



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : C12N 15/12, C07K 14/475, 16/22, A61K 38/18, C12Q 1/68, A61K 31/70, C12N 1/21, C07K 19/00, G01N 33/58, 33/68</p>	<p>A2</p>	<p>(11) International Publication Number: <b>WO 98/04692</b> (43) International Publication Date: 5 February 1998 (05.02.98)</p>
<p>(21) International Application Number: PCT/US97/12735 (22) International Filing Date: 21 July 1997 (21.07.97) (30) Priority Data: 08/690,094 31 July 1996 (31.07.96) US (60) Parent Application or Grant (63) Related by Continuation US 08/690,094 (CIP) Filed on 31 July 1996 (31.07.96) (71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): COLEMAN, Roger [US/US]; 260 Mariposa #2, Mountain View, CA 94041 (US). STUART, Susan, G. [US/US]; 1256 Birch Street, Montara, CA 94037 (US). (74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).</p>	<p>(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i></p>	

(54) Title: HUMAN NEURONATIN

```

          9           18           27           36           45           54
5' NCC GAG ACC AGC GGA TCT CGG CAA ACC CTC TTT CTC GAC CAC CCA CCT ACC ATT

          63           72           81           90           99           108
CTT GGA ACC ATG GCG GCA GTG GCG GCG GCC TCG GCT GAA CTG CTC ATC ATC GGC
          M A A V A A A S A E L L I I G

          117          126          135          144          153          162
TGG TAC ATC TTC CGC GTG CTG CTG CAG GTG TTC AGG TAC TCC CTG CAG AAG CTG
W Y I F R V L L Q V F R Y S L Q K L

          171          180          189          198          207          216
GCA TAC ACG GTG TCG CGG ACC GGG CGG CAG GTG TTG GGG GAG CGC AGG CAG CGA
A Y T V S R T G R Q V L G E R R Q R

          225          234          243          252
GCC CCC AAC TGA GGC CCC AGC TTC CAG CCC TGG GCG GCC 3'
A P N
    
```

(57) Abstract

The present invention provides a polynucleotide which identifies and encodes a novel human neuronatin (HNEU). The invention provides for genetically engineered expression vectors and host cells comprising the nucleic acid sequence encoding HNEU. The invention also provides for the use of substantially purified HNEU and its agonists in the commercial production of recombinant proteins and in pharmaceutical compositions for the treatment of diseases associated with the expression of HNEU. Additionally, the invention provides for the use of antisense molecules to HNEU in pharmaceutical compositions for treatment of diseases associated with the expression of HNEU. The invention also describes diagnostic assays which utilize diagnostic compositions comprising the polynucleotide, fragments or the complement thereof, which hybridize with the genomic sequence or the transcript of polynucleotides encoding HNEU or anti-HNEU antibodies which specifically bind to HNEU.

**FOR THE PURPOSES OF INFORMATION ONLY**

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

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakistan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**HUMAN NEURONATIN****TECHNICAL FIELD**

The present invention relates to nucleic acid and amino acid sequences of a novel human neuronatin and to the use of these sequences in the diagnosis, study, prevention and treatment of  
5 disease.

**BACKGROUND ART**

Neuronatin is a brain specific gene that was first identified in neonatal rat brain (Joseph, R. et al (1994) *Biochem. Biophys. Res. Commun.* 201(3):1227-1234). Subsequent studies revealed that neuronatin mRNA exists as two isoforms, alpha and beta (neu-a and neu-b in both  
10 mouse and rat (Joseph, R. et al (1995) *Brain Research* 690:92-98; Wijnholds, J. et al (1995) *Dev. Biol.* 171:73-84). The alpha-form encodes a protein of 81 amino acids, and the beta-form, a protein of 54 amino acids.

Both isoforms have the same open reading frame and differ only with respect to an additional 81 base pair region in neu-a encoding 27 amino acids embedded in the middle of its  
15 coding region. The two proteins have identical hydrophobic N-terminal and hydrophilic C-terminal ends, suggesting that they may be membrane bound and function as transmembrane proteins in signal transduction or intercellular communication. The two isoforms are both expressed during early brain development although differences in the time of appearance indicate that they are regulated differently. In particular, neuronatin is highly expressed in the developing  
20 hindbrain and pituitary. The mouse studies further demonstrate strong expression of neuronatin in the central and peripheral nervous system during later embryogenesis. Together these studies suggest a role for neuronatin in both hindbrain and pituitary development and in the development and differentiation of the central and peripheral nervous system. Both neu-a and neu-b are undetected in adult rat brain. The two isoforms have also been found in human fetal brain (Dou,  
25 D. And Joseph, R. (1996) *Genomics* 33:292-297,) and the deduced proteins are both highly conserved compared to rat or mouse.

Based upon its localization to brain, and the close correlation of its expression with early brain development, neuronatin appears to play a role in both normal and abnormal brain development. Furthermore, the loss of neuronatin expression in adulthood suggests that its  
30 downregulation may be associated with brain senescence.

**DISCLOSURE OF THE INVENTION**

The present invention discloses a novel human neuronatin-beta (hereinafter referred to as

HNEU) characterized as having homology to other neuronatin-beta isoforms. Accordingly, the invention features a substantially purified neuronatin, encoded by the amino acid sequence of SEQ ID NO:1, and having characteristics of neuronatin beta isoform family members.

One aspect of the invention features isolated and substantially purified polynucleotides  
5 which encode HNEU. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO:2. In addition, the invention features polynucleotide sequences that hybridize under stringent conditions to SEQ ID NO:2.

The invention further relates to nucleic acid sequence encoding HNEU, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof. The present  
10 invention also relates to an expression vector which includes polynucleotide encoding HNEU and its use to transform host cells or organisms.

The present invention also relates to methods for making HNEU, anti-HNEU antibodies and pharmaceutical compositions comprising HNEU or nucleic acid encoding HNEU.

#### **BRIEF DESCRIPTION OF DRAWINGS**

15 Figure 1 shows the nucleic acid sequence (SEQ ID NO:2) and amino acid sequence (SEQ ID NO:1) of the human neuronatin, HNEU, and was produced using MacDNAsis software (Hitachi Software Engineering Co Ltd). SEQ ID NO:2 is derived from the partial cDNAs, Incyte Clones 674382 and 926572 (SEQ ID NO:3 and 4.)

Figure 2 shows the amino acid sequence alignments among HNEU (SEQ ID NO:1) and  
20 other known neuronatin-betas; mouse neuronatin-beta (GI 619500; SEQ ID NO:5) and human neuronatin-beta (GI 1244410; SEQ ID NO:6), produced using the multisequence alignment program of DNASTar software (DNASTar Inc, Madison WI).

Figure 3 shows the hydrophobicity plot (generated using MacDNAsis software) for HNEU, SEQ ID NO:1; the X axis reflects amino acid position, and the negative Y axis,  
25 hydrophobicity.

Figure 4 shows the hydrophobicity plot for mouse neuronatin-beta, SEQ ID NO:5.

#### **MODES FOR CARRYING OUT THE INVENTION**

##### **Definitions**

“Nucleic acid sequence” as used herein refers to an oligonucleotide, nucleotide or  
30 polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, amino acid sequence as used herein refers to protein or peptide sequence.

“Peptide nucleic acid” as used herein refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary (template) strand of nucleic acid (Nielsen PE et al (1993) Anticancer Drug Des 5 8:53-63).

As used herein, HNEU refers to the amino acid sequence of substantially purified HNEU from any source whether natural, synthetic, semi-synthetic or recombinant.

A “variant” of HNEU is defined as an amino acid sequence that is different by one or more amino acid “substitutions”. The variant may have “conservative” changes, wherein a substituted amino acid has similar structural or chemical properties, eg, replacement of leucine with isoleucine. More rarely, a variant may have “nonconservative” changes, eg, replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTar software.

The term “biologically active” refers to a HNEU having structural, regulatory or biochemical functions of the naturally occurring HNEU. Likewise, “immunologically active” defines the capability of the natural, recombinant or synthetic HNEU, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term “derivative” as used herein refers to the chemical modification of a nucleic acid sequence encoding HNEU or the encoded HNEU. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of natural HNEU.

