

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 May 2002 (02.05.2002)

PCT

(10) International Publication Number
WO 02/34920 A2

(51) International Patent Classification⁷: **C12N 15/53**,
15/62, 15/11, 5/10, C07K 14/81, 16/38, C12Q 1/68, G01N
33/50, A61K 39/395, 48/00

(21) International Application Number: PCT/EP01/12158

(22) International Filing Date: 22 October 2001 (22.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/242,426 24 October 2000 (24.10.2000) US
60/313,796 22 August 2001 (22.08.2001) US

(71) Applicant (for all designated States except US): **BAYER
AKTIENGESSELLSCHAFT** [DE/DE]; 51368 Lev-
erkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **ZHU, Zhimin**
[CN/US]; 45 Hinckley Road, Waban, MA 02468 (US).

(74) Common Representative: **BAYER AKTIENGE-
SELLSCHAFT**; 51368 Leverkusen (DE).

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

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

Published:

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

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

(54) Title: REGULATION OF HUMAN CERULOPLASMIN-LIKE PROTEIN

(57) Abstract: Reagents which regulate human ceruloplasmin-like protein activity and reagents which bind to human ceruloplasmin-like protein gene products can be used to regulate ceruloplasmin activity. Such regulation is particularly useful for treating cardiovascular disorders, liver disorders, such as Wilson's disease, and CNS disorders, such as Parkinson's disease.



WO 02/34920 A2

REGULATION OF HUMAN CERULOPLASMIN-LIKE PROTEIN

This application claims the benefit of and incorporates by reference co-pending provisional applications Serial No. 60/242,426 filed October 11, 2000 and Serial No.
5 60/313,796 filed August 22, 2001.

TECHNICAL FIELD OF THE INVENTION

The invention relates to the area of regulation of human ceruloplasmin-like protein.
10 More particularly, the invention relates to the regulation of human ceruloplasmin-like protein activity to provide therapeutic effects.

BACKGROUND OF THE INVENTION

15 Ceruloplasmin is a copper-containing glycoprotein plasma protein found in a wide variety of human tissues, including the liver, brain, breast, fetal heart, colon, and blood. Structurally, ceruloplasmin has homology to a copper-binding domain found in lactase and ascorbate oxidase. Ceruloplasmin exhibits a range of enzymatic activities, such as ferroxidase activity, amine oxidase activity, copper transport and
20 homeostasis, and superoxide dismutase activity. It is postulated to play an important role in the metabolism of copper and iron homeostasis. *See Harris et al., Proc. Natl. Acad. Sci. U.S.A.* 96:10812- 17, 1999. Ceruloplasmin is found in the β_2 macroglobulin fraction. Normal human serum contains 20-50 mg/dl ceruloplasmin. *See U.S. Patent Nos. 6,010,903 and 4,772, 894.*

25 Inappropriate ceruloplasmin activity has been implicated in several diseases, particularly diseases in which copper metabolism is defective. For example, decreased ceruloplasmin activity is correlated with liver disorders such as Wilson's disease, an autosomal recessive disorder of copper metabolism in which the liver is
30 unable to properly eliminate copper. Increased serum ceruloplasmin levels have been correlated to an increased risk of myocardial infarction and cardiovascular disease.

Ceruloplasmin has also been shown to play a role in low density lipoprotein (LDL) oxidation, leading to atherosclerosis. In addition to the aforementioned disorders, there are many physiological disorders related to inappropriate accumulation of copper in tissues.

5

Thus, there is a need in the art for identifying new ceruloplasmin-like proteins and methods of regulating their activity.

SUMMARY OF THE INVENTION

10

It is an object of the invention to provide reagents and methods of regulating ceruloplasmin activity. These and other objects of the invention are provided by one or more of the embodiments described below.

15

One embodiment of the invention is a ceruloplasmin-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2; and

20

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

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

25

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2; and

30

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

- 3 -

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

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

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

the nucleotide sequence shown in SEQ ID NO: 1.

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

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a ceruloplasmin-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2; and

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

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

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

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

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

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a ceruloplasmin-like protein polypeptide or the product
30 encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

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

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

Ceruloplasmin-like protein activity in the cell is thereby decreased.

10 The invention thus provides a human ceruloplasmin-like protein which can be used to identify test compounds which may act, for example, as agonists or antagonists at the enzyme's active site. Human ceruloplasmin-like protein and fragments thereof also are useful in raising specific antibodies which can effectively reduce the protein's activity.

15 **BRIEF DESCRIPTION OF THE DRAWINGS**

Fig. 1 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (part of SEQ ID NO:1).

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

20

Fig. 3 shows the DNA-sequence including the CDS of a ceruloplasmin-like protein Polypeptide (SEQ ID NO:1).

Fig. 4 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:3)

25

Fig. 5 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:4).

30 Fig. 6 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:5).

- 6 -

- Fig. 7 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:6).
Fig. 8 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:7).
- 5 Fig. 9 shows the amino acid sequence of the protein identified by SwissProt Accession No. P00450 (SEQ ID NO:8).
- Fig. 10 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:9).
- 10 Fig. 11 shows the DNA-sequence encoding a ceruloplasmin-like protein Polypeptide (SEQ ID NO:10).
- Fig. 12 shows the amino acid sequence of a ceruloplasmin-like protein Polypeptide (SEQ ID NO:11).
- 15 Fig. 13 shows the BLASTP alignment of human ceruloplasmin-like protein (SEQ ID NO:2) against Swiss Prot Accession No. P00450 (SEQ ID NO:8).
- 20 Fig. 14 shows the Prosite search results.
- Fig. 15 shows the BLOCKS search results.
- Fig. 16 shows the alignment of human ceruloplasmin-like protein (amino acids 1-99 of SEQ ID NO:2) against pfam_hmm_Cu-oxidase multicopper oxidase.
- 25 Fig. 17 shows the alignment of human ceruloplasmin-like protein (amino acids 313-415 of SEQ ID NO:2) against pfam_hmm_Cu-oxidase multicopper oxidase.
- 30

Fig. 18 shows the alignment of human ceruloplasmin-like protein (amino acids 656-798 of SEQ ID NO:2) against pfam_hmm_Cu-oxidase multicopper oxidase.

5 Fig. 19 shows the BLASTP - alignment of human ceruloplasmin-like protein (SEQ ID NO:11) against swiss|P00450|CERU_HUMAN (SEQ ID NO:8).

DETAILED DESCRIPTION OF THE INVENTION

10

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

- a) a polynucleotide encoding a ceruloplasmin-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:
 - 15 amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2; and
 - the amino acid sequence shown in SEQ ID NO: 2.
- b) a polynucleotide comprising the sequence of SEQ ID NO: 1;
- c) a polynucleotide which hybridizes under stringent conditions to a
- 20 polynucleotide specified in (a) and (b);
- d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and
- e) a polynucleotide which represents a fragment, derivative or allelic variation of
- 25 a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that a novel protein, particularly a human ceruloplasmin-like protein, is a discovery of the present invention. Human ceruloplasmin-like protein as shown in SEQ ID NO:11 is 100% identical over 1102 amino acids to the *Homo sapiens* protein identified by Accession

30 No. AF148860. Human ceruloplasmin-like protein as shown in SEQ ID NO:11 is

49% identical over 1052 amino acids to the *Homo sapiens* protein identified by Swiss Prot Accession No. P00450 (SEQ ID NO:8) and annotated as a putative human ceruloplasmin precursor (EC 1.16.3.1) (Fig. 19). Human ceruloplasmin-like protein contains a multicopper oxidase domain which is shown in bold in Fig. 15.

5

A coding sequence for SEQ ID NO:11 is shown in SEQ ID NO:10. Related ESTs (SEQ ID NOS:3-7) are expressed in brain, and fetal heart.

Human ceruloplasmin-like protein is expected to be useful for the same purposes as previously identified ceruloplasmin (*see* Mohrenweiser and Decker, *Hum Hered* 32(6): 369-73, 1982). Regulators of a human ceruloplasmin-like protein can be used to regulate ceruloplasmin activity. Human ceruloplasmin-like protein is expected to be especially useful for treating cardiovascular diseases, liver disorders, and CNS disorders.

15

Polypeptides

Ceruloplasmin-like protein polypeptides according to the invention comprise at least 6, 10, 15, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 250, 275, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150 or 1158 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO:2 or 11 or a biologically active variant of the amino acid sequence shown in SEQ ID NO:2 or 11, as defined below. A ceruloplasmin-like protein polypeptide of the invention therefore can be a portion of a ceruloplasmin-like protein molecule, a full-length ceruloplasmin-like protein molecule, or a fusion protein comprising all or a portion of a ceruloplasmin-like protein molecule.

25

Biologically Active Variants

Ceruloplasmin-like protein variants which are biologically active, *i.e.*, retain a ceruloplasmin-like protein activity, also are ceruloplasmin-like protein polypeptides. Preferably, naturally or non-naturally occurring ceruloplasmin-like protein variants have amino acid sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to an amino acid sequence shown in SEQ ID NO:2 or 11. Percent identity between a putative ceruloplasmin-like protein variant and an amino acid sequence of SEQ ID NO:2 or 11 is determined using the Blast2 alignment program.

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

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active ceruloplasmin-like protein polypeptide can readily be determined by assaying for ceruloplasmin activity, as is known in the art and described, for example, in U.S. Patent No. 4,722,894.

Fusion Proteins

Fusion proteins are useful for generating antibodies against ceruloplasmin-like protein amino acid sequences and for use in various assay systems. For example,

fusion proteins can be used to identify proteins which interact with portions of a ceruloplasmin-like protein polypeptide, including its active site. Methods such as protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A ceruloplasmin-like protein fusion protein comprises two protein segments fused together by means of a peptide bond. Contiguous amino acids for use in a fusion protein can be selected from the amino acid sequence shown in SEQ ID NO:2 or 11 or from a biologically active variants of those sequences, such as those described above. For example, the first protein segment can comprise at least 6, 10, 15, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150 or 1158 contiguous amino acids of SEQ ID NO:2 or 11 or a biologically active variant. Preferably, a fusion protein comprises the enzyme's active site. The first protein segment also can comprise full-length ceruloplasmin-like protein.

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

protein sequence, so that the ceruloplasmin-like protein polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a
5 fusion protein is produced by covalently linking two protein segments or by standard
procedures in the art of molecular biology. Recombinant DNA methods can be used
to prepare fusion proteins, for example, by making a DNA construct which
comprises ceruloplasmin-like protein coding sequences disclosed herein in proper
reading frame with nucleotides encoding the second protein segment and expressing
10 the DNA construct in a host cell, as is known in the art. Many kits for constructing
fusion proteins are available from companies such as Promega Corporation
(Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA),
Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC;
Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-
15 KITS).

Identification of Species Homologs

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

25 Polynucleotides

A ceruloplasmin-like protein polynucleotide can be single- or double-stranded and
comprises a coding sequence or the complement of a coding sequence for a
ceruloplasmin-like protein polypeptide. A coding sequence of a ceruloplasmin-like
30 protein polynucleotide is shown in SEQ ID NO:1 or 10.

- 12 -

Degenerate nucleotide sequences encoding human ceruloplasmin-like protein polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to the ceruloplasmin-like protein coding sequences nucleotide sequence shown in SEQ ID NO:1 or 10 also
5 are ceruloplasmin-like protein polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of ceruloplasmin-like protein
10 polynucleotides which encode biologically active ceruloplasmin-like protein polypeptides also are ceruloplasmin-like protein polynucleotides.

