP to the target site or desired organ, where they cross the cell membrane and/or the nucleus where the NHP can exert its functional activity. This goal may be achieved by coupling of the NHP to a cytokine or other ligand that provides targeting specificity, and/or to a protein transducing domain (see generally U.S. applications Ser. No. 60/111,701 and 60/056,713, both of which are herein incorporated by reference, for examples of such transducing sequences) to facilitate passage across cellular membranes and can optionally be engineered to include nuclear localization.

15

5.3 ANTIBODIES TO NHP PRODUCTS

Antibodies that specifically recognize one or more epitopes of a NHP, or epitopes of conserved variants of a NHP, or peptide fragments of a NHP are also encompassed by the invention. Such antibodies include, but are not limited to, polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The antibodies of the invention may be used, for example, in the detection of NHP in a biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal amounts of NHP. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes for the evaluation of the effect of test compounds on expression and/or

activity of a NHP expression product. Additionally, such antibodies can be used in conjunction gene therapy to, for example, evaluate the normal and/or engineered NHP-expressing cells prior to their introduction into the patient. Such
5 antibodies may additionally be used as a method for the inhibition of abnormal NHP activity. Thus, such antibodies may, therefore, be utilized as part of treatment methods.

For the production of antibodies, various host animals may be immunized by injection with a NHP, an NHP peptide (e.g.,
10 one corresponding to a functional domain of an NHP), truncated NHP polypeptides (NHP in which one or more domains have been deleted), functional equivalents of the NHP or mutated variant of the NHP. Such host animals may include, but are not limited to, pigs, rabbits, mice, goats, and rats, to name but a few.
15 Various adjuvants may be used to increase the immunological response, depending on the host species, including, but not limited to, Freund's adjuvant (complete and incomplete), mineral salts such as aluminum hydroxide or aluminum phosphate, chitosan, surface active substances such as lysolecithin,
20 pluronic polyols, polyanions, peptides, oil emulsions, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Alternatively, the immune response could be enhanced by combination and or coupling with molecules such as keyhole limpet hemocyanin,
25 tetanus toxoid, diphtheria toxoid, ovalbumin, cholera toxin or fragments thereof. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Monoclonal antibodies, which are homogeneous populations
30 of antibodies to a particular antigen, can be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include,

but are not limited to, the hybridoma technique of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor *et al.*, 1983, Immunology Today 4:72; Cole *et al.*, 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole *et al.*, 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. Production of high titers of mAbs *in vivo* makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger *et al.*, 1984, Nature, 312:604-608; Takeda *et al.*, 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. Such technologies are described in U.S. Patent Nos. 6,075,181 and 5,877,397 and their respective disclosures which are herein incorporated by reference in their entirety. Also encompassed by the present invention is the use of fully humanized monoclonal antibodies as described in US Patent No. 6,150,584 and respective disclosures which are herein incorporated by reference in their entirety.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242:423-426; Huston *et al.*, 1988, Proc. Natl. Acad.

Sci. USA 85:5879-5883; and Ward *et al.*, 1989, Nature 341:544-546) can be adapted to produce single chain antibodies against NHP expression products.

Alternatively, techniques described for the production of
5 single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242:423-426; Huston *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward *et al.*, 1989, Nature 341:544-546) can be adapted to produce single chain antibodies against NHP gene products. Single chain antibodies are formed by
10 linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments
15 include, but are not limited to: the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed (Huse *et al.*, 1989,
20 Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibodies to a NHP can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" a given NHP, using techniques well-known to those skilled in the art. (See, *e.g.*,
25 Greenspan & Bona, 1993, FASEB J 7(5):437-444; and Nissinoff, 1991, J. Immunol. 147(8):2429-2438). For example antibodies which bind to a NHP domain and competitively inhibit the binding of NHP to its cognate receptor can be used to generate anti-idiotypes that "mimic" the NHP and, therefore, bind and
30 activate or neutralize a receptor. Such anti-idiotypic antibodies or Fab fragments of such anti-idiotypes can be used in therapeutic regimens involving a NHP-mediated pathway.