As used herein, the term “substantially purified” refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

“Stringency” typically occurs in a range from about  $T_m - 5^\circ\text{C}$  ( $5^\circ\text{C}$  below the  $T_m$  of the probe) to about  $20^\circ\text{C}$  to  $25^\circ\text{C}$  below  $T_m$ . As will be understood by those of skill in the art, a stringency hybridization can be used to identify or detect identical polynucleotide sequences or to identify or detect similar or related polynucleotide sequences.

The term "hybridization" as used herein shall include "any process by which a strand of nucleic acid joins with a complementary strand through base pairing" (Coombs J (1994) Dictionary of Biotechnology, Stockton Press, New York NY). Amplification as carried out in the polymerase chain reaction technologies is described in Dieffenbach CW and GS Dveksler (1995, 5 PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview NY).

A "deletion" is defined as a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent.

An "insertion" or "addition" is that change in a nucleotide or amino acid sequence which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as 10 compared to the naturally occurring HKLP.

A "substitution" results from the replacement of one or more nucleotides or amino acids by different nucleotides or amino acids, respectively.

### **Description**

The present invention relates to a novel human neuronatin, HNEU, initially identified 15 among the partial cDNAs from a brain cerebellum library (CRBLNOTO1) and to the use of the nucleic acid and amino acid sequences disclosed herein in the study, diagnosis, prevention and treatment of disease. In addition to being found in adult human cerebellum, nucleic acid encoding a portion of HNEU was also found in cDNA libraries from human fetal brain and pituitary.

The present invention also encompasses HNEU variants. A preferred HNEU variant is 20 one having at least 80% amino acid sequence similarity to the HNEU amino acid sequence (SEQ ID NO:1), a more preferred HNEU variant is one having at least 90% amino acid sequence similarity to SEQ ID NO:1 and a most preferred HNEU variant is one having at least 95% amino acid sequence similarity to SEQ ID NO:1. Although HNEU encoding nucleotide sequences are found transcribed in these tissues, the naturally occurring expression is not necessarily limited to 25 these tissues.

Nucleic acid encoding a portion of HNEU was first identified in the cDNA, Incyte Clone 674382 (SEQ ID NO:3), through a computer-generated search for amino acid sequence alignments. The nucleic acid sequence, SEQ ID NO:2; disclosed herein and designated in lower case (hneu) encodes the amino acid sequence, SEQ ID NO:1, disclosed hereinafter as HNEU. 30 The present invention is based, in part, on the chemical and structural homology among HNEU, and mouse neuronatin-beta (GI 619500; Wijnholds, J. et al. (1995) Dev. Biol. 171:73-84), and human neuronatin-beta (GI 1244410; Dou, D. and Joseph, R. (1996) Genomics 33:292-297).

HNEU has amino acid sequence homology to both mouse and human neuronatin-beta (Figure 2). The hydrophobicity plots comparing HNEU and mouse neuronatin-beta (Figures 3 and 4 respectively) are similar and reflect the hydrophobic N-terminal regions and hydrophilic C-terminal characteristic of a putative transmembrane protein.

## 5 The HNEU Coding Sequences

The nucleic acid and deduced amino acid sequences of HNEU are shown in Figure 1. In accordance with the invention, any nucleic acid sequence which encodes the amino acid sequence of HNEU can be used to generate recombinant molecules which express HNEU. In a specific embodiment described herein, a partial sequence of HNEU was first isolated as Incyte Clone  
10 674382 from a human cerebellum cDNA library (CRBLNOTO1).

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of HNEU-encoding nucleotide sequences, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene may be produced. The invention contemplates each and every possible variation of nucleotide sequence  
15 that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring HNEU, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HNEU and its variants are preferably  
20 capable of hybridizing to the nucleotide sequence of the naturally occurring HNEU under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HNEU or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic expression host in accordance with the frequency with which  
25 particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HNEU and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

It is now possible to produce a DNA sequence, or portions thereof, encoding a HNEU and  
30 its derivatives entirely by synthetic chemistry, after which the synthetic gene may be inserted into any of the many available DNA vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to

introduce mutations into a HNEU sequence or any portion thereof.

Also included within the scope of the present invention are polynucleotide sequences that are capable of hybridizing to the nucleotide sequence of Figure 1 under various conditions of stringency. Hybridization conditions are based on the melting temperature ( $T_m$ ) of the nucleic acid binding complex or probe, as taught in Berger and Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol 152, Academic Press, San Diego CA) incorporated herein by reference, and may be used at a defined stringency.

Altered nucleic acid sequences encoding HNEU which may be used in accordance with the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent HNEU. The protein may also show deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HNEU. Deliberate 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 as long as the biological activity of HNEU is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine; glycine, alanine; asparagine, glutamine; serine, threonine phenylalanine, and tyrosine.

Included within the scope of the present invention are alleles of hneu. As used herein, an "allele" or "allelic sequence" is an alternative form of hneu. Alleles result from a mutation, ie, a change in the nucleic acid sequence, and generally produce altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions or substitutions of amino acids. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

Methods for DNA sequencing are well known in the art and employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland OH), Taq polymerase (Perkin Elmer, Norwalk CT), thermostable T7 polymerase (Amersham, Chicago IL), or combinations of recombinant polymerases and proofreading exonucleases such as the ELONGASE Amplification System marketed by Gibco BRL (Gaithersburg MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno NV),



Peltier Thermal Cycler (PTC200; MJ Research, Watertown MA) and the ABI 377 DNA sequencers (Perkin Elmer).

### **Extending the Polynucleotide Sequence**

The polynucleotide sequence encoding HNEU may be extended utilizing partial  
5 nucleotide sequence and various methods known in the art to detect upstream sequences such as promoters and regulatory elements. Gobinda et al (1993; PCR Methods Applic. 2:318-22) disclose "restriction-site" polymerase chain reaction (PCR) as a direct method which uses universal primers to retrieve unknown sequence adjacent to a known locus. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known  
10 region. The amplified sequences are 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 can be used to amplify or extend sequences using divergent primers based on a known region (Triglia T et al (1988) Nucleic Acids Res 16:8186). The primers may be  
15 designed using OLIGO® 4.06 Primer Analysis Software (1992; National Biosciences Inc, Plymouth MN), or another appropriate program, 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  
20 template.

Capture PCR (Lagerstrom M et al (1991) PCR Methods Applic 1:111-19) is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into an unknown portion of the DNA  
25 molecule before PCR.

Another method which may be used to retrieve unknown sequences is that of Parker JD et al (1991; Nucleic Acids Res 19:3055-60). Additionally, one can use PCR, nested primers and PromoterFinder libraries to walk in genomic DNA (PromoterFinder™ Clontech, Palo Alto CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.  
30 Preferred libraries for screening for full length cDNAs are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes. A randomly primed library may

be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region.

Capillary electrophoresis may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. Systems for rapid sequencing are available from  
5 Perkin Elmer, Beckman Instruments (Fullerton CA), and other companies. Capillary sequencing may 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 is converted to electrical signal using appropriate software (eg. Genotyper™ and Sequence Navigator™ from Perkin Elmer) and the  
10 entire process from loading of samples to computer analysis and electronic data display is computer controlled. Capillary electrophoresis is particularly suited to the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample. The reproducible sequencing of up to 350 bp of M13 phage DNA in 30 min has been reported (Ruiz-Martinez MC et al (1993) Anal Chem 65:2851-8).

#### 15 **Expression of the Nucleotide Sequence**

In accordance with the present invention, polynucleotide sequences which encode HNEU, fragments of the polypeptide, fusion proteins or functional equivalents thereof may be used in recombinant DNA molecules that direct the expression of HNEU in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially  
20 the same or a functionally equivalent amino acid sequence, may be used to clone and express HNEU. As will be understood by those of skill in the art, it may be advantageous to produce HNEU-encoding nucleotide sequences possessing non-naturally occurring codons. Codons preferred by a particular prokaryotic or eukaryotic host (Murray E et al (1989) Nuc Acids Res 17:477-508) can be selected, for example, to increase the rate of hneu expression or to produce  
25 recombinant RNA transcripts having desirable properties, such as a longer half-life, than transcripts produced from naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered in order to alter a hneu coding sequence for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of the gene product. For example, mutations  
30 may be introduced using techniques which are well known in the art, eg, site-directed mutagenesis to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to produce splice variants, etc.