Identification of Variants and Homologs

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

Species homologs of the ceruloplasmin-like protein polynucleotides disclosed herein can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human
30 variants of ceruloplasmin-like protein polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m

of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol.* 81, 123 (1973). Variants of human ceruloplasmin-like protein polynucleotides or ceruloplasmin-like protein polynucleotides of other species can therefore be identified, for example, by hybridizing a putative homologous ceruloplasmin-like protein polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or 10. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising ceruloplasmin-like protein polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

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

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

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

where l = the length of the hybrid in basepairs.

- 14 -

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

5 Preparation of Polynucleotides

A naturally occurring ceruloplasmin-like protein polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, synthesized using an amplification technique, such as the
10 polymerase chain reaction (PCR), or synthesized using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated ceruloplasmin-like protein polynucleotides. For example, restriction enzymes and probes
15 can be used to isolate polynucleotide fragments which comprise ceruloplasmin-like protein nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

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

Alternatively, synthetic chemistry techniques can be used to synthesize ceruloplasmin-like protein polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a ceruloplasmin-like protein polypeptide having, for example, the amino acid sequence shown in SEQ
30 ID NO:2 or 11 or a biologically active variant of that sequence.

Extending Polynucleotides

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

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

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

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

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

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

Obtaining Polypeptides

Ceruloplasmin-like protein polypeptides can be obtained, for example, by purification from human cells, by expression of ceruloplasmin-like protein polynucleotides, or by direct chemical synthesis.

Protein Purification

Ceruloplasmin-like protein polypeptides can be purified from human cells, such as primary cells or cell lines. A purified ceruloplasmin-like protein polypeptide is separated from other compounds which normally associate with the ceruloplasmin-like protein polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified ceruloplasmin-like protein polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis. Enzymatic activity of the purified preparations can be assayed, for example, as described in U.S. Patent No. 4,722,894.

15

Expression of Polynucleotides

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

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a ceruloplasmin-like protein polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant

30

bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a ceruloplasmin-like protein polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the ceruloplasmin-like protein polypeptide. For example, when a large quantity of a ceruloplasmin-like protein polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT

(Stratagene), in which the sequence encoding the ceruloplasmin-like protein polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem.* 264, 5503-5509, 1989 or pGEX vectors (Promega, Madison, Wis.) can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or Factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

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

Plant and Insect Expression Systems

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

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

Mammalian Expression Systems

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

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

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding ceruloplasmin-like protein polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a ceruloplasmin-like protein polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ.* 20, 125-162, 1994).

Host Cells

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

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express ceruloplasmin-like protein

polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are
5 switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced ceruloplasmin-like protein sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

10

Any number of selection systems can be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell* 11, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell* 22, 817-23, 1980). Genes which can be employed in *tk* or *aprt* cells,
15 respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci.* 77, 3567-70, 1980); *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol.* 150, 1-14, 1981); and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin
20 acetyltransferase, respectively (Murray, 1992 *supra*). Additional selectable genes have been described, for example *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci.* 85, 8047-51, 1988). Visible markers
25 such as anthocyanins, β -glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol.* 55, 121-131, 1995).

Detecting Expression of Polypeptides

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

Alternatively, host cells which contain a ceruloplasmin-like protein polynucleotide and which express a ceruloplasmin-like protein polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein.

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

A variety of protocols for detecting and measuring the expression of a ceruloplasmin-like protein polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immuno-

sorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a ceruloplasmin-like protein polypeptide can be used, or a competitive binding assay can be employed. These and
5 other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med.* 158, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the
10 art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding ceruloplasmin-like protein polypeptides include oligo-labeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a ceruloplasmin-like protein
15 polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase, such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega,
20 and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

25 Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a ceruloplasmin-like protein polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed
30 cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors

containing polynucleotides which encode ceruloplasmin-like protein polypeptides can be designed to contain signal sequences which direct secretion of ceruloplasmin-like protein polypeptides through a prokaryotic or eukaryotic cell membrane.

5 Other constructions can be used to join a sequence encoding a ceruloplasmin-like protein polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow
10 purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the ceruloplasmin-like protein polypeptide can be used to facilitate purification. One
15 such expression vector provides for expression of a fusion protein containing a ceruloplasmin-like protein polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography as described in Porath *et al.*, *Prot. Exp. Purif.* 3, 263-281, 1992), while the enterokinase cleavage
20 site provides a means for purifying the ceruloplasmin-like protein polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol.* 12, 441-453, 1993).

Chemical Synthesis

25 Sequences encoding a ceruloplasmin-like protein polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al.* *Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a ceruloplasmin-like protein polypeptide itself can be
30 produced using chemical methods to synthesize its amino acid sequence. For example, ceruloplasmin-like protein polypeptides can be produced by direct peptide

synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc.* 85, 2149-2154, 1963; Roberge *et al.*, *Science* 269, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of ceruloplasmin-like protein polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

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

Production of Altered Polypeptides

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

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter ceruloplasmin-like protein polypeptide-encoding sequences for a variety of reasons, including modification of the cloning, processing,

and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

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

An antibody which specifically binds to an epitope of a ceruloplasmin-like protein polypeptide can be used therapeutically, as well as in immunochemical assays, including but not limited to Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a ceruloplasmin-like protein polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to ceruloplasmin-like protein poly-

peptides do not detect other proteins in immunochemical assays and can immunoprecipitate a ceruloplasmin-like protein polypeptide from solution.

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

15 Monoclonal antibodies which specifically bind to a ceruloplasmin-like protein polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature* 256, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods* 81, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci.* 80, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol.* 62, 109-120, 1984).
20

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.*,
25 *Proc. Natl. Acad. Sci.* 81, 6851-6855, 1984; Neuberger *et al.*, *Nature* 312, 604-608, 1984; Takeda *et al.*, *Nature* 314, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be
30 sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent

antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, one can produce humanized antibodies using recombinant methods, as described in
5 GB2188638B. Antibodies which specifically bind to a ceruloplasmin-like protein polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

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

15 Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev.* 5, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain
20 antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol.* 15, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem.* 269, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using
25 manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology. Verhaar *et al.*, 1995, *Int. J. Cancer* 61, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth.* 165, 81-
30 91.

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

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

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

Antisense Oligonucleotides

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

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an

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

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

Precise complementarity is not required for successful duplex formation between an antisense oligonucleotide and the complementary sequence of a ceruloplasmin-like protein polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a ceruloplasmin-like protein polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent ceruloplasmin-like protein nucleotides, can provide targeting specificity for ceruloplasmin-like protein mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in

length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular ceruloplasmin-like protein polynucleotide sequence.

5

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

15

Ribozymes

20

Ribozymes are RNA molecules with catalytic activity. See, e.g., Cech, *Science* 236, 1532-1539; 1987; Cech, *Ann. Rev. Biochem.* 59, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol.* 2, 605-609; 1992, Couture & Stinchcomb, *Trends Genet.* 12, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (e.g., Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

25

30

The coding sequence of a ceruloplasmin-like protein polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the ceruloplasmin-like protein polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence
5 specific manner have been developed and described in the art (*see* Haseloff *et al. Nature* 334, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).
10

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

25 Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease ceruloplasmin-like protein
30 expression. Alternatively, if it is desired that the cells stably retain the DNA construct, it can be supplied on a plasmid and maintained as a separate element or

integrated into the genome of the cells, as is known in the art. The DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

5

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

10

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human ceruloplasmin-like protein. Such genes may represent genes which are differentially expressed in disorders including, but not limited to, cardiovascular disorders, liver disorders, such as Wilson's disease, and CNS disorders, such as Parkinson's disease. Further, such genes may represent genes which are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human ceruloplasmin-like gene or gene product may itself be tested for differential expression.

15

20

25

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

30

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed.,, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature* 308, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science* 257, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human ceruloplasmin-like protein. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human ceruloplasmin-like protein. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human ceruloplasmin-like gene or gene product are up-regulated or down-regulated.

Screening Methods

The invention provides methods for identifying modulators, *i.e.*, candidate or test compounds which bind to ceruloplasmin-like protein polypeptides or polynucleotides and/or have a stimulatory or inhibitory effect on, for example, expression or activity of the ceruloplasmin-like protein polypeptide or polynucleotide. Increased ceruloplasmin-like activity may be desired, for example, in developmental disorders characterized by inappropriately low levels of ceruloplasmin-like activity or in regeneration.

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

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. Such compounds also may include, but are not limited to, other cellular proteins, peptides such as, for example, soluble peptides, including but not limited to, Ig-tailed fusion peptides, comprising extracellular portions of target gene product transmembrane receptors, and members of random peptide libraries (Lam, *et al.*, *Nature* 354, 82-84, 1991; Houghten *et al.*, *Nature* 354, 84-86, 1991), made of D- and/or L-configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries (Songyang *et al.*, *Cell* 72, 767-78, 1993), antibodies (including, but not limited to, polyclonal, monoclonal,

humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')₂ and Fab expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

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

Methods for the synthesis of molecular libraries are well known in the art (see, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90, 6909, 1993; Erb *et al.* *Proc. Natl. Acad. Sci. U.S.A.* 91, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem.* 37, 2678,
20 1994; Cho *et al.*, *Science* 261, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2061; Gallop *et al.*, *J. Med. Chem.* 37, 1233, 1994). Libraries of compounds can be presented in solution (see, e.g., Houghten, *BioTechniques* 13, 412-421, 1992), or on beads (Lam, *Nature* 354, 82-84, 1991), chips (Fodor, *Nature* 364, 555-556, 1993), bacteria or spores
25 (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89, 1865-1869, 1992), or phage (Scott & Smith, *Science* 249, 386-390, 1990; Devlin, *Science* 249, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci.* 97, 6378-6382, 1990; Felici, *J. Mol. Biol.* 222, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

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

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

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at
25 the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the
30 compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

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

5

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support.

10 When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

15 For binding assays, the test compound is preferably a small molecule which binds to and occupies the active site or the multicopper oxidase domain of the ceruloplasmin-like protein polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules. In binding assays, either the test compound or the ceruloplasmin-like
20 protein polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to the ceruloplasmin-like protein polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining con-
25 version of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a ceruloplasmin-like protein polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a target poly-
30 peptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable

potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a ceruloplasmin-like protein polypeptide. (McConnell *et al.*, *Science* 257, 1906-1912, 1992).

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

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

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct a polynucleotide
25 encoding a ceruloplasmin-like protein polypeptide is fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence that encodes an unidentified protein ("prey" or "sample") is fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in*
30 *vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity

allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the ceruloplasmin-like protein polypeptide.