Additionally given the high degree of relatedness of mammalian NHPs, the presently described knock-out mice (having never seen NHP, and thus never been tolerized to NHP) have a unique utility, as they can be advantageously applied to the
5 generation of antibodies against the disclosed mammalian NHP (*i.e.*, NHP will be immunogenic in NHP knock-out animals).

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention,
10 and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall
15 within the scope of the appended claims. All cited publications, patents, and patent applications are herein incorporated by reference in their entirety.

SEQUENCE LISTING

<110> LEXICON GENETICS INCORPORATED

<120> Novel Human EGF-Family Proteins and Polynucleotides Encoding the Same

<130> LEX-0317-PCT

<150> US 60/275,013

<151> 2001-03-12

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3423

<212> DNA

<213> homo sapiens

<400> 1

```

atggttattt ctttgaactc atgcctgagc tttatttggt tattgttatg ccactggatt 60
gggacagcat cacctctgaa tcttgaagac cctaagtgtg gtagccactg ggaaagctac 120
tcagtgactg tgcaagagtc ataccacat ccctttgatc aaatttacta cacgagctgc 180
actgacattc taaactgggt taaatgcacg cggcacagag tcagctatcg gacagcctat 240
cgacatgggg agaagactat gtataggcgc aagtctcagt gttgtcctgg attttatgaa 300
agcggggaaa tgtgtgtccc ccactgtgct gataaatgtg tccatggctg ctgtattgct 360
ccaaacacct gtcagtgtga gcctggctgg ggagggacca actgctccag tgcttgcgat 420
ggtgatcact ggggtcccca ctgcaccagc cggtgccagt gcaaaaatgg ggctctgtgc 480
aaccocatca ccggggcttg ccactgtgct gcgggcttcc ggggctggcg ctgcgaggac 540
cgctgtgagc agggcaccta tggtaacgac tgtcatcaga gatgccagtg ccagaatgga 600
gccacctgcg accacgtcac gggggaatgc cgctgcccac caggatacac cggagccttc 660
tgtgaggatc tttgtcctcc tggtaaacad ggtccacagt gtgagcagag atgcccttgt 720
caaaatggag gagtgtgtca tcacgtcact ggagaatgct cttgcccttc tggctggatg 780
ggcacagtgt gtggtcagcc ttgccccgag ggtcgctttg gaaagaactg ttccaagaa 840
tgccagtgcc ataatggagg gacgtgtgat gctgccacag gccaatgtca ttgcagtcca 900
ggatacacag gggaacggtg ccaggatgag tgtcctgttg ggacctatgg cgttctctgt 960
gctgagacct gccagtgtgt caacggaggg aagtgttacc acgtgagcgg cgcattgcctc 1020
tgtgaagcag gctttgctgg cgagcgctgc gaagcacgcc tgtgtcctga ggggctctac 1080
ggcatcaa atgtgacaaac gtgtccctgc cacttgaaa acactcatag ctgtcaccac 1140
atgtctggag agtgtgcctg caagccgggc tggtcaggac tctactgtaa tgagacatgt 1200
tctcctggat tctacgggga agcttgccag cagatctgca gctgccaaa tggggcagac 1260
tgtgacagtg tgactggaaa gtgcacctgt gcccaggat tcaaaggaat tgactgctct 1320
acccatgcc ctctgggaac ctatgggata aactgttctc ctgctgtggg ctgtaaaaat 1380
gatgcagtct gctctcctgt ggacgggtct tgtacttgca aggcaggctg gcacgggggtg 1440
gactgctcca tcagatgtcc cagtggcaca tggggctttg gctgtaactt aacatgccag 1500
tgctcaacg ggggagcctg caacaccctg gacgggacct gcacgtgtgc acctggatgg 1560
cgcggggaga aatgcgaact tcctgccag gatggcacgt acgggctgaa ctgtgctgag 1620
cgctgcgact gcagccacgc agatggctgc caccctacca cgggccattg ccgctgcctc 1680
ccgggatggt cagggtgtcca ctgtgacagc gtgtgtgtgt agggacgctg gggccccaac 1740
tgctccctgc cctgctactg taaaaatggg gcttcatgct ccctgatga tggcatctgc 1800
gagtgtgcac caggcttccg aggcaccact tgtcagagga tctgctcccc tggtttttat 1860
gggcatcgct gcagccagac atgcccacag tgcgttcaca gcagcgggac ctgccaccac 1920
atcaccggcc tgtgtgactg cttgcctggc ttcacaggcg ccctctgcaa tgaagtgtgt 1980
cccagtggca gatgtgggaa aaactgtgca ggaatttgta cctgcaccaa caacggaacc 2040