In another embodiment of the invention, a natural, modified or recombinant sequence encoding HNEU may be ligated to a heterologous sequence to encode a fusion protein. For example, for screening of peptide libraries for inhibitors of HNEU activity, it may be useful to encode a chimeric HNEU protein that is recognized by a commercially available antibody. A  
5 fusion protein may also be engineered to contain a cleavage site located between a HNEU sequence and the heterologous protein sequence, so that the HNEU may be cleaved and substantially purified away from the heterologous moiety.

In an alternate embodiment of the invention, the coding sequence of hneu may be synthesized, whole or in part, using chemical methods well known in the art (see Caruthers MH  
10 et al (1980) Nuc Acids Res Symp Ser 215-23, Horn T et al(1980) Nuc Acids Res Symp Ser 225-32, etc). Alternatively, the protein itself could be produced using chemical methods to synthesize a HNEU amino acid sequence, whole or in part. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge JY et al (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A  
15 Peptide Synthesizer (Perkin Elmer) in accordance with the instructions provided by the manufacturer.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (eg, Creighton (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co, New York NY). The composition of the synthetic peptides  
20 may be confirmed by amino acid analysis or sequencing (eg, the Edman degradation procedure; Creighton, supra). Additionally the amino acid sequence of HNEU, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

### **Expression Systems**

25 In order to express a biologically active HNEU, the nucleotide sequence encoding HNEU or its functional equivalent, is inserted into an appropriate expression vector, ie, a 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  
30 expression vectors containing a HNEU coding sequence and appropriate transcriptional or translational controls. These methods include in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination or genetic recombination. Such techniques are described in

Sambrook et al (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY and Ausubel FM et al (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY.

A variety of expression vector/host systems may be utilized to contain and express a  
5 HNEU coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (eg, baculovirus); plant cell systems transfected with virus expression vectors (eg, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with bacterial  
10 expression vectors (eg, Ti or pBR322 plasmid); or animal cell systems.

The "control elements" or "regulatory sequences" of these systems vary in their strength and specificities and are those nontranslated regions of the vector, enhancers, promoters, and 3' untranslated regions, which interact with host cellular proteins to carry out transcription and translation. Depending on the vector system and host utilized, any number of suitable  
15 transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, LaJolla CA) or pSport1 (Gibco BRL) and ptrp-lac hybrids and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (eg, heat  
20 shock, RUBISCO; and storage protein genes) or from plant viruses (eg, viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from the mammalian genes or from mammalian viruses are most appropriate. If it is necessary to generate a cell line that contains multiple copies of H hneu, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

25 In bacterial systems, a number of expression vectors may be selected depending upon the use intended for HNEU. For example, when large quantities of HNEU are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as Bluescript® (Stratagene), in which  
30 the hneu coding sequence may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of  $\beta$ -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke & Schuster (1989) J Biol Chem 264:5503-5509); and the like.

pGEX vectors (Promega, Madison WI) may 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. Proteins made in such systems are 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 may be used. For reviews, see Ausubel et al (supra) and Grant et al (1987) *Methods in Enzymology* 153:516-544.

10 In cases where plant expression vectors are used, the expression of a sequence encoding HNEU may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV (Brisson et al (1984) *Nature* 310:511-514) may be used alone or in combination with the omega leader sequence from TMV (Takamatsu et al (1987) *EMBO J* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO  
15 (Coruzzi et al (1984) *EMBO J* 3:1671-1680; Broglie et al (1984) *Science* 224:838-843); or heat shock promoters (Winter J and Sinibaldi RM (1991) *Results Probl Cell Differ* 17:85-105) may be used. These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. For reviews of such techniques, see Hobbs S or Murry LE in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill New York NY, pp  
20 191-196 or Weissbach and Weissbach (1988) Methods for Plant Molecular Biology, Academic Press, New York NY, pp 421-463.

An alternative expression system which could be used to express hneu is an insect system. 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. The  
25 hneu coding sequence may be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of hneu will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect S. frugiperda cells or Trichoplusia larvae in which HNEU is expressed (Smith et al (1983) *J Virol* 46:584; Engelhard EK et al (1994) *Proc*  
30 *Nat Acad Sci* 91:3224-7).

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, a hneu coding sequence may be

ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing HNEU in infected host cells (Logan and Shenk (1984) Proc Natl Acad Sci 81:3655-59). In addition, transcription enhancers, such as the rous  
5 sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be required for efficient translation of a hneu sequence. These signals include the ATG initiation codon and adjacent sequences. In cases where hneu, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases  
10 where only coding sequence, or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers  
15 appropriate to the cell system in use (Scharf D et al (1994) Results Probl Cell Differ 20:125-62; Bittner et al (1987) Methods in Enzymol 153:516-544).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation,  
20 glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 293, WI38, etc have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

25 For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express hneu may be transformed using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose  
30 of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell

type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler M et al (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy I et al (1980) Cell 22:817-23) genes which can be employed in tk- or apt- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler M et al (1980) Proc Natl Acad Sci 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin F et al (1981) J Mol Biol 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, 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 SC and RC Mulligan (1988) Proc Natl Acad Sci 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins,  $\beta$  glucuronidase and its substrate, GUS, and luciferase and its substrate, luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes CA et al (1995) Methods Mol Biol 55:121-131).

#### **Identification of Transformants Containing the Polynucleotide Sequence**

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression should be confirmed. For example, if the hneu is inserted within a marker gene sequence, recombinant cells containing hneu can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a HNEU sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem hneu as well.

Alternatively, host cells which contain the coding sequence for hneu and express HNEU may 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 hybridization and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of the nucleic acid or protein.

The presence of the polynucleotide sequence encoding HNEU can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes, portions or fragments of

hneu. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequence encoding HNEU to detect transformants containing hneu DNA or RNA. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of HNEU, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HNEU is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton R et al (1990, Serological Methods, a Laboratory Manual, APS Press, St Paul MN) and Maddox DE et al (1983, J Exp Med 158:1211).

A wide variety of labels and conjugation techniques are known by those skilled in the 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 hneu include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the hneu sequence, or any portion of it, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3 or SP6 and labeled nucleotides.

A number of companies such as Pharmacia Biotech (Piscataway NJ), Promega (Madison WI), and US Biochemical Corp (Cleveland OH) supply commercial kits and protocols for these procedures. Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. Patents teaching the use of such labels include US Patents 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in US Patent No. 4,816,567 incorporated herein by reference.

### 30 Purification of HNEU

Host cells transformed with a nucleotide sequence encoding NHEU may be cultured under conditions suitable for the expression and recovery of the encoded protein from cell



culture. The protein produced by a recombinant cell may 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 encoding HNEU can be designed with signal sequences which direct secretion of HNEU through a prokaryotic or eukaryotic cell membrane.

- 5 Other recombinant constructions may join hneu to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins (Kroll DJ et al (1993) DNA Cell Biol 12:441-53).

HNEU may also be expressed as a recombinant protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating  
10 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 WA). The inclusion of a cleavable linker sequences such as Factor XA or enterokinase (Invitrogen, San Diego CA) between the purification domain  
15 and HNEU is useful to facilitate purification. One such expression vector provides for expression of a fusion protein comprising an HNEU and contains nucleic acid encoding 6 histidine residues followed by thioredoxin and an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography as described in Porath et al (1992) Protein Expression and Purification 3: 263-281) while the enterokinase  
20 cleavage site provides a means for purifying the neuronatin from the fusion protein.

In addition to recombinant production, fragments of HNEU may be produced by direct peptide synthesis using solid-phase techniques (cf Stewart et al (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco; Merrifield J (1963) J Am Chem Soc 85:2149-2154).  
In vitro protein synthesis may be performed using manual techniques or by automation.  
25 Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City CA) in accordance with the instructions provided by the manufacturer. Various fragments of HNEU may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

#### Uses of HNEU

30 The rationale for the use of nucleotide and polypeptide sequences disclosed herein is based in part on the chemical and structural homology among the novel HNEU disclosed herein and known family members such as the mouse neuronatin-beta (GI 619500; Wijnholds, supra)

and human neuronatin-beta (GI 1244410; Dou, supra).

Accordingly, decreased expression of HNEU may suggest abnormal brain development or conditions associated with brain senescence, such as Alzheimer's disease. Increasing levels of HNEU in the brain either by gene therapy, using polynucleotides encoding HNEU or through  
5 delivery of the protein HNEU, in suitable form, may be a useful therapeutic strategy for treating these conditions. HNEU may be also useful in the diagnosis and treatment of diseases associated with the central and peripheral nervous system such as Parkinson's disease.