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

In one embodiment, a ceruloplasmin-like protein polypeptide is a fusion protein comprising a domain that allows the ceruloplasmin-like protein polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed ceruloplasmin-like protein polypeptide; the mixture is then incubated under conditions conducive to complex

formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

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

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the ceruloplasmin-like protein polypeptide (or polynucleotides) or test compound, enzyme-linked assays which rely on detecting a ceruloplasmin-like protein activity of the ceruloplasmin-like protein polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a ceruloplasmin-like protein polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a ceruloplasmin-like protein polynucleotide or polypeptide can be used in a cell-based assay system. A ceruloplasmin-like protein polynucleotide can be naturally

occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, including neoplastic cell lines such as the colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, and the H392 glioblastoma cell line, can be used. An intact cell is contacted with a test compound. Binding of the test compound to a ceruloplasmin-like protein polypeptide or polynucleotide is determined as described above, after lysing the cell to release the ceruloplasmin-like protein polypeptide-test compound complex.

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease a ceruloplasmin-like protein activity of a ceruloplasmin-like protein polypeptide. Ceruloplasmin-like protein activity can be measured, for example, as described in U.S. Patent No. 4,722,894. Ceruloplasmin-like protein activity can be measured after contacting either a purified ceruloplasmin-like protein polypeptide, a cell extract, or an intact cell with a test compound. A test compound which decreases ceruloplasmin-like protein activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing ceruloplasmin activity. A test compound which increases ceruloplasmin-like protein activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing ceruloplasmin activity.

Gene Expression

In another embodiment, test compounds which increase or decrease ceruloplasmin-like protein gene expression are identified. A ceruloplasmin-like protein polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the ceruloplasmin-like protein polynucleotide is determined. The level of

expression of ceruloplasmin-like protein mRNA or polypeptide in the presence of the test compound is compared to the level of expression of ceruloplasmin-like protein mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of ceruloplasmin-like protein mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of ceruloplasmin-like protein mRNA or polypeptide is less expression. Alternatively, when expression of the mRNA or protein is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of ceruloplasmin-like protein mRNA or polypeptide expression.

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

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a ceruloplasmin-like protein polynucleotide can be used in a cell-based assay system. The ceruloplasmin-like protein polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, including neoplastic cell lines such as the colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, and the H392 glioblastoma cell line, can be used.

Pharmaceutical Compositions

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

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

30 Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to

obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

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

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

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes.

Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular
5 barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

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

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

25

Therapeutic Indications and Methods

1. Neurodegenerative diseases. Ceruloplasmin-like protein activity provides a therapeutic target for treating or preventing CNS disorders such as brain injuries,
30 cerebrovascular diseases and their consequences, Parkinson's disease, corticobasal degeneration, motor neuron disease, dementia, including ALS, multiple sclerosis,

traumatic brain injury, stroke, post-stroke, post-traumatic brain injury, and small-vessel cerebrovascular disease. Dementias, such as Alzheimer's disease, vascular dementia, dementia with Lewy bodies, frontotemporal dementia and Parkinsonism linked to chromosome 17, frontotemporal dementias, including Pick's disease, progressive nuclear palsy, corticobasal degeneration, Huntington's disease, thalamic degeneration, Creutzfeld-Jakob dementia, HIV dementia, schizophrenia with dementia, and Korsakoff's psychosis also can be treated. Similarly, it is possible to treat cognitive-related disorders, such as mild cognitive impairment, age-associated memory impairment, age-related cognitive decline, vascular cognitive impairment, attention deficit disorders, attention deficit hyperactivity disorders, and memory disturbances in children with learning disabilities, by regulating the activity of human ceruloplasmin-like protein.

2. Cardiovascular diseases. Cardiovascular diseases include the following disorders of the heart and the vascular system: congestive heart failure, myocardial infarction, ischemic diseases of the heart, all kinds of atrial and ventricular arrhythmias, hypertensive vascular diseases and peripheral vascular diseases.

Heart failure is defined as a pathophysiologic state in which an abnormality of cardiac function is responsible for the failure of the heart to pump blood at a rate commensurate with the requirement of the metabolizing tissue. It includes all forms of pumping failure such as high-output and low-output, acute and chronic, right-sided or left-sided, systolic or diastolic, independent of the underlying cause.

Myocardial infarction (MI) is generally caused by an abrupt decrease in coronary blood flow that follows a thrombotic occlusion of a coronary artery previously narrowed by arteriosclerosis. MI prophylaxis (primary and secondary prevention) is included as well as the acute treatment of MI and the prevention of complications.

Ischemic diseases are conditions in which the coronary flow is restricted resulting in an perfusion which is inadequate to meet the myocardial requirement for oxygen.

This group of diseases include stable angina, unstable angina and asymptomatic ischemia.

5 Arrhythmias include all forms of atrial and ventricular tachyarrhythmias (atrial tachycardia, atrial flutter, atrial fibrillation, atrio-ventricular reentrant tachycardia, preexcitation syndrome, ventricular tachycardia, ventricular flutter, ventricular fibrillation) as well as bradycardic forms of arrhythmias.

10 Hypertensive vascular diseases include primary as well as all kinds of secondary arterial hypertension (renal, endocrine, neurogenic, others). The genes may be used as drug targets for the treatment of hypertension as well as for the prevention of all complications. Peripheral vascular diseases are defined as vascular diseases in which arterial and/or venous flow is reduced resulting in an imbalance between blood supply and tissue oxygen demand. It includes chronic peripheral arterial occlusive
15 disease (PAOD), acute arterial thrombosis and embolism, inflammatory vascular disorders, Raynaud's phenomenon and venous disorders.

Increased ceruloplasmin activity is associated with oxidation of low density lipoproteins, and atherosclerosis. Thus, a therapeutic strategy aimed at decreasing
20 ceruloplasmin activity may be a viable means of treating cardiovascular disorders.

3. Iron homeostasis disorders. Hereditary hemochromatosis is associated with the absence of ceruloplasmin, a disorder whose symptoms include diabetes, retinal degeneration, and neurological symptoms. Long term iron accumulation due to the
25 lack of ceruloplasmin has been identified as a primary cause of these symptoms. Thus, a therapeutic regimen that increases ceruloplasmin activity could be used to treat iron homeostasis disorders.

4. Liver disorders. Wilson's disease is caused by inactive ceruloplasmin, *i.e.*,
30 sharp decrease in active ceruloplasmin. See 36th Japan Human Genetics Society, October, 1991. Wilson's disease is associated with defects in copper metabolism.

These defects lead to a variety of debilitating symptoms including cirrhosis of the liver and neurological disorders. Thus, enhancing ceruloplasmin-like protein activity may provide a method of treating Wilson's disease.

5 The invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a polypeptide-binding partner) can be used in an animal model to determine the efficacy, toxicity, 10 or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

15 A reagent which affects ceruloplasmin-like protein activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce ceruloplasmin-like protein activity. The reagent preferably binds to an expression product of a human ceruloplasmin-like protein gene. If the expression product is a polypeptide, the reagent is preferably an 20 antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

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

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about
5 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably
10 between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to
15 polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a tumor cell, such as a tumor cell ligand exposed on the outer surface of the liposome.

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

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al.* *Trends in Biotechnol.* 11, 202-05 (1993);
30 Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem.* 263, 621-24 (1988);

Wu *et al.*, *J. Biol. Chem.* 269, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem.* 266, 338-42 (1991).

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

Determination of a Therapeutically Effective Dose

15 The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases ceruloplasmin activity relative to that which occurs in the absence of the therapeutically effective dose.

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

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

30

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

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

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

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

- 54 -

about 1 μ g to about 2 mg, about 5 μ g to about 500 μ g, and about 20 μ g to about 100 μ g of DNA.

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

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

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

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

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

EXAMPLE 1

Detection of ceruloplasmin-like protein activity

The polynucleotide of SEQ ID NO: 1 is inserted into the expression vector pCEV4 and the expression vector pCEV4-ceruloplasmin-like protein polypeptide obtained is transfected into human embryonic kidney 293 cells. 1-3 Days post-transfection, cells are harvested and lysed by freeze-thawing. Ceruloplasmin-like protein activity is measured as described in U.S. Patent No. 4,722,894. Briefly, a reagent solution of 11.3 ml 0.2 M acetate buffer, pH 5.0; 38.7 ml 0.2 M sodium acetate solution; 0.1 g Triton-X 100; and 50 ml H₂O is made. 3 ml of the reagent solution is incubated at 37°C for 5 minutes. 30 µl of the cell lysate is added to the reagent solution and incubated at 37 °C for 2 minutes. Absorbance changed at 663 nm are measured for 1 minute. It is shown that the polypeptide of SEQ ID NO: 2 has a ceruloplasmin-like protein activity.

EXAMPLE 2

Identification of a test compound which binds to a ceruloplasmin-like protein polypeptide

Purified ceruloplasmin-like protein polypeptides comprising a glutathione-S-transferase protein are absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Ceruloplasmin-like protein polypeptides comprise the amino acid sequence shown in SEQ ID NO:2 or 11. The test compounds comprise a

fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a ceruloplasmin-like protein polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound was not incubated is identified as a compound which binds to a ceruloplasmin-like protein polypeptide.

EXAMPLE 3

Identification of a test compound which decreases ceruloplasmin-like protein activity

HEK293 cells are seeded at 6×10^5 /well in poly-L-lysine-coated 6-well plates and cultured for 24 hours. Cells are then transfected with vectors containing ceruloplasmin-like constructs and 6 μ l LipofectAMINE Plus reagent. An additional 4 μ l LipofectAMINE Plus reagent is added to each well as well as a suitable amount of a test compound from a small molecule library. Control wells receive no test compound. 1-3 days post-transfection, cells are harvested and lysed by freeze-thawing. Ceruloplasmin-like protein activity is measured as described in U.S. Patent No. 4,722,894. Briefly, a reagent solution of 11.3 ml 0.2 M acetate buffer, pH 5.0; 38.7 ml 0.2 M sodium acetate solution; 0.1 g Triton-X 100; and 50 ml H₂O is made. 3 ml of the reagent solution is incubated at 37 ° C for 5 minutes. 30 μ l of the cell lysate is added to the reagent solution and incubated at 37 ° C for 2 minutes. Absorbance changed at 663 nm are measured for 1 minute. A test compound which decreases ceruloplasmin-like protein activity of the extract relative to the control extract by at least 20% is identified as a ceruloplasmin-like protein inhibitor.

EXAMPLE 4

Identification of a test compound which decreases ceruloplasmin-like protein gene expression

5 A test compound is administered to a primary culture of vascular smooth muscle cells and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells incubated for the same time without the test compound provides a negative control.

10 RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 µg total RNA and hybridized with a ³²P-labeled ceruloplasmin-like protein-specific probe at 65 ° C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO:1 or 10. A test compound which decreases the ceruloplasmin-like protein-specific signal relative to the signal obtained
15 in the absence of the test compound is identified as an inhibitor of ceruloplasmin-like protein gene expression.

EXAMPLE 5

Tissue-specific expression of ceruloplasmin-like protein

20

The qualitative expression pattern of ceruloplasmin-like protein in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR). To demonstrate that ceruloplasmin-like protein is involved in CNS disorders, the following tissues are screened: fetal and adult brain, muscle, heart, lung, kidney,
25 liver, thymus, testis, colon, placenta, trachea, pancreas, kidney, gastric mucosa, colon, liver, cerebellum, skin, cortex (Alzheimer's and normal), hypothalamus, cortex, amygdala, cerebellum, hippocampus, choroid, plexus, thalamus, and spinal cord.

30 *Quantitative expression profiling.* Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in

Higuchi *et al.*, *BioTechnology* 10, 413-17, 1992, and Higuchi *et al.*, *BioTechnology* 11, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

5

If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al.*, *Genome Res.* 6, 986-94, 1996, and Gibson *et al.*, *Genome Res.* 6, 995-1001, 1996).