```

```

tgtaacccca ttgacagatc ttgtcagtgt taccgccggtt ggattggcag tgactgctct 2100
caacctatgtc cacctgcccc ctgggggcccc aactgcatcc acacgtgcaa ctgccataat 2160
ggagcttttct gcagcgccta cgatgggggaa tgtaaattgca ctcctggctg gacagggctc 2220
tactgcactc agagatgtcc tctaggggtt tatggaaaag attgtgact gatatgcca 2280
tgtcaaaacg gagctgactg cgaccacatt tctgggcagt gtacttgccg cactggattc 2340
atgggacggc actgtgagca gaagtgcctt tcaggaacat atggctatgg ctgtcgccag 2400
atatgtgatt gtctgaacaa ctccacctgc gaccacatca ctgggacctg ttactgcagc 2460
cccggatgga agggagcgag atgtgatcaa gctggtgtta tcatagttgg aaatctgaac 2520
agcttaagcc gaaccagtac tgctctccct gctgattcct accagatcgg ggccattgca 2580
ggcatcatca ttcttgcctt agttgttctc ttctactgg cattgttcat tatttataga 2640
cacaagcaga agggaaagga atcaagcatg ccagcagtta cctacacccc tgctatgagg 2700
gtcgtcaatg cagattatac catttcagga acccttcctc acagcaatgg tggaaacgct 2760
aatagccact acttcaccaa tcccagttac cacacgctca cccagtgtgc cacatcccct 2820
cacgtcaaca acagggacag gatgactgtc acgaagtcaa aaaacaatca actgtttgtg 2880
aatcttaaaa atgtgaaccc tgggaagaga ggccctgtgg gggactgcac tgggacattg 2940
ccggctgact ggaaacatgg cggctacctc aacgagctcg gtgcttttgg acttgacaga 3000
agctatatgg gaaaatcctt aaaagacctg ggaaagaatt ctgaatataa ttcaagtaac 3060
tgctccctaa gcagttctga gaaccataat gccactatta aagacccacc tgtacttata 3120
ccgaaaagct cagagtgtgg ttatgtggag atgaaatcgc cggcacgaag agattcccca 3180
tatgcagaga tcaataactc aacttcagcc aacaggaatg tctatgaagt tgaacctaca 3240
gtgagtgttg tccaaggagt attcagcaat aatgggcgtc tctcccagga tccatatgac 3300
ctcccaaaga acagtcacat cccttgtcat tatgacctgc tgccagtccg agacagttca 3360
tcctccccta agcaagagga cagtggaggt agcagcagca acagcagcag cagcagtga 3420
tga 3423

```

<210> 2

<211> 1140

<212> PRT

<213> homo sapiens

<400> 2

```

Met Val Ile Ser Leu Asn Ser Cys Leu Ser Phe Ile Cys Leu Leu Leu
1          5          10          15
Cys His Trp Ile Gly Thr Ala Ser Pro Leu Asn Leu Glu Asp Pro Asn
20          25          30
Val Cys Ser His Trp Glu Ser Tyr Ser Val Thr Val Gln Glu Ser Tyr
35          40          45
Pro His Pro Phe Asp Gln Ile Tyr Tyr Thr Ser Cys Thr Asp Ile Leu
50          55          60
Asn Trp Phe Lys Cys Thr Arg His Arg Val Ser Tyr Arg Thr Ala Tyr
65          70          75          80
Arg His Gly Glu Lys Thr Met Tyr Arg Arg Lys Ser Gln Cys Cys Pro
85          90          95
Gly Phe Tyr Glu Ser Gly Glu Met Cys Val Pro His Cys Ala Asp Lys
100         105         110
Cys Val His Gly Arg Cys Ile Ala Pro Asn Thr Cys Gln Cys Glu Pro
115         120         125
Gly Trp Gly Gly Thr Asn Cys Ser Ser Ala Cys Asp Gly Asp His Trp
130         135         140
Gly Pro His Cys Thr Ser Arg Cys Gln Cys Lys Asn Gly Ala Leu Cys
145         150         155         160
Asn Pro Ile Thr Gly Ala Cys His Cys Ala Ala Gly Phe Arg Gly Trp
165         170         175
Arg Cys Glu Asp Arg Cys Glu Gln Gly Thr Tyr Gly Asn Asp Cys His
180         185         190
Gln Arg Cys Gln Cys Gln Asn Gly Ala Thr Cys Asp His Val Thr Gly