HNEU and/or a cell line that expresses HNEU may be used to evaluate, screen and identify compounds, synthetic drugs, antibodies, peptides or other molecules that modulate the  
10 activity of HNEU and may therefore be useful in the treatment of disease conditions associated with HNEU.

### **HNEU Antibodies**

HNEU-specific antibodies are useful for the diagnosis of conditions and diseases associated with expression of HNEU. Such antibodies may include, but are not limited to,  
15 polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library. Neutralizing antibodies, ie, those which inhibit dimer formation, are especially preferred for diagnostics and therapeutics.

HNEU for antibody induction does not require biological activity; however, the protein fragment, or oligopeptide must be antigenic. Peptides used to induce specific antibodies may  
20 have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids. Preferably they should mimic a portion of the amino acid sequence of the natural protein and may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HNEU amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule. Procedures well  
25 known in the art can be used for the production of antibodies to HNEU.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, etc may be immunized by injection with HNEU or any portion, fragment or oligopeptide which retains immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include but are not limited to Freund's,  
30 mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are potentially

useful human adjuvants.

Monoclonal antibodies to HNEU may be prepared using 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 originally described by Koehler and Milstein (1975 Nature 256:495-497), the human B-cell hybridoma technique (Kosbor et al (1983) Immunol Today 4:72; Cote et al (1983) Proc Natl Acad Sci 80:2026-2030) and the EBV-hybridoma technique (Cole et al (1985) Monoclonal Antibodies and Cancer Therapy, Alan R Liss Inc, New York NY, pp 77-96).

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 (1984) Proc Natl Acad Sci 81:6851-6855; Neuberger et al (1984) Nature 312:604-608; Takeda et al (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies (US Patent No. 4,946,778) can be adapted to produce HNEU-specific single chain antibodies

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi et al (1989, Proc Natl Acad Sci 86: 3833-3837), and Winter G and Milstein C (1991; Nature 349:293-299).

Antibody fragments which contain specific binding sites for HNEU may also be generated. For example, such fragments include, but are not limited to, the F(ab')<sub>2</sub> 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')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse WD et al.(1989) Science 256:1275-1281).

A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the formation of complexes between HNEU and its specific antibody and the measurement of complex formation. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two noninterfering epitopes on a specific HNEU protein is preferred, but a competitive binding assay may also be employed. These assays are described in Maddox DE et al (1983, J Exp Med 158:1211).

### **Diagnostic Assays Using HNEU Specific Antibodies**

Particular HNEU antibodies are useful for the diagnosis of conditions or diseases characterized by expression of HNEU or in assays to monitor patients being treated with HNEU, agonists or inhibitors. Diagnostic assays for HNEU include methods utilizing the antibody and a  
5 label to detect HNEU in human body fluids or extracts of cells or tissues. The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, the polypeptides and antibodies will be labeled by joining them, either covalently or noncovalently, with a reporter molecule. A wide variety of reporter molecules are known, several of which were described above.

10 A variety of protocols for measuring HNEU, using either polyclonal or monoclonal antibodies specific for the respective protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HNEU is preferred, but a competitive  
15 binding assay may be employed. These assays are described, among other places, in Maddox, DE et al (1983, J Exp Med 158:1211).

In order to provide a basis for diagnosis, normal or standard values for HNEU expression must be established. This is accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with antibody to HNEU under conditions suitable for  
20 complex formation which are well known in the art. The amount of standard complex formation may be quantified by comparing various artificial membranes containing known quantities of HNEU with both control and disease samples from biopsied tissues. Then, standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by disease. Deviation between standard and subject values  
25 establishes the presence of disease state.

### **Drug Screening**

HNEU, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening therapeutic compounds in any of a variety of drug screening techniques. The fragment employed in such a test may be free in solution, affixed to a solid support, borne on a cell surface,  
30 or located intracellularly. The formation of binding complexes, between HNEU and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput

screening of compounds having suitable binding affinity to the HNEU is described in detail in "Determination of Amino Acid Sequence Antigenicity" by Geysen HN, WO Application 84/03564, published on September 13, 1984, and incorporated herein by reference. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with fragments of HNEU and washed. Bound HNEU is then detected by methods well known in the art. Substantially purified HNEU can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

10 This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding HNEU specifically compete with a test compound for binding HNEU. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HNEU.

#### **Uses of the Polynucleotide Encoding HNEU**

15 A polynucleotide designated herein as hneu, or any part thereof, may be used for diagnostic and/or therapeutic purposes. For diagnostic purposes, the hneu of this invention may be used to detect and quantitate gene expression in biopsied tissues in which expression of hneu may be implicated. The diagnostic assay is useful to distinguish between absence, presence, and excess expression of hneu and to monitor regulation of hneu levels during therapeutic intervention. The association of HNEU with disorders and disease conditions in specific tissues would greatly facilitate studies aimed at determining HNEU function in these conditions and the development of therapeutic strategies to treat them. Included in the scope of the invention are oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs.

25 Another aspect of the subject invention is to provide for hybridization or PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HNEU or closely related molecules. The specificity of the probe, whether it is made from a highly specific region, eg, 10 unique nucleotides in the 5' regulatory region, or a less specific region, eg, especially in the 3' region, and the stringency of the hybridization or amplification (maximal, high, intermediate or low) will determine whether the probe identifies only naturally occurring hneu, alleles or related sequences.

30 Probes may also be used for the detection of related sequences and should preferably contain at least 50% of the nucleotides from any of these HNEU encoding sequences. The

hybridization probes of the subject invention may be derived from the nucleotide sequence of SEQ ID NO:2 or from genomic sequence including promoter, enhancer elements and introns of the naturally occurring hneu. Hybridization probes may be labeled by a variety of reporter groups, including radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Other means for producing specific hybridization probes for hneu DNAs include the cloning of nucleic acid sequences encoding HNEU or HNEU derivatives into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides.

#### Diagnostic Use

Polynucleotide sequences encoding HNEU may be used for the diagnosis of conditions or diseases with which the expression of HNEU is associated. For example, polynucleotide sequences encoding HNEU may be used in hybridization or PCR assays of fluids or tissues from biopsies to detect hneu expression. The form of such qualitative or quantitative methods may include southern or northern analysis, dot blot or other membrane-based technologies; PCR technologies; dip stick, pin, chip and ELISA technologies. All of these techniques are well known in the art and are the basis of many commercially available diagnostic kits.

The hneu nucleotide sequences disclosed herein provide the basis for assays that detect activation or induction of hneu associated with specific diseases. The hneu nucleotide sequence may be labeled by methods known in the art and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After an incubation period, the sample is washed with a compatible fluid which optionally contains a dye (or other label requiring a developer) if the nucleotide has been labeled with an enzyme. After the compatible fluid is rinsed off, the dye is quantitated and compared with a standard. If the amount of dye in the biopsied or extracted sample is significantly elevated over that of a comparable control sample, the nucleotide sequence has hybridized with nucleotide sequences in the sample, and the presence of elevated levels of hneu nucleotide sequences in the sample indicates the presence of the associated disease.

Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regime in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of disease, a normal or standard profile for hneu expression must be established. This is accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with hneu, or a portion thereof, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified  
5 by comparing the values obtained for normal subjects with a dilution series of hneu run in the same experiment where a known amount of substantially purified hneu is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients afflicted with HNEU-associated diseases. Deviation between standard and subject values is used to establish the presence of disease.

10 Once disease is established, a therapeutic agent is administered and a treatment profile is generated. Such assays may be repeated on a regular basis to evaluate whether the values in the profile progress toward or return to the normal or standard pattern. Successive treatment profiles may be used to show the efficacy of treatment over a period of several days or several months.