15

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

20

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

RNA extraction and cDNA preparation. Total RNA from the tissues listed above are used for expression quantification. RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

25

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl

30

- 59 -

RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10 mM MgCl₂; 50 mM NaCl; and 1 mM DTT.

5 After incubation, RNA is extracted once with 1 volume of phenol:chloroform:iso-
amyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume
of 3 M NaAcetate, pH 5.2, and 2 volumes of ethanol.

10 Fifty µg of each RNA from the autaptic tissues are DNase treated with the DNA-free
kit purchased from Ambion (Ambion, TX). After resuspension and spectrophoto-
metric quantification, each sample is reverse transcribed with the TaqMan Reverse
Transcription Reagents (PE Applied Biosystems, CA) according to the manu-
facturer's protocol. The final concentration of RNA in the reaction mix is
200 ng/µL. Reverse transcription is carried out with 2.5µM of random hexamer
primers.

15

TaqMan quantitative analysis. Specific primers and probe are designed according to
the recommendations of PE Applied Biosystems and are listed below:

forward primer: 5'-(gene specific sequence)-3'

20

reverse primer: 5'-(gene specific sequence)-3'

probe: 5'-(FAM) -(gene specific sequence) (TAMRA)-3'

25

where FAM = 6-carboxy-fluorescein

and TAMRA = 6-carboxy-tetramethyl-rhodamine.

The expected length of the PCR product is -(gene specific length)bp.

30

- 60 -

Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

5 Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

10 The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 µl.

15 Each of the following steps are carried out once: pre PCR, 2 minutes at 50° C, and 10 minutes at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

20 The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

EXAMPLE 6

25 *The effect of test compounds which regulate ceruloplasmin-like activity in models of Parkinson's disease*

The effect of test compounds identified in Examples 1-3, above, in the following *in vivo* models of Parkinson's disease is assessed.

6-Hydroxydopamine (6-OH-DA) Lesion

30

Degeneration of the dopaminergic nigrostriatal and striatopallidal pathways is the central pathological event in Parkinson's disease. This disorder has been mimicked experimentally in rats using single/sequential unilateral stereotaxic injections of 6-OH-DA into the medium forebrain bundle (MFB).

5

Male Wistar rats (Harlan Winkelmann, Germany), weighing 200Å250 g at the beginning of the experiment, are used. The rats are maintained in a temperature- and humidity-controlled environment under a 12 h light/dark cycle with free access to food and water when not in experimental sessions. In the following *in vivo* protocols, all efforts are made to minimize animal suffering, to reduce the number of animals used, and to utilize alternatives to *in vivo* techniques.

10

Animals are administered pargyline on the day of surgery (Sigma, St. Louis, MO, USA; 50 mg/kg i.p.) to inhibit metabolism of 6-OHDA by monoamine oxidase and desmethylinipramine HCl (Sigma; 25 mg/kg i.p.), to prevent uptake of 6-OHDA by noradrenergic terminals. Thirty minutes later the rats are anesthetized with sodium pentobarbital (50 mg/kg) and placed in a stereotaxic frame. To make a lesion in the DA nigrostriatal pathway, 4 µl of 0.01% ascorbic acid-saline containing 8 µg of 6-OHDA HBr (Sigma) are injected into the left medial fore-brain bundle at a rate of 1 µl/min (2.4 mm anterior, 1.49 mm lateral, -2.7 mm ventral to Bregma and the skull surface). The needle is left in place an additional 5 min to allow diffusion to occur.

15

20

Stepping Test. Forelimb akinesia is assessed three weeks following lesion placement using a modified stepping test protocol. In brief, the animals are held by the experimenter with one hand fixing the hindlimbs and slightly raising the hind part above the surface. One paw is touching the table, and is then moved slowly sideways (5 s for 1 m), first in the forehand and then in the backhand direction. The number of adjusting steps is counted for both paws in the backhand and forehand direction of movement. The sequence of testing was right paw forehand and backhand adjusting stepping, followed by left paw forehand and backhand directions. The test is repeated three times on three consecutive days, after an initial training period of three days

25

30

prior to the first testing. Forehand adjusted stepping reveals no consistent differences between lesioned and healthy control animals. Analysis is therefore restricted to backhand adjusted stepping.

5 Balance Test. Balance adjustments following postural challenge also are measured during the stepping test sessions. The rats are held in the same position as described in the stepping test and, instead of being moved sideways, tilted by the experimenter towards the side of the paw touching the table. This maneuver results in loss of balance and the ability of the rats to regain balance by forelimb movements was
10 scored on a scale ranging from 0 to 3. Score 0 is given for a normal forelimb placement. When the forelimb movement is delayed but recovery of postural balance detected, score 1 is given. Score 2 represents a clear, yet insufficient, forelimb reaction, as evidenced by muscle contraction, but lack of success in recovering balance, and score 3 is given for no reaction of movement. The test is repeated three
15 times a day on each side for three consecutive days after an initial training period of three days prior to the first testing.

Staircase Test (Paw Reaching). A modified version of the staircase test is used for evaluation of paw reaching behavior three weeks following primary and secondary
20 lesion placement. Plexiglass test boxes with a central platform and a removable staircase on each side are used. The apparatus is designed such that only the paw on the same side at each staircase can be used, thus providing a measure of independent forelimb use. For each test the animals are left in the test boxes for 15 min. The double staircase is filled with 7 x 3 chow pellets (Precision food pellets, formula: P, purified rodent diet, size 45 mg; Sandown Scientific) on each side. After each test,
25 the number of pellets eaten (successfully retrieved pellets) and the number of pellets taken (touched but dropped) for each paw and the success rate (pellets eaten/pellets taken) are counted separately. After three days of food deprivation (12 g per animal per day), the animals are tested for 11 days. Full analysis is conducted only for the
30 last five days.

MPTP treatment

The neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydro-pyridine (MPTP) causes degeneration of mesencephalic dopaminergic (DAergic) neurons in rodents, non-human primates, and humans and, in so doing, reproduces many of the symptoms of Parkinson's disease. MPTP leads to a marked decrease in the levels of dopamine and its metabolites, and in the number of dopaminergic terminals in the striatum as well as severe loss of the tyrosine hydroxylase (TH)-immunoreactive cell bodies in the substantia nigra, pars compacta.

10

To obtain severe and long-lasting lesions, and to reduce mortality, animals receive single injections of MPTP, and are then tested for severity of lesion 7–10 days later. Successive MPTP injections are administered on days 1, 2 and 3. Animals receive application of 4 mg/kg MPTP hydrochloride (Sigma) in saline once daily. All injections are intraperitoneal (i.p.) and the MPTP stock solution is frozen between injections. Animals are decapitated on day 11.

15

Immunohistology. At the completion of behavioral experiments, all animals are anaesthetized with 3 ml thiopental (1 g/40 ml i.p., Tyrol Pharma). The mice are perfused transcardially with 0.01 M PBS (pH 7.4) for 2 min, followed by 4% paraformaldehyde (Merck) in PBS for 15 min. The brains are removed and placed in 4% paraformaldehyde for 24 h at 4°C. For dehydration, they are then transferred to a 20% sucrose (Merck) solution in 0.1 M PBS at 4°C until they sink. The brains are frozen in methylbutan at -20°C for 2 min and stored at -70°C. Using a sledge micro-

20

25

30

A series of sections is processed for free-floating tyrosine hydroxylase (TH) immunohistochemistry. Following three rinses in 0.1 M PBS, endogenous peroxidase

- 64 -

activity is quenched for 10 min in 0.3% H₂O₂ ÅPBS. After rinsing in PBS, sections are preincubated in 10% normal bovine serum (Sigma) for 5 min as blocking agent and transferred to either primary anti-rat TH rabbit antiserum (dilution 1:2000).

5 Following overnight incubation at room temperature, sections for TH immuno-reactivity are rinsed in PBS (2 x 10 min) and incubated in biotinylated anti-rabbit immunoglobulin G raised in goat (dilution 1:200) (Vector) for 90 min, rinsed repeatedly and transferred to Vectastain ABC (Vector) solution for 1 h. 3,3' –
10 Diaminobenzidine tetrahydrochloride (DAB; Sigma) in 0.1 M PBS, supplemented with 0.005% H₂O₂ , serves as chromogen in the subsequent visualization reaction. Sections are mounted onto gelatin-coated slides, left to dry overnight, counter-stained with hematoxylin dehydrated in ascending alcohol concentrations, and cleared in butylacetate. Coverslips were mounted on entellan.

15 Rotarod Test. A modification of the procedure described by Rozas and Labandeira-Garcia (1997), with a CR-1 Rotamex system (Columbus Instruments, Columbus, OH) comprising an IBM-compatible personal computer, a CIO-24 data acquisition card, a control unit, and a four-lane rotarod unit is used. The rotarod unit consists of a
20 rotating spindle (diameter 7.3 cm) and individual compartments for each mouse. The system software allows preprogramming of session protocols with varying rotational speeds (0–80 rpm). Infrared beams are used to detect when a mouse has fallen onto the base grid beneath the rotarod. The system logs the fall as the end of the experiment for that mouse, and the total time on the rotarod, as well as the time of the fall and all the set-up parameters, are recorded. The system also allows a weak current to
25 be passed through the base grid, to aid training.

EXAMPLE 7

The effect of test compounds which regulate ceruloplasmin-like activity in models of dementia

- 5 The effect of test compounds identified in Examples 1-3, above, in the following *in vivo* models of dementia is assessed.

The object recognition task. The object recognition task has been designed to assess the effects of experimental manipulations on the cognitive performance of rodents.

10 A rat is placed in an open field, in which two identical objects are present. The rat inspects both objects during the first trial of the object recognition task. In a second trial, after a retention interval of for example 24 hours, one of the two objects used in the first trial, the 'familiar' object, and a novel object are placed in the open field. The inspection time at each of the objects is registered. The basic measures in the

15 OR task is the time spent by a rat exploring the two object the second trial. Good retention is reflected by higher exploration times towards the novel than the 'familiar' object.

Administration of a test compound prior to the first trial predominantly allows

20 assessment of its the effects on acquisition, and eventually on consolidation processes. Administration of the test compound after the first trial allows to assess the effects on consolidation processes, whereas administration before the second trial allows to measure effects on retrieval processes.

25 ***The passive avoidance task***

The passive avoidance task assesses memory performance in rats and mice. The inhibitory avoidance apparatus consists of a two-compartment box with a light compartment and a dark compartment. The two compartments are separated by a

30 guillotine door that can be operated by the experimenter. A threshold of 2 cm separates the two compartments when the guillotine door is raised. When the door is

- 66 -

open, the illumination in the dark compartment is about 2 lux. The light intensity is about 500 lux at the center of the floor of the light compartment.

Two habituation sessions, one shock session, and a retention session are given, separated by inter-session intervals of 24 hours. In the habituation sessions and the retention session the rat is allowed to explore the apparatus for 300 sec. The rat is placed in the light compartment, facing the wall opposite to the guillotine door. After an accommodation period of 15 sec. the guillotine door is opened so that all parts of the apparatus can be visited freely. Rats normally avoid brightly lit areas and will enter the dark compartment within a few seconds.