```

		195					200					205				
Glu	Cys	Arg	Cys	Pro	Pro	Gly	Tyr	Thr	Gly	Ala	Phe	Cys	Glu	Asp	Leu	
	210					215					220					
Cys	Pro	Pro	Gly	Lys	His	Gly	Pro	Gln	Cys	Glu	Gln	Arg	Cys	Pro	Cys	
225					230					235					240	
Gln	Asn	Gly	Gly	Val	Cys	His	His	Val	Thr	Gly	Glu	Cys	Ser	Cys	Pro	
				245					250					255		
Ser	Gly	Trp	Met	Gly	Thr	Val	Cys	Gly	Gln	Pro	Cys	Pro	Glu	Gly	Arg	
			260					265					270			
Phe	Gly	Lys	Asn	Cys	Ser	Gln	Glu	Cys	Gln	Cys	His	Asn	Gly	Gly	Thr	
	275						280					285				
Cys	Asp	Ala	Ala	Thr	Gly	Gln	Cys	His	Cys	Ser	Pro	Gly	Tyr	Thr	Gly	
	290					295					300					
Glu	Arg	Cys	Gln	Asp	Glu	Cys	Pro	Val	Gly	Thr	Tyr	Gly	Val	Leu	Cys	
305					310					315					320	
Ala	Glu	Thr	Cys	Gln	Cys	Val	Asn	Gly	Gly	Lys	Cys	Tyr	His	Val	Ser	
				325					330					335		
Gly	Ala	Cys	Leu	Cys	Glu	Ala	Gly	Phe	Ala	Gly	Glu	Arg	Cys	Glu	Ala	
			340					345					350			
Arg	Leu	Cys	Pro	Glu	Gly	Leu	Tyr	Gly	Ile	Lys	Cys	Asp	Lys	Arg	Cys	
	355						360					365				
Pro	Cys	His	Leu	Glu	Asn	Thr	His	Ser	Cys	His	Pro	Met	Ser	Gly	Glu	
	370					375					380					
Cys	Ala	Cys	Lys	Pro	Gly	Trp	Ser	Gly	Leu	Tyr	Cys	Asn	Glu	Thr	Cys	
385					390					395					400	
Ser	Pro	Gly	Phe	Tyr	Gly	Glu	Ala	Cys	Gln	Gln	Ile	Cys	Ser	Cys	Gln	
			405						410					415		
Asn	Gly	Ala	Asp	Cys	Asp	Ser	Val	Thr	Gly	Lys	Cys	Thr	Cys	Ala	Pro	
		420						425					430			
Gly	Phe	Lys	Gly	Ile	Asp	Cys	Ser	Thr	Pro	Cys	Pro	Leu	Gly	Thr	Tyr	
	435						440					445				
Gly	Ile	Asn	Cys	Ser	Ser	Arg	Cys	Gly	Cys	Lys	Asn	Asp	Ala	Val	Cys	
	450					455					460					
Ser	Pro	Val	Asp	Gly	Ser	Cys	Thr	Cys	Lys	Ala	Gly	Trp	His	Gly	Val	
465					470					475					480	
Asp	Cys	Ser	Ile	Arg	Cys	Pro	Ser	Gly	Thr	Trp	Gly	Phe	Gly	Cys	Asn	
				485					490					495		
Leu	Thr	Cys	Gln	Cys	Leu	Asn	Gly	Gly	Ala	Cys	Asn	Thr	Leu	Asp	Gly	
			500					505					510			
Thr	Cys	Thr	Cys	Ala	Pro	Gly	Trp	Arg	Gly	Glu	Lys	Cys	Glu	Leu	Pro	
	515						520						525			
Cys	Gln	Asp	Gly	Thr	Tyr	Gly	Leu	Asn	Cys	Ala	Glu	Arg	Cys	Asp	Cys	
	530					535					540					
Ser	His	Ala	Asp	Gly	Cys	His	Pro	Thr	Thr	Gly	His	Cys	Arg	Cys	Leu	
545					550					555					560	
Pro	Gly	Trp	Ser	Gly	Val	His	Cys	Asp	Ser	Val	Cys	Ala	Glu	Gly	Arg	
				565					570					575		
Trp	Gly	Pro	Asn	Cys	Ser	Leu	Pro	Cys	Tyr	Cys	Lys	Asn	Gly	Ala	Ser	
			580					585					590			
Cys	Ser	Pro	Asp	Asp	Gly	Ile	Cys	Glu	Cys	Ala	Pro	Gly	Phe	Arg	Gly	
	595					600						605				
Thr	Thr	Cys	Gln	Arg	Ile	Cys	Ser	Pro	Gly	Phe	Tyr	Gly	His	Arg	Cys	
	610					615						620				
Ser	Gln	Thr	Cys	Pro	Gln	Cys	Val	His	Ser	Ser	Gly	Pro	Cys	His	His	
625					630					635					640	
Ile	Thr	Gly	Leu	Cys	Asp	Cys	Leu	Pro	Gly	Phe	Thr	Gly	Ala	Leu	Cys	