PCR, as described in US Patent Nos. 4,683,195 and 4,965,188, provides additional uses  
15 for oligonucleotides based upon the HNEU sequence. Such oligomers are generally chemically synthesized, but they may be generated enzymatically or produced from a recombinant source. Oligomers generally comprise two nucleotide sequences, one with sense orientation (5'→3') and one with antisense (3'←5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool  
20 of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Additionally, methods which may be used to quantitate the expression of a particular molecule include radiolabeling (Melby PC et al 1993 J Immunol Methods 159:235-44) or biotinylating (Duplaa C et al 1993 Anal Biochem 229-36) nucleotides, coamplification of a  
25 control nucleic acid, and standard curves onto which the experimental results are interpolated. Quantitation of multiple samples may be speeded up by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation. A definitive diagnosis of this type may allow health professionals to begin aggressive treatment and prevent further worsening of the condition.  
30 Similarly, further assays can be used to monitor the progress of a patient during treatment. Furthermore, the nucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of

nucleotide sequences that are currently known such as the triplet genetic code, specific base pair interactions, and the like.

### Therapeutic Use

For therapeutic purposes, an antisense molecule of hneu may provide a basis for treatment  
5 where down-regulation of the gene and consequent inhibition of its activity is desirable. Such  
may be the case where a mutation in hneu may be involved in an abnormality in brain  
development. Alternatively, sequences encoding HNEU itself may provide the basis for gene  
therapy in conditions where it may be desirable to increase expression of HNEU and hence  
increase its activity. This condition might exist where loss of brain function, such as occurs in  
10 Alzheimer's disease, might be reversed by increased expression of HNEU.

Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or  
from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted  
organ, tissue or cell population. Methods which are well known to those skilled in the art can be  
used to construct recombinant vectors which will express antisense HNEU. See, for example, the  
15 techniques described in Sambrook et al (supra) and Ausubel et al (supra).

The polynucleotides comprising full length cDNA sequence and/or its regulatory  
elements enable researchers to use hneu as an investigative tool in sense (Youssoufian H and HF  
Lodish 1993 Mol Cell Biol 13:98-104) or antisense (Eguchi et al (1991) Annu Rev Biochem  
60:631-652) regulation of gene function. Such technology is now well known in the art, and  
20 sense or antisense oligomers, or larger fragments, can be designed from various locations along  
the coding or control regions.

Genes encoding HNEU can be turned off by transfecting a cell or tissue with expression  
vectors which express high levels of a desired hneu fragment. Such constructs can flood cells  
with untranslatable sense or antisense sequences. Even in the absence of integration into the  
25 DNA, such vectors may continue to transcribe RNA molecules until all copies are disabled by  
endogenous nucleases. Transient expression may last for a month or more with a non-replicating  
vector (Mettler I, personal communication) and even longer if appropriate replication elements  
are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing  
30 antisense molecules, DNA, RNA or PNA, to the control regions of hneu, ie, the promoters,  
enhancers, and introns. Oligonucleotides derived from the transcription initiation site, eg,  
between -10 and +10 regions of the leader sequence, are preferred. The antisense molecules may



also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic  
5 advances using triplex DNA were reviewed by Gee JE et al (In: Huber BE and BI Carr (1994) Molecular and Immunologic Approaches, Futura Publishing Co, Mt Kisco NY).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage.  
10 Within the scope of the invention are engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of hneu.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20  
15 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method  
20 known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HNEU. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, antisense  
25 cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase  
30 linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine and wybutosine as well as acetyl-, methyl-, thio- and similarly modified

forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Methods for introducing vectors into cells or tissues include those methods discussed *infra* and which are equally suitable for *in vivo*, *in vitro* and *ex vivo* therapy. For *ex vivo* therapy, vectors are introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient is presented in US Patent Nos. 5,399,493 and 5,437,994, disclosed herein by reference. Delivery by transfection and by liposome are quite well known in the art.

Furthermore, the nucleotide sequences for hneu disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including but not limited to such properties as the triplet genetic code and specific base pair interactions.

#### **Detection and Mapping of Related Polynucleotide Sequences**

The nucleic acid sequence for HNEU can also be used to generate hybridization probes for mapping the naturally occurring genomic sequence. The sequence may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. These include *in situ* hybridization to chromosomal spreads, flow-sorted chromosomal preparations, or artificial chromosome constructions such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price CM (1993; Blood Rev 7:127-34) and Trask BJ (1991; Trends Genet 7:149-54).

The technique of fluorescent *in situ* hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY. Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a hneu on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

*In situ* hybridization of chromosomal preparations and physical mapping techniques such

as linkage analysis using established chromosomal markers may be used for extending genetic maps. For example, a sequence tagged site based map of the human genome was recently published by the Whitehead-MIT Center for Genomic Research (Hudson TJ et al (1995) Science 270:1945-1954). Often the placement of a gene on the chromosome of another mammalian species such as mouse (Whitehead Institute/MIT Center for Genome Research, Genetic Map of the Mouse, Database Release 10, April 28, 1995) may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome, such as ataxia telangiectasia (AT), has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti et al (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

#### **Pharmaceutical Compositions**

The present invention relates to pharmaceutical compositions which may comprise nucleotides, proteins, antibodies, agonists, antagonists, or inhibitors, alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. Any of these molecules can be administered to a patient alone, or in combination with other agents, drugs or hormones, in pharmaceutical compositions where it is mixed with excipient(s) or pharmaceutically acceptable carriers. In one embodiment of the present invention, the pharmaceutically acceptable carrier is pharmaceutically inert.

#### **Administration of Pharmaceutical Compositions**

Administration of pharmaceutical compositions is accomplished orally or parenterally. Methods of parenteral delivery include topical, intra-arterial (directly to the tumor), intramuscular, subcutaneous, intramedullary, intrathecal, intraventricular, intravenous, intraperitoneal, or intranasal administration. In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation

and administration may be found in the latest edition of "Remington's Pharmaceutical Sciences" (Maack Publishing Co, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral  
5 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.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the  
10 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, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; and gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired,  
15 disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores are provided with suitable coatings such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent  
20 mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, ie, 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 filler or binders such as lactose or  
25 starches, lubricants such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations for parenteral administration include aqueous solutions of active compounds. For injection, the pharmaceutical compositions of the invention may be  
30 formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as

sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection 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. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated 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.

### **Manufacture and Storage**

10 The pharmaceutical compositions of the present invention may be manufactured in a manner that known in the art, eg, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

The pharmaceutical composition may 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 aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder in 1mM-50 mM histidine, 0.1%-2% sucrose, 2%-7% mannitol at a pH range of 4.5 to 5.5 that is combined with buffer prior to use.

After pharmaceutical compositions comprising a compound of the invention formulated in a acceptable carrier have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HNEU, such labeling would include amount, frequency and method of administration.

### **Therapeutically Effective Dose**

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

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

A therapeutically effective dose refers to that amount of protein or its antibodies, antagonists, or inhibitors which ameliorate the symptoms or condition. Therapeutic efficacy and toxicity of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, eg, ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. 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 of such compounds lies preferably within a range of circulating concentrations that include the ED50 with 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 is chosen by the individual physician in view of the patient to be treated. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Additional factors which may be taken into account include the severity of the disease state, eg, tumor size and location; age, weight and gender of the patient; diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long acting pharmaceutical compositions might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may 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. See US Patent Nos. 4,657,760; 5,206,344; or 5,225,212. 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.

It is contemplated, for example, that molecules or compounds that modulate HNEU activity, such as antibodies of HNEU, or an HNEU derivative can be delivered in a suitable formulation as a therapeutic agent. Similarly, administration of agonists should also improve the activity or lifespan of this protein and lessen the onset and progression of senescence.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

## INDUSTRIAL APPLICABILITY

### I cDNA Library Construction

The normal tissue used for cerebellum cDNA library construction was obtained from a 69 year-old, Caucasian male (Lot #RT95-05-0301; International Institute for Advanced Medicine, 5 Exton PA). The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury NJ). The reagents and extraction procedures were used as supplied in the Stratagene RNA Isolation Kit (Catalog # 200345; Stratagene). The lysate was centrifuged over a 5.7 M CsCl cushion using an Beckman SW28 rotor in a Beckman L8-70M Ultracentrifuge (Beckman Instruments) for 18 hours at 25,000 rpm 10 at ambient temperature. The RNA was extracted twice with phenol chloroform pH 8.0 and once with acid phenol pH 4.0, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, and resuspended in water and DNase treated for 15 min at 37°C. The RNA was isolated using the Qiagen Oligotex kit (QIAGEN Inc, Chatsworth CA) and used to construct the cDNA library.