In the shock session, the guillotine door between the compartments is lowered as soon as the rat has entered the dark compartment with its four paws, and a scrambled 1 mA footshock is administered for 2 sec. The rat is removed from the apparatus and put back into its home cage. The procedure during the retention session is identical to that of the habituation sessions.

The step-through latency, that is the first latency of entering the dark compartment (in sec.) during the retention session is an index of the memory performance of the animal; the longer the latency to enter the dark compartment, the better the retention is. A test compound is given half an hour before the shock session, together with 1 mg*kg⁻¹ scopolamine. Scopolamine impairs the memory performance during the retention session 24 hours later. If the test compound increases the enter latency compared with the scopolamine-treated controls, is likely to possess cognition enhancing potential.

The Morris water escape task

The Morris water escape task measures spatial orientation learning in rodents. It is a test system that has extensively been used to investigate the effects of putative therapeutic on the cognitive functions of rats and mice. The performance of an

animal is assessed in a circular water tank with an escape platform that is submerged about 1 cm below the surface of the water. The escape platform is not visible for an animal swimming in the water tank. Abundant extra-maze cues are provided by the furniture in the room, including desks, computer equipment, a second water tank, the presence of the experimenter, and by a radio on a shelf that is playing softly.

The animals receive four trials during five daily acquisition sessions. A trial is started by placing an animal into the pool, facing the wall of the tank. Each of four starting positions in the quadrants north, east, south, and west is used once in a series of four trials; their order is randomized. The escape platform is always in the same position. A trial is terminated as soon as the animal had climbs onto the escape platform or when 90 seconds have elapsed, whichever event occurs first. The animal is allowed to stay on the platform for 30 seconds. Then it is taken from the platform and the next trial is started. If an animal does not find the platform within 90 seconds it is put on the platform by the experimenter and is allowed to stay there for 30 seconds. After the fourth trial of the fifth daily session, an additional trial is given as a probe trial: the platform is removed, and the time the animal spends in the four quadrants is measured for 30 or 60 seconds. In the probe trial, all animals start from the same start position, opposite to the quadrant where the escape platform had been positioned during acquisition.

Four different measures are taken to evaluate the performance of an animal during acquisition training: escape latency, traveled distance, distance to platform, and swimming speed. The following measures are evaluated for the probe trial: time (s) in quadrants and traveled distance (cm) in the four quadrants. The probe trial provides additional information about how well an animal learned the position of the escape platform. If an animal spends more time and swims a longer distance in the quadrant where the platform had been positioned during the acquisition sessions than in any other quadrant, one concludes that the platform position has been learned well.

To assess the effects of test compounds, rats or mice with specific brain lesions which impair cognitive functions, or animals treated with compounds such as scopolamine or MK-801, which interfere with normal learning, or aged animals which suffer from cognitive deficits, are used.

5

The T-maze spontaneous alternation task

The T-maze spontaneous alternation task (TeMCAT) assesses the spatial memory performance in mice. The start arm and the two goal arms of the T-maze are provided with guillotine doors which can be operated manually by the experimenter. A mouse is put into the start arm at the beginning of training. The guillotine door is closed. In the first trial, the 'forced trial', either the left or right goal arm is blocked by lowering the guillotine door. After the mouse has been released from the start arm, it will negotiate the maze, eventually enter the open goal arm, and return to the start position, where it will be confined for 5 seconds, by lowering the guillotine door. Then, the animal can choose freely between the left and right goal arm (all guillotine-doors opened) during 14 'free choice' trials. As soon as the mouse has entered one goal arm, the other one is closed. The mouse eventually returns to the start arm and is free to visit whichever goal arm it wants after having been confined to the start arm for 5 seconds. After completion of 14 free choice trials in one session, the animal is removed from the maze. During training, the animal is never handled.

The percent alternations out of 14 trials is calculated. This percentage and the total time needed to complete the first forced trial and the subsequent 14 free choice trials (in s) is analyzed. Cognitive deficits are usually induced by an injection of scopolamine, 30 min before the start of the training session. Scopolamine reduces the percent alternations to chance level, or below. A successful test compound, which is always administered before the training session, will at least partially, antagonize the scopolamine-induced reduction in the spontaneous alternation rate.

REFERENCES

1. Fox *et al.*, *Free Radical Biol Med* 28(12): 1735-44 (2000).
2. Yamaguchi *et al.*, *Pediatr International* 41(4): 405-08 (1999).
- 5 3. Harris *et al.*, *PNAS* 96: 10812-17 (1999).
4. Yuce *et al.*, *Turk. J. Pediatr.* 41(1):99-102 (1999).
5. Pfeil and Lynn, *J. Clin. Gastroenterol.* 29(1): 22-31 (1999).

CLAIMS

1. An isolated polynucleotide encoding a ceruloplasmin-like protein polypeptide and being selected from the group consisting of:
 - 5 a) a polynucleotide encoding a ceruloplasmin-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to
the amino acid sequence shown in SEQ ID NO: 2; and
10 the amino acid sequence shown in SEQ ID NO: 2.
 - b) a polynucleotide comprising the sequence of SEQ ID NO: 1;
 - c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);
 - d) a polynucleotide the sequence of which deviates from the
15 polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and
 - e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).
- 20 2. An expression vector containing any polynucleotide of claim 1.
3. A host cell containing the expression vector of claim 2.
4. A substantially purified ceruloplasmin-like protein polypeptide encoded by a
25 polynucleotide of claim 1.
5. A method for producing a ceruloplasmin-like protein polypeptide, wherein the method comprises the following steps:
 - a) culturing the host cell of claim 3 under conditions suitable for the
30 expression of the ceruloplasmin-like protein polypeptide; and

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

- contacting a test compound with a ceruloplasmin-like protein polypeptide encoded by any polynucleotide of claim 1; and detecting a ceruloplasmin-like protein activity of the polypeptide, wherein a test compound which increases the ceruloplasmin-like protein activity is identified as a potential therapeutic agent for increasing the activity of the ceruloplasmin-like protein, and wherein a test compound which decreases the ceruloplasmin-like protein activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the ceruloplasmin-like protein.
- 5
- 10 12. A method of screening for agents which decrease the activity of a ceruloplasmin-like protein, comprising the steps of:
contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of ceruloplasmin-like protein.
- 15
13. A method of reducing the activity of ceruloplasmin-like protein, comprising the steps of:
contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any ceruloplasmin-like protein polypeptide of claim 4, whereby the activity of ceruloplasmin-like protein is reduced.
- 20
14. A reagent that modulates the activity of a ceruloplasmin-like protein polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.
- 25
15. A pharmaceutical composition, comprising:
the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.
- 30

- 73 -

16. Use of the expression vector of claim 2 or the reagent of claim 14 for the preparation of a medicament for modulating the activity of a ceruloplasmin-like protein in a disease.
- 5 17. Use of claim 16 wherein the disease is Wilson's disease or a CNS disorder.
18. A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
- 10 19. The cDNA of claim 18 which comprises SEQ ID NO:1.
20. The cDNA of claim 18 which consists of SEQ ID NO:1.
- 15 21. An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
22. The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO:1.
- 20 23. A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
24. The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO:1.
- 25 25. A purified polypeptide comprising the amino acid sequence shown in SEQ ID NO:2.
26. The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NO:2.
- 30

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

34. The method of claim 33 wherein the reagent is an antibody.
35. A kit for detecting a polypeptide comprising the amino acid sequence shown
5 in SEQ ID NO:2, comprising:
an antibody which specifically binds to the polypeptide; and instructions for
the method of claim 33.
36. A method of screening for agents which can modulate the activity of a human
10 ceruloplasmin-like protein, comprising the steps of:
contacting a test compound with a polypeptide comprising an amino acid
sequence selected from the group consisting of: (1) amino acid sequences
which are at least about 50% identical to the amino acid sequence shown in
SEQ ID NO:2 and (2) the amino acid sequence shown in SEQ ID NO:2; and
15 detecting binding of the test compound to the polypeptide, wherein a test
compound which binds to the polypeptide is identified as a potential agent for
regulating activity of the human ceruloplasmin-like protein.
37. The method of claim 36 wherein the step of contacting is in a cell.
20
38. The method of claim 36 wherein the cell is *in vitro*.
39. The method of claim 36 wherein the step of contacting is in a cell-free
system.
25
40. The method of claim 36 wherein the polypeptide comprises a detectable label.
41. The method of claim 36 wherein the test compound comprises a detectable
label.
30

42. The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.
43. The method of claim 36 wherein the polypeptide is bound to a solid support.
- 5 44. The method of claim 36 wherein the test compound is bound to a solid support.
- 10 45. A method of screening for agents which modulate an activity of a human ceruloplasmin-like protein, comprising the steps of:
contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2 and (2) the amino acid sequence shown in SEQ ID NO:2; and
15 detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human ceruloplasmin-like protein, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential agent for decreasing the activity of the human ceruloplasmin-like
20 protein.
46. The method of claim 45 wherein the step of contacting is in a cell.
47. The method of claim 45 wherein the cell is *in vitro*.
- 25 48. The method of claim 45 wherein the step of contacting is in a cell-free system.
49. A method of screening for agents which modulate an activity of a human ceruloplasmin-like protein, comprising the steps of:
- 30

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

50. The method of claim 49 wherein the product is a polypeptide.
51. The method of claim 49 wherein the product is RNA.
- 10 52. A method of reducing activity of a human ceruloplasmin-like protein,
 comprising the step of:
 contacting a cell with a reagent which specifically binds to a product encoded
 by a polynucleotide comprising the nucleotide sequence shown in SEQ ID
15 NO:1, whereby the activity of a human ceruloplasmin-like protein is reduced.
53. The method of claim 52 wherein the product is a polypeptide.
54. The method of claim 53 wherein the reagent is an antibody.
- 20 55. The method of claim 52 wherein the product is RNA.
56. The method of claim 55 wherein the reagent is an antisense oligonucleotide.
- 25 57. The method of claim 56 wherein the reagent is a ribozyme.
58. The method of claim 52 wherein the cell is *in vitro*.
59. The method of claim 52 wherein the cell is *in vivo*.
- 30 60. A pharmaceutical composition, comprising:

- 78 -

a reagent which specifically binds to a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2; and a pharmaceutically acceptable carrier.

5 61. The pharmaceutical composition of claim 60 wherein the reagent is an antibody.

10 62. A pharmaceutical composition, comprising:
a reagent which specifically binds to a product of a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1; and a pharmaceutically acceptable carrier.

63. The pharmaceutical composition of claim 62 wherein the reagent is a ribozyme.

15 64. The pharmaceutical composition of claim 62 wherein the reagent is an antisense oligonucleotide.

20 65. The pharmaceutical composition of claim 62 wherein the reagent is an antibody.

66. A pharmaceutical composition, comprising:
an expression vector encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2; and a pharmaceutically acceptable carrier.

25 67. The pharmaceutical composition of claim 66 wherein the expression vector comprises SEQ ID NO:1.

30 68. A method of treating a ceruloplasmin-like protein dysfunction related disease, wherein the disease is selected from Wilson's disease or a CNS disorder comprising the step of:

- 79 -

administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human ceruloplasmin-like protein, whereby symptoms of the ceruloplasmin-like protein dysfunction related disease are ameliorated.