				645					650					655			
Asn	Glu	Val	Cys	Pro	Ser	Gly	Arg	Phe	Gly	Lys	Asn	Cys	Ala	Gly	Ile		
				660					665					670			
Cys	Thr	Cys	Thr	Asn	Asn	Gly	Thr	Cys	Asn	Pro	Ile	Asp	Arg	Ser	Cys		
				675					680					685			
Gln	Cys	Tyr	Pro	Gly	Trp	Ile	Gly	Ser	Asp	Cys	Ser	Gln	Pro	Cys	Pro		
				690					695					700			
Pro	Ala	His	Trp	Gly	Pro	Asn	Cys	Ile	His	Thr	Cys	Asn	Cys	His	Asn		
705						710					715				720		
Gly	Ala	Phe	Cys	Ser	Ala	Tyr	Asp	Gly	Glu	Cys	Lys	Cys	Thr	Pro	Gly		
				725							730				735		
Trp	Thr	Gly	Leu	Tyr	Cys	Thr	Gln	Arg	Cys	Pro	Leu	Gly	Phe	Tyr	Gly		
				740										750			
Lys	Asp	Cys	Ala	Leu	Ile	Cys	Gln	Cys	Gln	Asn	Gly	Ala	Asp	Cys	Asp		
				755										765			
His	Ile	Ser	Gly	Gln	Cys	Thr	Cys	Arg	Thr	Gly	Phe	Met	Gly	Arg	His		
				770										780			
Cys	Glu	Gln	Lys	Cys	Pro	Ser	Gly	Thr	Tyr	Gly	Tyr	Gly	Cys	Arg	Gln		
785						790									800		
Ile	Cys	Asp	Cys	Leu	Asn	Asn	Ser	Thr	Cys	Asp	His	Ile	Thr	Gly	Thr		
				805										815			
Cys	Tyr	Cys	Ser	Pro	Gly	Trp	Lys	Gly	Ala	Arg	Cys	Asp	Gln	Ala	Gly		
				820										830			
Val	Ile	Ile	Val	Gly	Asn	Leu	Asn	Ser	Leu	Ser	Arg	Thr	Ser	Thr	Ala		
				835										845			
Leu	Pro	Ala	Asp	Ser	Tyr	Gln	Ile	Gly	Ala	Ile	Ala	Gly	Ile	Ile	Ile		
				850										860			
Leu	Val	Leu	Val	Val	Leu	Phe	Leu	Leu	Ala	Leu	Phe	Ile	Ile	Tyr	Arg		
865						870									880		
His	Lys	Gln	Lys	Gly	Lys	Glu	Ser	Ser	Met	Pro	Ala	Val	Thr	Tyr	Thr		
				885											895		
Pro	Ala	Met	Arg	Val	Val	Asn	Ala	Asp	Tyr	Thr	Ile	Ser	Gly	Thr	Leu		
				900										910			
Pro	His	Ser	Asn	Gly	Gly	Asn	Ala	Asn	Ser	His	Tyr	Phe	Thr	Asn	Pro		
				915										925			
Ser	Tyr	His	Thr	Leu	Thr	Gln	Cys	Ala	Thr	Ser	Pro	His	Val	Asn	Asn		
				930										940			
Arg	Asp	Arg	Met	Thr	Val	Thr	Lys	Ser	Lys	Asn	Asn	Gln	Leu	Phe	Val		
945						950									960		
Asn	Leu	Lys	Asn	Val	Asn	Pro	Gly	Lys	Arg	Gly	Pro	Val	Gly	Asp	Cys		
				965										975			
Thr	Gly	Thr	Leu	Pro	Ala	Asp	Trp	Lys	His	Gly	Gly	Tyr	Leu	Asn	Glu		
				980										990			
Leu	Gly	Ala	Phe	Gly	Leu	Asp	Arg	Ser	Tyr	Met	Gly	Lys	Ser	Leu	Lys		
				995										1005			
Asp	Leu	Gly	Lys	Asn	Ser	Glu	Tyr	Asn	Ser	Ser	Asn	Cys	Ser	Leu	Ser		
				1010										1020			
Ser	Ser	Glu	Asn	Pro	Tyr	Ala	Thr	Ile	Lys	Asp	Pro	Pro	Val	Leu	Ile		
1025						1030									1040		
Pro	Lys	Ser	Ser	Glu	Cys	Gly	Tyr	Val	Glu	Met	Lys	Ser	Pro	Ala	Arg		
				1045										1055			
Arg	Asp	Ser	Pro	Tyr	Ala	Glu	Ile	Asn	Asn	Ser	Thr	Ser	Ala	Asn	Arg		
				1060										1070			
Asn	Val	Tyr	Glu	Val	Glu	Pro	Thr	Val	Ser	Val	Val	Gln	Gly	Val	Phe		
				1075										1085			
Ser	Asn	Asn	Gly	Arg	Leu	Ser	Gln	Asp	Pro	Tyr	Asp	Leu	Pro	Lys	Asn		