The poly-A<sup>+</sup> RNA was handled according to the recommended protocols in the 15 SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (Catalog #18248-013; Gibco/BRL). First strand cDNA synthesis was accomplished using oligo d(T) priming and second strand synthesis was performed using a combination of DNA polymerase I, *E. coli* ligase and RNase H. The cDNA was blunted with T4 polymerase, and a Sal I linker was added to the blunt ended cDNA. The Sal I adapted, double stranded cDNAs were the digested with Not I and 20 fractionated on a Sepharose CL4B column (catalog #275105, Pharmacia), and those cDNAs exceeding 400 bp were ligated into pSport I. The plasmid pSport I was subsequently transformed into DH5a<sup>TM</sup> competent cells (Cat. #18258-012, Gibco/BRL).

### II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the Miniprep Kit (Catalog 25 #77468; Advanced Genetic Technologies Corporation, Gaithersburg MD). This kit consists of a 96-well block with reagents for 960 purifications. The recommended protocol was employed except for the following changes: 1) the 96 wells were each filled with only 1 ml of sterile Terrific Broth (Catalog #22711, LIFE TECHNOLOGIES<sup>TM</sup>) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) the bacteria were cultured for 24 hours after the wells were inoculated and 30 then lysed with 60  $\mu$ l of lysis buffer; 3) a centrifugation step employing the Beckman GS-6R rotor at 2900 rpm for 5 minutes was performed before the contents of the block were added to the primary filter plate; and 4) the optional step of adding isopropanol to TRIS buffer was not

routinely performed. After the last step in the protocol, samples were transferred to a Beckman 96-well block for storage.

The cDNAs were sequenced by the method of Sanger F and AR Coulson (1975; J Mol Biol 94:441f), using a Hamilton Micro Lab 2200 (Hamilton, Reno NV) in combination with four  
5 Peltier Thermal Cyclers (PTC200 from MJ Research, Watertown MA) and Applied Biosystems 377 or 373 DNA Sequencing Systems, and the reading frame was determined.

### III Homology Searching of cDNA Clones and Their Deduced Proteins

Each cDNA was compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 Sequence Analysis System. In  
10 this algorithm, Pattern Specification Language (TRW Inc, Los Angeles CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value.  
15 Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT™ 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern  
20 Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

BLAST, which stands for Basic Local Alignment Search Tool (Altschul SF (1993) J Mol  
25 Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10), was used to search for local sequence alignments. BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. BLAST is useful for matches which do not contain gaps. The fundamental unit of BLAST algorithm output is the  
30 High-scoring Segment Pair (HSP).

An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or



cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

#### IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labelled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al. supra).

Analogous computer techniques using BLAST (Altschul SF 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte, Palo Alto CA). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

100

and it takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

#### V Extension of HNEU to Full Length or to Recover Regulatory Elements

The nucleic acid sequence of full length hneu (SEQ ID NO:2) is used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5' sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized to extend sequence in the sense direction (XLF). Primers allow the extension of the known hneu sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest (US Patent Application 08/487,112, filed June 7, 1995, specifically incorporated by reference). The initial

primers are designed from the cDNA using OLIGO<sup>®</sup> 4.06 Primer Analysis Software (National Biosciences), or another appropriate program, 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. Any stretch of nucleotides which would result in hairpin structures and primer-primer

5 dimerizations is avoided.

The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5' upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing  
10 the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR is performed using the Peltier Thermal Cycler (PTC200; MJ Research, Watertown MA) and the following parameters:

	Step 1	94% C for 1 min (initial denaturation)
15	Step 2	65% C for 1 min
	Step 3	68% C for 6 min
	Step 4	94% C for 15 sec
	Step 5	65% C for 1 min
	Step 6	68% C for 7 min
20	Step 7	Repeat step 4-6 for 15 additional cycles
	Step 8	94% C for 15 sec
	Step 9	65% C for 1 min
	Step 10	68% C for 7:15 min
	Step 11	Repeat step 8-10 for 12 cycles
25	Step 12	72% C for 8 min
	Step 13	4% C (and holding)

A 5-10  $\mu$ l aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in  
30 extending the sequence. Bands thought to contain the largest products were selected and cut out of the gel. Further purification involves using a commercial gel extraction method such as QIAQuick<sup>™</sup> (QIAGEN Inc). After recovery of the DNA, Klenow enzyme was used to trim

single-stranded, nucleotide overhangs creating blunt ends which facilitate religation and cloning.

After ethanol precipitation, the products are redissolved in 13  $\mu$ l of ligation buffer, 1  $\mu$ l T4-DNA ligase (15 units) and 1  $\mu$ l T4 polynucleotide kinase are added, and the mixture is incubated at room temperature for 2-3 hours or overnight at 16 $^{\circ}$  C. Competent *E. coli* cells (in 5 40  $\mu$ l of appropriate media) are transformed with 3  $\mu$ l of ligation mixture and cultured in 80  $\mu$ l of SOC medium (Sambrook J et al, supra). After incubation for one hour at 37 $^{\circ}$  C, the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook J et al, supra) containing 2xCarb. The following day, several colonies are randomly picked from each plate and cultured in 150  $\mu$ l of liquid LB/2xCarb medium placed in an individual well of an appropriate, 10 commercially-available, sterile 96-well microtiter plate. The following day, 5  $\mu$ l of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5  $\mu$ l of each sample is transferred into a PCR array.

For PCR amplification, 18  $\mu$ l of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer and one or both of the gene specific primers used for 15 the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1	94 $^{\circ}$ C for 60 sec
Step 2	94 $^{\circ}$ C for 20 sec
Step 3	55 $^{\circ}$ C for 30 sec
20 Step 4	72 $^{\circ}$ C for 90 sec
Step 5	Repeat steps 2-4 for an additional 29 cycles
Step 6	72 $^{\circ}$ C for 180 sec
Step 7	4 $^{\circ}$ C (and holding)

Aliquots of the PCR reactions are run on agarose gels together with molecular weight 25 markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid and sequenced.

## VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 30 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 mCi of [ $^{-32}$ P]

adenosine triphosphate (Amersham, Chicago IL) and T4 polynucleotide kinase (DuPont NEN®, Boston MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 super fine resin column (Pharmacia). A portion containing 10<sup>7</sup> counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of  
5 human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II; DuPont NEN®).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room  
10 temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester NY) is exposed to the blots in a Phosphorimager cassette (Molecular Dynamics, Sunnyvale CA) for several hours, hybridization patterns are compared visually.

#### VII Antisense Molecules

The hneu sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring hneu. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequence of HNEU as shown in Figure 1 is used to inhibit expression of naturally occurring HNEU. The complementary oligonucleotide is  
20 designed from the most unique 5' sequence as shown in Figure 1 and used to inhibit translation of an hneu transcript by preventing the ribosome from binding. Using an appropriate portion of the leader and 5' sequence of SEQ ID NO:2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or early coding sequence of the polypeptide as shown in Figure 1.

#### 25 VIII Expression of HNEU

Expression of the HNEU is accomplished by subcloning the cDNAs into appropriate vectors and transfecting the vectors into host cells. In this case, the cloning vector, pSport, previously used for the generation of the cDNA library is used to express HNEU in E. coli. Upstream of the cloning site, this vector contains a promoter for  $\beta$ -galactosidase, followed by  
30 sequence containing the amino-terminal Met and the subsequent 7 residues of  $\beta$ -galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

Induction of an isolated, transfected bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first seven residues of  $\beta$ -galactosidase, about 5 to 15 residues of linker, and the full length HNEU. The signal sequence directs the secretion of HNEU into the bacterial growth media which can be used directly in the following assay for  
5 activity.

### **IX HNEU Activity**

HNEU activity may be measured through its association with senescence in neuronal cells exhibiting decreased expression of neuronatin. A cultured human neuronal cell line, such as hNT2, expressing neuronatin is treated with an inhibitor of neuronatin and the senescence of cells  
10 due to decreased neuronatin activity is measured. Cell senescence is measured by measuring changes in cell morphology using light or phase-contrast microscopy.

### **X Production of HNEU Specific Antibodies**

HNEU substantially purified using PAGE electrophoresis (Sambrook, supra) is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence  
15 translated from HNEU is analyzed using DNASTar software (DNASTar Inc) to determine regions of high immunogenicity and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Analysis to select appropriate epitopes, such as those near the C-terminus or in hydrophilic regions (shown in Figures 4 and 5) is described by Ausubel FM et al (supra).