5

69. The method of claim 68 wherein the reagent is identified by the method of claim 36.

10

70. The method of claim 68 wherein the reagent is identified by the method of claim 45.

71. The method of claim 68 wherein the reagent is identified by the method of claim 49.

Fig. 1

atgcatgcaatcaatggcctttgtttttgggaatttacctgagctgaacatgtg
tgcacagaaacgtgtggcctggcacttggttggcatgggcaatgaaattgatg
tccacacagcatttttccatggacagatgctgactacccgtggacaccacact
gatgtggctaacaatcctttccagccacctttgtgactgctgagatgggtgccctg
ggaacctggtacctgggttaattagctgccaaagtgaacagtcactttcgagatg
gcatgcaggcactctacaaggtcaagtcttgctccatggccccctcctgtggac
ctgctcacaggcaaaagttcgacagtaacttcattgaggcccatgagattcaatg
ggactatggcccgatggggcatgatgggagtaactgggaagaatttgagagagc
caggcagtatctcagataagtttttccagaagagctccagccgaattgggggc
acttactggaaagtgcgatatgaagcctttcaagatgagacattccaagagaa
gatgcattttggaggaagataggcatcttggaatcctggggccagtgatccggg
ctgaggtgggtgacaccattcaggtgggtcttctacaaccgtgcctcccagcca
ttcagcatgcagccccatggggctcttttatgagaaagactatgaaggcactgt
gtacaatgatgggtcatcttaccctggcttggttgccaagccctttgagaaag
taacataccgctggacagtcccccctcatgccggtcccactgctcaggatcct
gcttgtctcacttggtatgtacttctctgctgcagatcccataagagacacaaa
ttctggcctggtgggccccgctgctgggtgtgcagggtggtgccttgggtgcag
atggcaagcagaaaggggtggataaagaattctttcttctcttactgtgttg
gatgagaacaagagctggtacagcaatgccaatcaagcagctgctatgttgga
tttccgactgctttcagaggatattgagggcttccaagactccaatcggatgc
atgccattaatgggtttctgttctctaacctgccaggctggacatgtgcaag
ggtgacacagtggcctggcacctgctcggcctgggcacagagactgatgtgca
tggagtcatgttccagggaacactgtgcagcttcagggcacatgaggaaggggtg
cagctatgctctttcctcatacctttgtcatggccatcatgcagcctgacaac
cttgggacatttgagatttattgccaggcaggcagccatcgagaagcagggat
gagggcaatctataatgtctcccagtgtcctggccaccaagccacccctcgcc
aacgctaccaagctgcaagaatctactatatcatggcagaagaagtagagtgg
gactattgccctgaccggagctgggaacgggaatggcacaaccagtcctgagaa
ggacagttatgggttacattttcctgagcaacaaggatgggctcctgggttcca
gatacaagaaagctgtattcagggaatacactgatggtacattcaggatccct
cggccaaggactggaccagaagaacacttgggaatcttgggtccacttatcaa
aggtgaagttggtgatatacctgactgtggtattcaagaataatgccagccgcc
cctactctgtgcatgctcatggagtgttagaatctactactgtctggccactg
gctgctgagcctggtgaggtggtcacttatcagtggaaacatcccagagaggtc
tggccctgggcccacatgactctgcttgtgtttcctggatctattattctgcag
tggatcccatcaaggacatgtatagtggcctggtggggcccttggctatctgc
caaaagggcatcctggagccccatggaggacggagtgacatggatcggggaatt
tgcattgttgttcttgatttttgatgaaaataagtcttggtatttgaggaaa
atgtggcaacccatgggtcccaggatccaggcagtattaacctacaggatgaa

Fig. 1 (Continuation)

actttcttggagagcaataaaaatgcatgcaatcaatgggaaactctatgccaa
ccttaggggtcttaccatgtaccaaggagaacgagtggcctggtacatgctgg
ccatgggccaagatgtggatctacacaccatccactttcatgcagagagcttc
ctctatcggaatggcgagaactaccgggcagatgtggtggatctgttcccagg
gacttttgagggttgtggagatggtggccagcaaccctgggacatggctgatgc
actgccatgtgactgaccatgtccatgctggcatggagaccctcttcactgtt
ttttctcgaacagaacacttaagccctctcaccgtcatcaccaaagagactga
aaaagcagtgccccccagagacattgaagaaggcaatgtgaagatgctgggca
tgcagatccccataaagaatgttgagatgctggcctctgttttggttgccatt
agtgtcacccttctgctcgttggttctggctcttggtggagtggtttggtacca
acatcgacagagaaagctacgacgcaataggaggtccatcctggatgacagct
tcaagcttctgtctttcaaacagtaa

Fig. 2

MHAINGFVFGNLPELNMCAQKRVAWHLFGMGNEIDVHTAFFHGQMLTTRGHHT
DVANIFPATFVTAEMVPWEPGTWLI SCQVNSHFRDGMQALYKVKSCSMAPPVD
LLTGKVRQYFIEAHEIQWDYGPMGHDGSTGKNLREPGSISDKFFQKSSSRIGG
TYWKVRYEAFQDETFQEKMHLEEDRHLGILGPVIRAEVGD TIQVVFYNRASQP
FSMQPHGVFYEKDYEGTVYNDGSSYPGLVAKPF EKVTYRWTVP PHAGPTAQDP
ACLTWMYFSAADPIRDTNSGLVGPLLVC RAGALGADGKQKGVDKEFFLLFTVL
DENKSWYSNANQAAAMLD FRLLEDIEGFQDSNRMHAINGFLFSNLPR LDMCK
GDTVAVHLLGLGTETDVHGVMFQGN TVQLQGM RKGAAMLFPHTFVM AIMQPDN
LGTFEIYCQAGSHREAGMRAIYNVSQCPGHQATPRQRYQAARIYYIMAE EEW
DYCPDRSWEREWHNQSEKDSYGYIFLSNKDGLLGSRYKKAVFREYTDGTFRIP
RPRTGP EEHLGILGPLIKGEVGDILT VVFKNNASRPYSVHAHGVLESTTVWPL
AAEPGEVV TYQWNI PERSGPGPNDSACVSWIYYSAVDPIKDMYSGLVGPLAIC
QKGILEPHGGRSDMDREFALLFLIFDENKSWYLEENVATHGSQDPGSINLQDE
TFLESNKMHAINGKLYANLRGLTMYQGERVAWYMLAMGQDVDLHTIHFHAESF
LYRNGENYRADVVDLFPGT FEVVEMVASNP GTWLMHCHVTDH VHAGMETLFTV
FSRTEHLSPLTVITKETEKAVPPRDIEEGNVKMLGMQIPIKNVEMLASVLVAI
SVTLLLVLALGGVVWYQHRQRKLRRNRRSILDDSFKLLSFKQ

Fig. 3

ggaaaagaggggcacccagcccttccccctccctcatcctcccatcccagtaaa
ccctgccaaattggaatcctggacttaatttaggagaaaggccctgtaaccaa
gatactgactgaacatggctggcggactcaggctggggctctgcagtgacagcat
taatggggccgctgacatgaatatggagtagttttctctagcaaagagtggctt
ccagcttcttaaagtctgacaagaaccggataggggggaacctacaagaagacc
atctataaagaatacaaggatgactcatacacagatgaagtggcccagcctgc
ctgggttgggcttccctggggccagtggtgacaggctgaagtgggggatgtcattc
ttattcacctgaagaattttgccactcgtccctataccatccaccctcatggt
gtcttctacgagaaggactctgaagggtccctatacccagatggctcctctgg
gccactgaaagctgatgactctgttccccggggggcagccatatctacaact
ggaccattccagaaggccatgcacccaccgatgctgacccagcgtgcctcacc
tggtatctaccattctcatgtagatgctccacgagacattgcaactggccta
tgggcctctcatcacctgtaaaagaggagccctggatgggaactcccctcctc
aacgccaggatgtagaccatgatttcttccctcctcttcagtgtggtagatgag
aacctcagctggcatctcaatgagaacattgccacttactgctcagatcctgc
ttcagtggacaaagaagatgagacatttcaggagagcaataggatgcatgcaa
tcaatggctttgtttttgggaatttacctgagctgaacatgtgtgcacagaaa
cgtgtggcctggcacttggttggcatgggcaatgaaattgatgtccacacagc
atttttccatggacagatgctgactaccgctggacaccacactgatgtggcta
acatctttccagccacctttgtgactgctgagatgggtgccctgggaacctggt
acctggttaattagctgccaaagtgaacagtcactttcgagatggcatgcaggc
actctacaagggtcaagtcttgctccatggccccctcctgtggacctgctcacag
gcaaagttcgacagtacttcattgaggcccatgagattcaatgggactatggc
ccgatggggcatgatgggagtagtgggaagaatttgagagagccaggcagtat
ctcagataagtttttccagaagagctccagccgaattgggggacacttactgga
aagtgcgatatgaagcctttcaagatgagacattccaagagaagatgcatttg
gaggaagataggcatcttggaatcctggggccagtgatccgggctgaggtggg
tgacaccattcaggtggtcttctacaaccgtgcctcccagccattcagcatgc
agccccatgggggtcttttatgagaaagactatgaaggcactgtgtacaatgat
ggctcatcttaccctggcttggttgccaagccctttgagaaagtaacataccg
ctggacagtccccctcatgccgggtcccactgctcaggatcctgcttgtctca
cttggatgtacttctctgctgcagatcccataagagacacaaattctggcctg
gtggggcccgctgctggtgtgcagggtggtgccttgggtgcagatggcaagca
gaaaggggtggataaagaattctttcttctcttccactgtgttggtgagaaca
agagctggtacagcaatgccaatcaagcagctgctatgttggatttccgactg
ctttcagaggatattgagggcttccaagactccaatcgatgcatgccattaa
tggtttctgttcttaacctgccaggtggacatgtgcaaggggtgacacag
tggtctggcacctgctcggcctgggcacagagactgatgtgcatggagtcag
ttccagggcaacactgtgcagcttcagggcatgaggaaggggtgcagctatgct
ctttcctcatacctttgtcatggccatcatgcagcctgacaaccttgggacat