WO 02/072611

PCT/US02/07477

1090	1095	1100
Ser His Ile Pro Cys His Tyr Asp Leu Leu Pro Val Arg Asp Ser Ser		
1105	1110	1115
Ser Ser Pro Lys Gln Glu Asp Ser Gly Gly Ser Ser Ser Asn Ser Ser		1120
	1125	1130
Ser Ser Ser Glu		1135
1140		

<210> 3
 <211> 1761
 <212> DNA
 <213> homo sapiens

<400> 3
 atgggttattt ctttgaactc atgcctgagc tttatttgggt tattgttatg ccactggatt 60
 gggacagcat cacctctgaa tcttgaagac cctaattgtgt gtagccactg ggaaagctac 120
 tcagtgactg tgcaagagtc ataccacacat ccctttgatc aaatttacta cagagctgc 180
 actgacattc taaactgggt taaatgcacg cggcacagag tcagctatcg gacagcctat 240
 cgacatgggg agaagactat gtataggcgc aagtctcagt gttgtcctgg attttatgaa 300
 agcggggaaa tgtgtgtccc ccactgtgct gataaatgtg tccatggctg ctgtattgct 360
 ccaaacacct gtcagtgtga gcctggctgg ggagggacca actgctccag tgcttgcgat 420
 ggtgatcact ggggtcccca ctgcaccagc cgggtgccagt gcaaaaatgg ggctctgtgc 480
 aaccccatca ccggggcctt ccactgtgct gcgggcttcc ggggctggcg ctgcgaggac 540
 cgctgtgagc agggcaccta tggtaacgac tgtcatcaga gatgccagt ccagaatgga 600
 gccacctgcg accacgtcac gggggaatgc cgtgcccac caggatacac cggagccttc 660
 tgtgaggatc tttgtcctcc tggtaaacat ggtccacagt gtgagcagag atgcccttgt 720
 caaaatggag gagtgtgtca tcacgtcact ggagaatgct cttgcccttc tggctggatg 780
 ggcacagtgt gtggtcagcc ttgccccgag ggtcgctttg gaaagaactg ttccaagaa 840
 tgccagtgcc ataattggagg gacgtgtgat gctgccacag gccaatgtca ttgcagtcca 900
 ggatacacag gggaacgggt ccaggatgag tgtcctgttg ggacctatgg cgttctctgt 960
 gctgagacct gccagtgtgt caacggaggg aagtgttacc acgtgagcgg cgcctgcctc 1020
 tgtgaagcag gctttgctgg cgagcgctgc gaagcacgcc tgtgtcctga ggggctctac 1080
 ggcacacaaat gtgacaaacg gtgtccctgc cacttggaac aactcatag ctgtcacccc 1140
 atgtctggag agtgtgcctg caagccgggc tggtcaggac tctactgtaa tgagacatgt 1200
 tctcctggat tctacgggga agcttgccag cagatctgca gctgccaaaa tggggcagac 1260
 tgtgacagtg tgactggaaa gtgcacctgt gcccaggat tcaaaggaat tgactgctct 1320
 accccatgcc ctctgggaac ctatgggata aactgttcct ctgctgtgg ctgtaaaaat 1380