20 Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel FM et al, supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity,  
25 for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

### **XI Purification of Naturally Occurring HNEU Using Specific Antibodies**

Naturally occurring or recombinant HNEU is substantially purified by immunoaffinity chromatography using antibodies specific for HNEU. An immunoaffinity column is constructed  
30 by covalently coupling HNEU antibody to an activated chromatographic resin such as CnBr-activated Sepharose (Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HNEU is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HNEU (eg, high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HNEU binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope such as urea or thiocyanate ion), and HNEU is collected.

## **XII Identification of Molecules Which Interact with HNEU**

HNEU, or biologically active fragments thereof, are labelled with <sup>125</sup>I Bolton-Hunter reagent (Bolton, AE and Hunter, WM (1973) Biochem J 133:529). Candidate molecules previously arrayed in the wells of a 96 well plate are incubated with the labelled HNEU, washed and any wells with labelled HNEU complex are assayed. Data obtained using different concentrations of HNEU are used to calculate values for the number, affinity, and association of HNEU with the candidate molecules.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

## (1) GENERAL INFORMATION

- (i) APPLICANT: INCYTE PHARMACEUTICALS, INC.
- (ii) TITLE OF THE INVENTION: A NOVEL HUMAN NEURONATIN
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
  - (B) STREET: 3174 Porter Drive
  - (C) CITY: Palo Alto
  - (D) STATE: CA
  - (E) COUNTRY: U.S.
  - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: DOS
  - (D) SOFTWARE: FastSEQ Version 1.5
- (vi) CURRENT APPLICATION DATA:
  - (A) PCT APPLICATION NUMBER: To Be Assigned
  - (B) FILING DATE: Herewith
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/690,094
  - (B) FILING DATE: 31-JULY-1996
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Billings, Lucy J.
  - (B) REGISTRATION NUMBER: 36,749
  - (C) REFERENCE/DOCKET NUMBER: PF-0089 PCT
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 415-855-0555
  - (B) TELEFAX: 415-845-4166

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus

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

```

Met Ala Ala Val Ala Ala Ala Ser Ala Glu Leu Leu Ile Ile Gly Trp
 1               5              10             15
Tyr Ile Phe Arg Val Leu Leu Gln Val Phe Arg Tyr Ser Leu Gln Lys
                20              25             30
Leu Ala Tyr Thr Val Ser Arg Thr Gly Arg Gln Val Leu Gly Glu Arg
                35              40             45
Arg Gln Arg Ala Pro Asn
 50
  
```

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 254 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: Consensus

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

CCGAGACCAG	CGGATCTCGG	CAAACCTCT	TTCTCGACCA	CCCACCTACC	ATTCTTGAA	60
CCATGGCGGC	AGTGGCGGCG	GCCTCGGCTG	AACTGCTCAT	CATCGGCTGG	TACATCTTCC	120
GCGTGCTGCT	GCAGGTGTTT	AGGTACTCCC	TGCAGAAGCT	GGCATAACAG	GTGTCGCGGA	180
CCGGGCGGCA	GGTGTGGGG	GAGCGCAGGC	AGCGAGCCCC	CAACTGAGGC	CCCAGCTTCC	240
AGCCCTGGGC	GGCC					254

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 260 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: CRBLNOT01
  - (B) CLONE: 674382

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

GCGGATCTCG	GCAAACCCTC	TTTCTCGACC	ACCCACCTAC	CATTCTTGGA	ACCATGGCGG	60
CAGTGGCGGN	GGCCTNGGCT	GAAGTGGTCA	TCATCGGCTG	GTACATNTTC	CGCGTGCTGC	120
TGCAGGTGTT	CAGGTACTNC	CTGCAGAAGC	TGGCATAAC	GGTGTGNGG	ACCGGGCGGC	180
AGGTGTTGGG	GGAGCGCAGG	CAGCGAGCCC	NCAACTGAGG	NCCCAGCTTC	CCAGCCCTGG	240
GCGGCCGTAT	NATNAGGTGT					260

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 257 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: BRAINOT04
  - (B) CLONE: 926572

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

GCGGACTNCG	AGACCAGCGG	ATCTNGGCAA	ACCCTNTTTN	TNGACCACCC	ACCTACCATT	60
NTTGGAAACCA	TGGCGGCAGT	GGCGGCGGCC	TCGGCTGAAC	TGNTCATNAT	NGGNTGGTAC	120
ATNTCCGCG	TGCTGCTGCA	GGTGTTCAGT	ACTNCNTGCA	GAAGCTGGNA	TANACGGTGT	180
CGCGGACCGG	GCGGCAGGTG	TTGGGGGAGC	GCAGGCAGCG	AGCCCNAAATG	AGGCCCCAGT	240
NCNAGCCCTG	GNGGCC					257

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 54 amino acids



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

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 619500

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

```

Met Ala Ala Val Ala Ala Ala Ser Ala Glu Leu Leu Ile Ile Gly Trp
 1           5           10           15
Tyr Ile Phe Arg Val Leu Leu Gln Val Phe Arg Tyr Ser Leu Gln Lys
          20           25           30
Leu Ala His Thr Val Ser Arg Thr Gly Arg Gln Val Leu Gly Glu Arg
          35           40           45
Arg Gln Arg Ala Pro Asn
 50
    
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1244410

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

```

Met Ala Ala Val Ala Ala Ala Ser Ala Glu Leu Leu Ile Ile Gly Trp
 1           5           10           15
Tyr Ile Phe Arg Val Leu Leu Gln Val Phe Arg Tyr Ser Leu Gln Lys
          20           25           30
Leu Ala Tyr Thr Val Ser Arg Thr Gly Arg Gln Val Leu Gly Glu Arg
          35           40           45
Arg Gln Arg Ala Pro Asn
 50
    
```

## CLAIMS

1. A substantially purified human neuronatin polypeptide comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof.
2. The isolated and purified polynucleotide sequence encoding the polypeptide of claim 1.
3. The isolated and purified polynucleotide sequence of claim 2 consisting of the sequence of SEQ ID NO:2 or variants thereof.
4. A polynucleotide sequence fully complementary to the sequence of SEQ ID NO:2 or variants thereof.
5. The isolated and purified polynucleotide sequence of claim 2 consisting of a polynucleotide sequence that hybridizes under stringent hybridization conditions to the sequence of SEQ ID NO:2.
6. A recombinant expression vector containing the polynucleotide sequence of claim 2.
7. A recombinant host cell comprising the polynucleotide sequence of claim 2.
8. A method for producing a polypeptide comprising the amino acid sequence shown in SEQ ID NO:1, the method comprising the steps of:
  - a) culturing the host cell of Claim 7 under conditions suitable for the expression of the polypeptide; and
  - b) recovering the polypeptide from the host cell culture.
9. A purified antibody which binds specifically to the polypeptide of claim 1.
10. A pharmaceutical composition comprising a substantially purified human neuronatin (SEQ ID NO:1) in conjunction with a suitable pharmaceutical carrier.

```

9      18      27      36      45      54
5' NCC GAG ACC AGC GGA TCT CGG CAA ACC CTC TTT CTC GAC CAC CCA CCT ACC ATT

63      72      81      90      99      108
CTT GGA ACC ATG GCG GCA GTG GCG GCC TCG GCT GAA CTG CTC ATC ATC GGC
M A A V A A A S A E L L I I G

117      126      135      144      153      162
TGG TAC ATC TTC CGC GTG CTG CTG CAG GTG TTC AGG TAC TCC CTG CAG AAG CTG
W Y I F R V L L Q V F R Y S L L Q K L

171      180      189      198      207      216
GCA TAC ACG GTG TCG CGG ACC GGG CGG CAG GTG TTG GGG GAG CGC AGG CAG CGA
A Y T V S R T G R Q V L G E R R Q R