Fig. 3 (Continuation)

ttgagatTTtattgccaggcaggcagccatcgagaagcagggatgagggcaatc
tataatgtctcccagtgctcctggccaccaagccacccctcgccaacgctacca
agctgcaagaatctactatatcatggcagaagaagtagagtgggactattgcc
ctgaccggagctgggaacgggaatggcacaaccagtcctgagaaggacagttat
ggttacatTTTtctgagcaacaaggatgggctcctgggttccagatacaagaa
agctgtattcagggaatacactgatggtacattcaggatccctcggccaagga
ctggaccagaagaacacttgggaatcctgggtccacttatcaaaggtgaagtt
ggtgatatcctgactgtggtattcaagaataatgccagccgcccctactctgt
gcatgctcatggagtgctagaatctactactgtctggccactggctgctgagc
ctgggtgaggtggtcacttatcagtggaacatcccagagaggtctggccctggg
cccaatgactctgcttgtgtttcctggatctattattctgcagtggatcccat
caaggacatgtatagtggcctggtggggcccttggctatctgccaaaagggca
tcttgaggcccccatggaggacggagtgacatggatcggaatttgcattgttg
ttcttgattTTTtgatgaaaataagtccttggtatttggaggaaaatgtggca
acccatgggtcccaggatccaggcagtatTAacctacaggatgaaactttctt
ggagagcaataaaaatgcatgcaatcaatgggaaactctatgccaaccttaggg
gtcttaccatgtaccaaggagaacgagtggcctggtacatgctggccatgggc
caagatgtggatctacacaccatccactttcatgcagagagcttctctatcg
gaatggcgagaactaccgggcagatgtggtggatctgttcccagggaacttttg
aggttgtggagatggtggccagcaaccctgggacatggctgatgcactgccat
gtgactgaccatgtccatgctggcatggagaccctcttcactgtTTTTtctcg
aacagaacacttaagccctctcacgcgtcatcacaaagagactgaaaaagcag
tgccccccagagacattgaagaaggcaatgtgaagatgctgggcatgcagatc
cccataaagaatgTTtgagatgctggcctctgttttgggttgccattagtgtcac
ccttctgctcgttgttctggctccttggtggagtggtttggtagccaacatcgac
agagaaagctacgacgcaataggaggtccatcctggatgacagcttcaagctt
ctgtctttcaaacagtaacatctggagcctggagatatcctcaggaagcacat
ctgtagtgcactcccagcaggccatggactagtcactaaccaccacactcaaag
gggcatgggtggtggagaagcagaaggagcaatcaagcttatctggatatTTc
tttctttattttattttacatggaaataatatgatttcactTTTTtcttagttt
ctttgctctacgtgggcacctggcactaaggagtagcttattatcctacatc
gcaaatttcaacagctacattatatTTTccttctgacacttggaaggtattgaa
atttctagaaatgtatccttctcacaaagtagagaccaagagaaaaactcatt
gattgggttttctacttctttcaaggactcaggaaatttcactttgaactgagg
ccaagtgagctgttaagataaccacacttaaaactaaaggctaagaatatagg
cttgatgggaaattgaaggtaggctgagtatTTgggaatccaaattgaattttg
attctccttggcagtgaaactactTTTgaagaagtggccaatgggttgttgctgc
catgagcatgtacaacctctggagctagaagctcctcaggaaagccagttctc
caagttcttaacctgtggcactgaaaggaatgTTgagttacctcttcatgttt
tagacagcaaaccctatccattaaagtacttgTTtagaacact

Fig. 4

gaacttttcttggagagtaatatataatcatgcaatcaatgggaaactctatgcc
accttaggggtcttaccatgtaccaaggagaacgagtggcctggtacatgctg
gccatgggccaagatgtggatctacacaccatccactttcatgcagagagctt
cctctatcggaatggcgagaactaccgggcagatgtggtggatctgttcccag
ggacttttgagggttgtggagatgggtggccagcaaccctgggacatggctgatg
cactgccatgtgactgaccatgtccatgctggcatggagaccctcttcactgt
tttttctcgaacagaaacacttaagccctctcaccgatcatcaccaaagagactg
aaaaagcagtgccccccagagacattgaagaaggcaatgtgaagatgctgggc
atgcagatccccataaagaatgttgagatgctggcctctgttttggttgccat
tagtgtcacccttctgctcgttggttctggctcttggtggagtgggttggtacc
aacatcgacagagaaagctacgacgcaataggaggtc

- 7/28 -

Fig. 5

gaacagaacacttaagccctctcacccgtcatcaccaaagagactgaaaaagtg
ccccccagagacattgaagaaggcaatgtgaagatgctgggcatgcagatccc
cataaagaatggtgagatgctggcctctgttttggttgccattagtgtcaccc
ttctgctcgttggttctggctcttggtggagtggtttggtaccaacatcgacag
agaaagctacgacgcaataggagggtccatcctggatgacagcttcaagcttct
gtctttcaacagtacatctggagcctggagatatcctcaggaagcacatctgt
agtgcactcccagcaggccatggactagtactaacccccacactc

Fig. 6

gcgagcatgtaccaggccactcgtgctccttggtacatagtaagaccctaag
gttggcatagagtttcccattgattgcatgcattttattgctctccaagaaag
tttcatcctgtaggttaatactgcctggatcctgggacccatggggttgccaca
ttttcctccaaataccaagacttattttcatcaaaaatcaagaacaacaatgc
aaattcccgatccatgtcactccgtcctccatgggcat

- 9/28 -

Fig. 7

tcgaacagaaacttaagccctctgaccgtcatcaccaaagagactgaaaaag
cagtggcccccagagacattgaagaaggcaatgtgaagatgctgggcatgcag
atccccataaagaatggttgagatgctggcctctgttttggttgccattagtgt
cacccttctgctcgttggttctggctcttggtggagttggttggtaccaacatc
gacagagaaagctacgacgcaataggag

Fig. 8

atgaagatTTTtgataCTTggtatTTTTtctgtTTTTatgtagtaccCCcagcctg
ggcgaaagaaaagcattattacattggaattattgaaacgacttgggattatg
cctctgaccatgggggaaaagaaacttatttctgttgacacggaacattccaat
atctatcttcaaaatggcccagatagaattgggagactatataagaaggccct
ttatcttcagtacacagatgaaacctttaggacaactatagaaaaaccggtct
ggcttgggtTTTTtaggccctattatcaaagctgaaactggagataaagtTTat
gtacacttaaaaaaccttgcctctaggccctacacctttcattcacatggaat
aacttactataaggaacatgagggggccatctaccctgataacaccacagatt
ttcaaagagcagatgacaaagtatatccaggagagcagtatacatatgTTg
cttgccactgaagaacaaagtcttggggaaggagatggcaattgtgtgactag
gatttaccattcccacattgatgctccaaaagatatTgcctcaggactcatcg
gacctTTaataatctgtaaaaaagattctctagataaagaaaaagaaaaacat
attgaccgagaatttTgtggtgatgTTTTtctgtggtggatgaaaatttcagctg
gtacctagaagacaacattaaaacctactgctcagaaccagagaaagtTgaca
aagacaacgaagacttccaggagagtaacagaatgtattctgtgaatggatac
actTTTtggagctctcccaggactctccatgtgtgctgaagacagagtaaaatg
gtacctTTTTtgggtatgggtaatgaagttgatgtgcacgcagctttctttcacg
ggcaagcactgactaacaagaactaccgtattgacacaatcaacctctttcct
gctaccctgTTTtgatgcttatatggtggcccagaaccctggagaatggatgct
cagctgtcagaatctaaacctctgaaagccggtttgcaagcctTTTTtccagg
tccaggagtgtacaagtcttcatcaaaggataatatccgtgggaagcatgtt
agacactactacattgccgctgaggaaatcatctggaactatgctccctctgg
tatagacatcttcaactaaagaaaacttaacagcacctggaagtgactcagcgg
tgTTTTTTgaacaaggtaccacaagaattggaggctcttataaaaaagctgggt
tatcgtgagtacacagatgcctccttcacaaatcgaaaggagagaggccctga
agaagagcatcttggcatcctgggtcctgtcatttgggcagaggtgggagaca
ccatcagagtaaccttccataacaaaggagcatatcccctcagtattgagccg
attgggggtgagattcaataagaacaacgaggggcacatactattccccc aaatta
caacccccagagcagaagtgtgcctccttcagcctcccatgtggcacccacag
aaacattcacctatgaatggactgtccccc aaagaagtaggacccactaatgca
gatcctgtgtgtctagctaagatgtattattctgtgtggatcccactaaaga
tatattcactgggcttattgggccaatgaaaatatgcaagaaaggaagtTtac
atgcaaattgggagacagaaagatgtagacaaggaattctatttgtttcctaca
gtatttTgatgagaatgagagtttactcctggaagataatattagaatgtttac
aactgcacctgatcaggtggataaggaagatgaagactttcaggaatctaata
aaatgcactccatgaatggattcatgtatgggaatcagccgggtctcactatg
tgcaaaggagattcggtcgtgtggtacttattcagcgcgggaaatgagggccga
tgtacatggaatatactTTTTcaggaaacacatatctgtggagaggagaacgga
gagacacagcaaacctcttccctcaaacaagtcttacgctccacatgtggcct

Fig. 8 (Continuation)

gacacagaggggactttttaatggtgaatgccttacaactgatcattacacagg
cggcatgaagcaaaaatatactgtgaaccaatgcaggcggcagtcctgaggatt
ccaccttctacctgggagagaggacatactatatcgcagcagtgagggtggaa
tgggattattccccacaaagggagtgaggaaaaggagctgcatcatttacaaga
gcagaatgtttcaaatgcatttttagataagggagagttttacatagggtcaa
agtacaagaaaagttgtgtatcggcagttatactgatagcacattccgtgttcca
gtggagagaaaagctgaagaagaacatctgggaattctagggtccacaacttca
tgcagatgttggagacaaagtcaaaattatctttaaaaacatggccacaaggc
cctactcaatacatgcccattggggtacaaacagagagttctacagttactcca
acattaccaggtgaaactctcacttacgtatggaaaatcccagaaagatctgg
agctggaacagaggattctgcttgtattccatgggcttattattcaactgtgg
atcaagttaaggacctctacagtggattaattggccccctgattgtttgtcga
agaccttacttgaaagtattcaatcccagaaggaagctggaatttgcccttct
gtttctagtttttgatgagaatgaatcttggtacttagatgacaacatcaaaa
catactctgatcaccccgagaaagtaaacaaagatgatgaggaattcatag
aaagcaataaaaatgcatgctattaatggaagaatgtttggaaacctacaaggc
ctcacaatgcacgtgggagatgaagtcaactggtatctgatgggaatgggcaa
tgaaatagacttacacactgtacattttcacggccatagcttccaatacaagc
acaggggagtttatagttctgatgtctttgacattttccctggaacataccaa
accctagaaatgtttccaagaacacctggaatttggttactccactgccatgt
gaccgaccacattcatgctggaatggaaaccacttacaccgttctacaaaatg
aagacaccaaattctggctgaatgaaataaattggtgataagtggaaaaaagag
aaaaaccaatgattcataacaatgtatgtgaaagtgtaaaatagaatgttact
ttggaatgactataaacattaaaagagactggagcat

Fig. 9

MKILILGIFLFLCSTPAWAKEKHYYIGIIETTWDYASDHGEKKLISVDTEHSN
IYLQNGPDRIGRLYKKALYLQYTDETFRTTIEKPVWLGFLGPIIKAETGDKVY
VHLKNLASRPYTFHSHGITYYKEHEGAIYPDNTTDFQRADDKVYPGEQYTYML
LATEEQSPGEGDGNCVTRIYHSHIDAPKDIA SGLIGPLIICKKDSLDKEKEKH
IDREFVVMF SVVDENFSWYLEDNIKTYCSEPEKVDKDNEDFQESNRMYSVNGY
TFGSLPGLSMCAEDRVKWYLF GMGNEVDVHAAFFH GQALTNKNYRIDTINLFP
ATLFDAYMVAQNPGEWMLSCQNLNHLKAGLQAFFQVQECNKSSSKDNIRGKHV
RHYIIAAEEI IWNYP SGIDIFTKENLTAPGSDSAVFFEQGTTRIGGSYKKLV
YREYTDASFTNRKERGP EEEHLGILGPVIWAEVGD TIRVTFHNKGAYPLSIEP
IGVRFNKNNEGTYYS PNYNPQSR SVPPSASHVAPTETFTYEWTPKEVGPTNA
DPVCLAKMYYS AVDP TKDIFTGLIGPMKICKKGSLHANGRQKDVDKEFYLFPT
VFDENESLLEDNIRMFTTAPDQVDKEDEDFQESNKMHS MN GFM YGNQPGLTM
CKGDSVVWYLF SAGNEADVHGIYFSGNTYLWRGERRDTANLFPQTS LTLHMWP
DTEGTFNVECLTTDHYTGGMKQKYTVNQCRQSEDSTFYLGERTYYIAAVEVE
WDYSPQREWEKELHHLQE QNVSN AFLDKGEFYIGSKYKKVVYRQYTDSTFRVP
VERKAE EEEHLGILGPQLHADVGDKVKIIFKNMATRPYSIHAHGVQTESSTVTP
TLPGETLTYVWKI PERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCR
RPYLKVFNPRRKLEFALLFLVFDENESWYLD DNIKTYSDHPEKVNKDDEEFIE
SNKMHA INGRMFGNLQGLTMHVGDEVNWYLMGMGNEIDLHTVHFHGH SFQYKH
RGVYSSDVFDIFPGTYQTLEMFPRTPGIWLLHCHVTDHIHAGMETTYTVLQNE
DTKSG