 gatgcagtct gctctcctgt ggacgggtct tgtacttgca aggcaggctg gcacgggggtg 1440
 gactgctcca tcagatgtcc cagtggcaca tggggctttg gctgttaactt aacatgccag 1500
 tgcctcaacg ggggagcctg caacaccctg gacgggacct gcacgtgtgc acctggatgg 1560
 cgcgggggaga aatgcgaact tccctgccag gatggcacgt acgggctgaa ctgtgctgag 1620
 cgctgcgact gcagccacgc agatggctgc caccctacca cgggccattg ccgctgcctc 1680
 ccgggatggg caggtgtcca ctgtgacagc gtgtgtgctg agggacgctg gggccccaac 1740
 tgctccctgc cctgctactg a 1761

<210> 4
 <211> 586
 <212> PRT
 <213> homo sapiens

<400> 4
 Met Val Ile Ser Leu Asn Ser Cys Leu Ser Phe Ile Cys Leu Leu Leu
 1 5 10 15
 Cys His Trp Ile Gly Thr Ala Ser Pro Leu Asn Leu Glu Asp Pro Asn
 20 25 30

6/7

WO 02/072611

PCT/US02/07477

Asp	Cys	Ser	Ile	Arg	Cys	Pro	Ser	Gly	Thr	Trp	Gly	Phe	Gly	Cys	Asn	
				485					490					495		
Leu	Thr	Cys	Gln	Cys	Leu	Asn	Gly	Gly	Ala	Cys	Asn	Thr	Leu	Asp	Gly	
			500					505					510			
Thr	Cys	Thr	Cys	Ala	Pro	Gly	Trp	Arg	Gly	Glu	Lys	Cys	Glu	Leu	Pro	
		515					520					525				
Cys	Gln	Asp	Gly	Thr	Tyr	Gly	Leu	Asn	Cys	Ala	Glu	Arg	Cys	Asp	Cys	
	530					535					540					
Ser	His	Ala	Asp	Gly	Cys	His	Pro	Thr	Thr	Gly	His	Cys	Arg	Cys	Leu	
545					550					555					560	
Pro	Gly	Trp	Ser	Gly	Val	His	Cys	Asp	Ser	Val	Cys	Ala	Glu	Gly	Arg	
			565					570					575			
Trp	Gly	Pro	Asn	Cys	Ser	Leu	Pro	Cys	Tyr							
			580					585								

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising at least 24 contiguous bases of nucleotide sequence first
5 disclosed in SEQ ID NO: 1.

2. An isolated nucleic acid molecule comprising a nucleotide sequence that:

- 10 (a) encodes the amino acid sequence shown in SEQ ID NO: 2; and
- (b) hybridizes under stringent conditions to the nucleotide sequence of SEQ ID NO:1 or the complement thereof.

15 3. An isolated nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:2.

20 4. An isolated nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:4.