225      234      243      252
GCC CCC AAC TGA GGC CCC AGC TTC CAG CCC TGG GCG GCC 3'
A P N
    
```

FIGURE 1

1	M	A	V	A	A	A	S	A	E	L	L	I	I	G	W	Y	I	F	R	V	L	L	Q	V	F	R	Y	S	L	Q	K	L	A	Y	T	V	S	R	T	gi926572 HNEU
1	M	A	V	A	A	A	S	A	E	L	L	I	I	G	W	Y	I	F	R	V	L	L	Q	V	F	R	Y	S	L	Q	K	L	A	H	T	V	S	R	T	gi619500 Rat
1	M	A	V	A	A	A	S	A	E	L	L	I	I	G	W	Y	I	F	R	V	L	L	Q	V	F	R	Y	S	L	Q	K	L	A	Y	T	V	S	R	T	gi1244410 Human
41	G	R	Q	V	L	G	E	R	R	Q	R	A	P	N																									gi926572 HNEU	
41	G	R	Q	V	L	G	E	R	R	Q	R	A	P	N																									gi619500 Rat	
41	G	R	Q	V	L	G	E	R	R	Q	R	A	P	N																									gi1244410 Human	

FIGURE 2

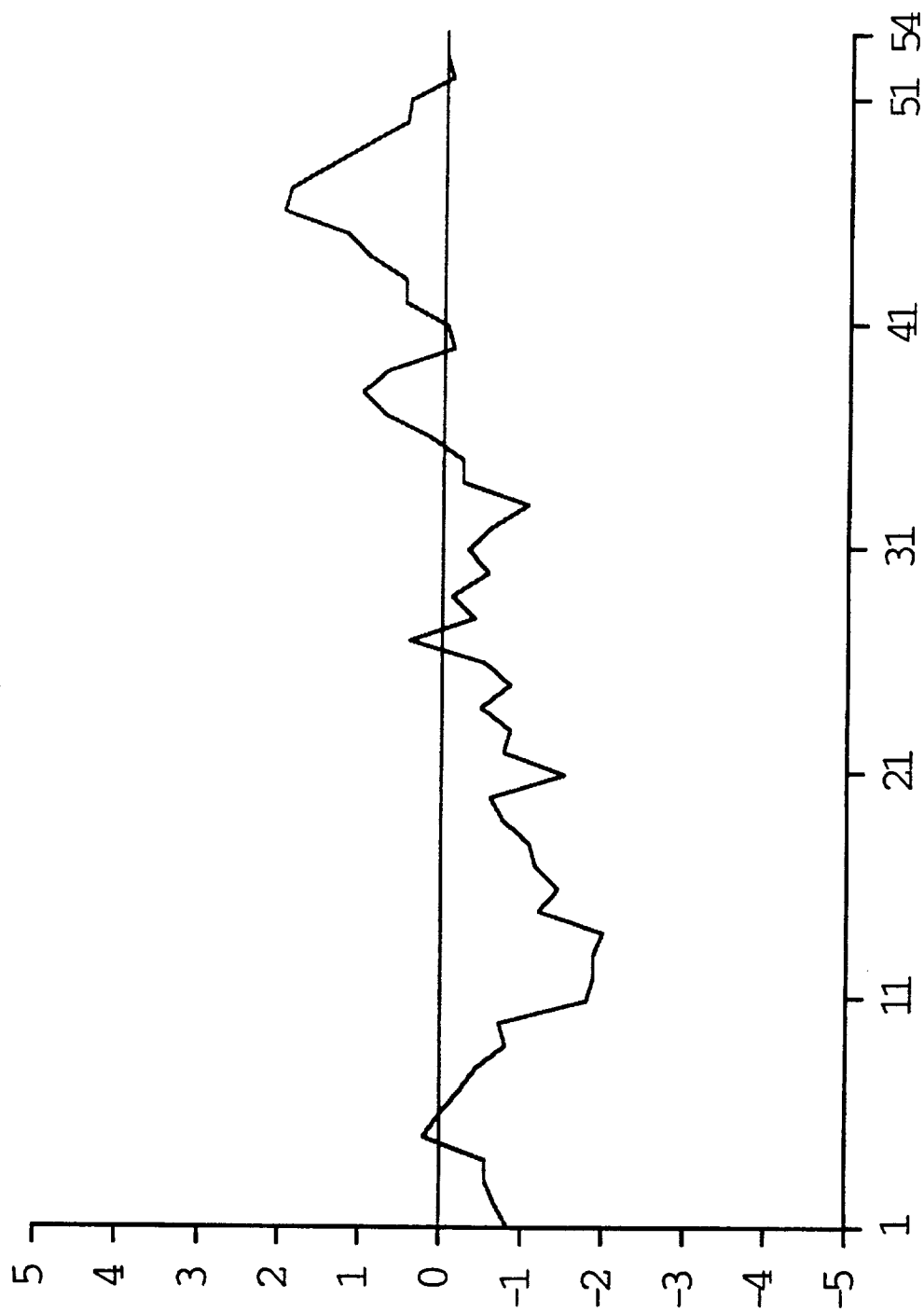


FIGURE 3

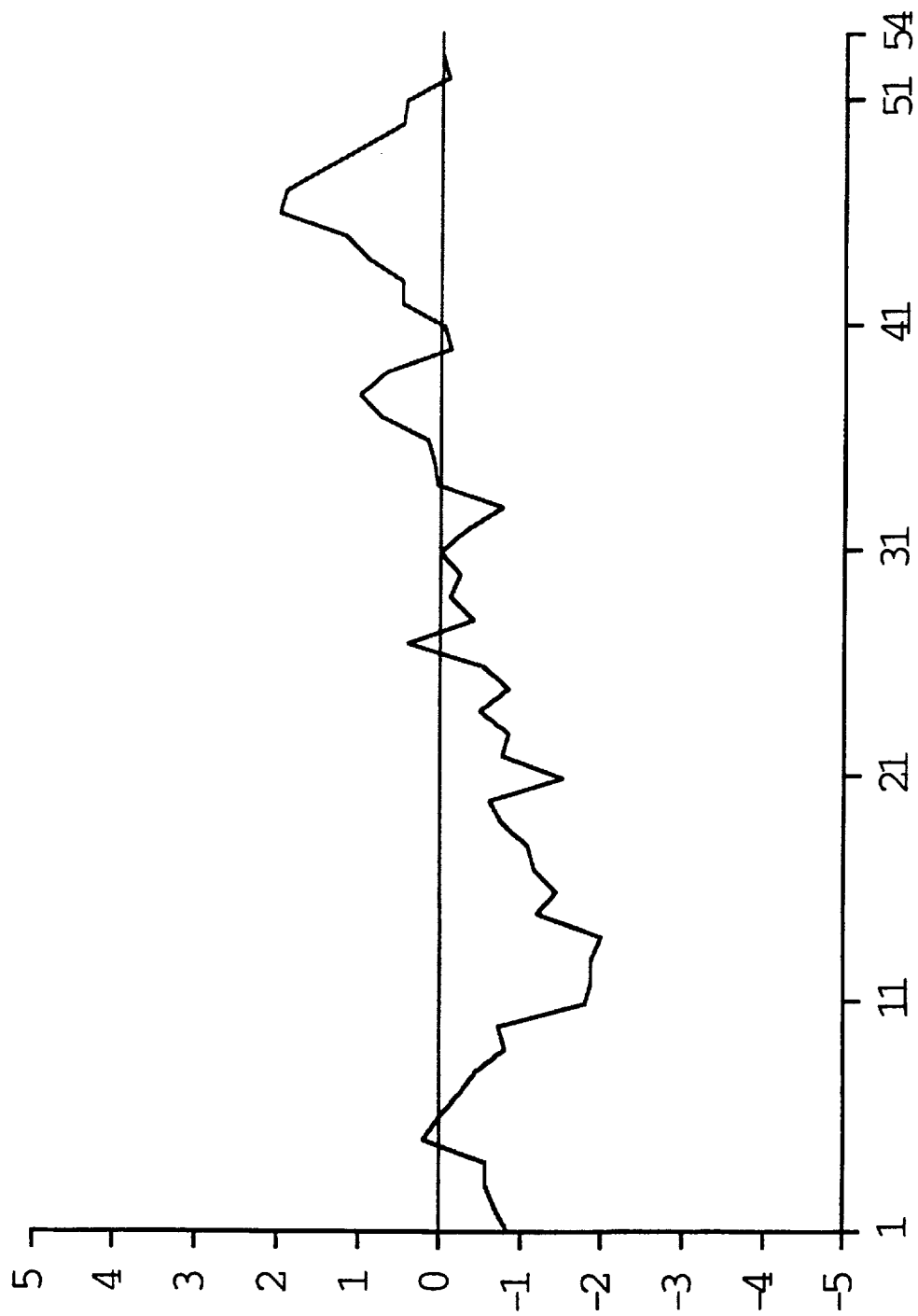


FIGURE 4