Fig. 10

atgaagatTTTTgatacttggtatTTTTctgtTTTTatgtagtaccCCagcctg
ggcgaaagaaaagcattattacattggaattattgaaacgacttgggattatg
cctctgaccatggggaaaagaaacttatttctgttgacacggaacattccaat
atctatcttcaaaatggcccagatagaattgggagactatataagaaggccct
ttatcttcagtacacagatgaaacctttaggacaactatagaaaaaccggtct
ggcttgggtTTTTtaggccctattatcaaagctgaaactggagataaagtttat
gtacacttaaaaaaccttgccctctaggccctacacctttcattcacatggaat
aacttactataaggaacatgagggggccatctaccctgataacaccacagatt
ttcaaagagcagatgacaaagtatatccaggagagcagtatacatatgttg
cttgccactgaagaacaaagtcttggggaaggagatggcaattgtgtgactag
gatttaccattcccacattgatgctccaaaagatattgcctcaggactcatcg
gacctttaataatctgtaaaaaagatttctctagataaagaaaaagaaaaacat
attgaccgagaatttTgtggtgatgttttctgtggtggatgaaaatttcagctg
gtacctagaagacaacattaaaacctactgctcagaaccagagaaagttgaca
aagacaacgaagacttccaggagagtaacagaatgtattctgtgaatggatac
acttttTggaagtctcccaggactctccatgtgtgctgaagacagagtaaaatg
gtacctTTTTTggtatgggtaatgaagttgatgtgcacgcagctttctttcacg
ggcaagcactgactaacaagaactaccgtattgacacaatcaacctctttcct
gctaccctgtttgatgcttatatgggtggcccagaaccctggagaatggatgct
cagctgtcagaatctaaacctctgaaagccggtttgcaagcctttttccagg
tccaggagtgtacaagtcttcatcaaaggataatatccgtgggaagcatgtt
agacactactacattgccgctgaggaaatcatctggaactatgctccctctgg
tatagacatcttTactaaagaaaacttaacagcacctggaagtgactcagcgg
tgtTTTTTgaacaaggtaccacaagaattggagggtctttataaaaagctgggt
tatcgtgagtacacagatgcctccttcacaaatcgaaaggagagaggccctga
agaagagcatcttggcatcctgggtcctgtcatttgggcagaggtgggagaca
ccatcagagtaaccttccataacaaaggagcatatcccctcagtattgagccg
attgggggtgagattcaataagaacaacgaggggcacatactattccccaaatta
caacccccagagcagaagtgtgcctccttcagcctcccatgtggcacccacag
aaacattcacctatgaatggactgtccccaagaagtaggaccactaatgca
gatcctgtgtgtctagctaagatgtattattctgtgtggatcccactaaaga
tatattcactgggcttattgggccaatgaaaatattgcaagaaaggaagtttac
atgcaaatgggagacagaaagatgtagacaaggaattctatttgtttcctaca
gtatttTgatgagaatgagagtttactcctggaagataatattagaatgtttac
aactgcacctgatcaggtggataaggaagatgaagactttcaggaatctaata
aaatgcactccatgaatggattcatgtatgggaatcagccgggtctcactatg
tgcaaaggagattcggctcgtgtggtacttattcagcgccggaaatgagggcga
tgtacatggaatatacttttcaggaaacacatatctgtggagaggagaacgga
gagacacagcaaacctcttccctcaaacaagtcttacgctccacatgtggcct

Fig. 10 (Continuation)

gacacagaggggactttttaatggtgaatgccttacaactgatcattacacagg
cggcatgaagcaaaaatatactgtgaaccaatgcaggcggcagtcctgaggatt
ccaccttctacctgggagagaggacatactatatcgcagcagtgagggtggaa
tgggattattccccacaaagggagtgaggaaaaggagctgcatcatttacaaga
gcagaatgtttcaaagtcatttttagataagggagagttttacatagggtcaa
agtacaagaaagttgtgtatcggcagtatactgatagcacattccgtgttcca
gtggagagaaaagctgaagaagaacatctgggaattctagggtccacaacttca
tgcagatgttggagacaaagtcaaaattatctttaaaaacatggccacaaggc
cctactcaatacatgcccattggggtacaaacagagagttctacagttactcca
acattaccaggtgaaactctcacttacgtatggaaaatcccagaaagatctgg
agctggaacagaggattctgcttgtattccatgggcttattattcaactgtgg
atcaagttaaggacctctacagtggattaattggccccctgattgtttgtcga
agaccttacttgaaagtattcaatcccagaaggaagctggaatttgcccttct
gtttctagtttttgatgagaatgaatcttgggtacttagatgacaacatcaaaa
catactctgatcaccccgagaaagtaaacaaagatgatgaggaattcatag
aaagcaataaaatgcatgctattaatggaagaatgtttggaaacctacaaggc
ctcacaatgcacgtgggagatgaagtcaactggtatctgatgggaatgggcaa
tgaaatagacttacacactgtacattttcacggccatagcttccaatacaagc
acaggggagtttatagttctgatgtctttgacattttccctggaacataccaa
accctagaaatgtttccaagaacacctggaatttggttactccactgccatgt
gaccgaccacattcatgctggaatggaaaccacttacaccgttctacaaaatg
aagacaccaaactctggctgaatgaaataaattggtgataagtggaaaaaagag
aaaaaccaatgattcataacaatgtatgtgaaagtgtaaaatagaatgttact
ttggaatgactataaacattaaaagagactggagcat

Fig. 11

cctgtttcccagagtaatgtgggcatggagtcaggccacctcctctgggctc
tgctgttcatgcagtccttgtggcctcaactgactgatggagccactcgagtc
tactacctgggcatccgggatgtgcagtggaactatgctcccaaggaagaaa
tgtcatcacgaaccagcctctggacagtgacatagtggcttccagcttcttaa
agtctgacaagaaccggataggggggaacctacaagaagaccatctataaagaa
tacaaggatgactcatacacagatgaagtggcccagcctgcctgggttgggctt
cctgggggccagtgttgcaggctgaagtgggggatgtcattcttattcacctga
agaattttgccactcgtccctataccatccaccctcatgggtgtcttctacgag
aaggactctgaagggttccctataccagatggctcctctggggccactgaaagc
tgatgactctgttcccccggggggcagccatatctacaactggaccattccag
aaggccatgcacccaccgatgctgaccagcgtgcctcacctggatctaccat
tctcatgtagatgctccacgagacattgcaactggcctaattgggcctctcat
cacctgtaaaagaggagccctggatgggaactccctcctcaacgccaggatg
tagaccatgatttcttccctcctcttcagtgtggttagatgagaacctcagctgg
catctcaatgagaacattgccacttactgctcagatcctgcttcagtggacaa
agaagatgagacatttccaggagagcaataggatgcatgcaatcaatggctttg
tttttgggaatttacctgagctgaacatgtgtgcacagaaacgtgtggcctgg
cacttgtttggcatgggcaatgaaattgatgtccacacagcatttttccatgg
acagatgctgactaccgctggacaccacactgatgtggctaacatctttccag
ccacctttgtgactgctgagatgggtgccctgggaacctgggtacctgggttaatt
agctgccaagtgaacagtcactttcgagatggcatgcaggcactctacaagg
caagtcttgctccatggccccctcctgtggacctgctcacaggcaaagttcgac
agtacttcattgaggcccatgagattcaatgggactatggcccgatggggcat
gatgggagtagtgggaagaatttgagagagccaggcagtatctcagataagtt
tttccagaagagctccagccgaattgggggacacttactggaaagtgcgatatg
aagcctttcaagatgagacattccaagagaagatgcatttggaggaagatagg
catcttgggaatcctggggccagtgatccgggctgagggtgggtgacaccattca
gggtgggtcttctacaaccgtgcctcccagccattcagcatgcagccccatgggg
tcttttatgagaaagactatgaaggcactgtgtacaatgatggctcatcttac
cctggcttgggttgccaagccctttgagaaagtaacataccgctggacagtccc
ccctcatgccggtcccactgctcaggatcctgcttgtctcacttggatgtact
tctctgctgcagatcccataagagacacaaattctggcctgggtggggccgctg
ctgggtgtgcagggtggtgccttgggtgcagatggcaagcagaaaggggtgga
taaagaattctttcttctcttactgtgttgatgagaacaagagctggtaca
gcaatgccaatcaagcagctgctatgttggatttccgactgctttcagaggat
attgagggttccaagactccaatcggatgcatgccattaatgggtttctgtt
ctctaacctgcccaggctggacatgtgcaagggtgacacagtggcctggcacc
tgctcggcctggggcacagagactgatgtgcatggagtcatgttccagggaac

Fig. 11 (Continuation)

actgtgcagcttcagggcatgaggaaggggtgcagctatgctctttcctcatac
ctttgtcatggccatcatgcagcctgacaaccttgggacatttgagatttatt
gccaggcaggcagccatcgagaagcagggatgagggcaatctataatgtctcc
cagtgtcctgg
ccaccaagccacccctcgccaacgctaccaagctgcaagaatctactatatca
tggcagaagaagtagagtgggactattgcccctgaccggagctgggaacgggaa
tggcacaaccagtctgagaaggacagttatgggttacattttcctgagcaacaa
ggatgggctcctgggttccagatacaagaaagctgtattcaggggaatacactg
atgggtacattcaggatccctcggccaaggactggaccagaagaacacttggga
atcttgggtccacttatcaaaggtgaagttgggtgatatcctgactgtgggtatt
caagaataatgccagccgcccctactctgtgcatgctcatggagtgtagaat
ctactactgtctggccactggctgctgagcctgggtgaggtgggtcacttatcag
tggaacatcccagagaggtctggccctggggcccaatgactctgcttgtgtttc
ctggatctattattctgcagtggatcccatcaaggacatgtatagtggcctgg
tggggcccttggctatctgccaaaagggcatcctggagcccatggaggacgg
agtgcacatggatcgggaatttgcattgttgttcttgatttttgatgaaaataa
gtcttgggtatttggaggaaaatgtggcaaccatgggtcccaggatccaggca
gtattaacctacaggatgaaactttcttggagagcaataaaaatgcatgcaatc
aatgggaaactctatgccaaccttaggggtcttaccatgtaccaaggagaacg
agtggcctgggtacatgctggccatgggccaagatgtggatctacacaccatcc
actttcatgcagagagcttcctctatcggaatggcgagaactaccgggcagat
gtggtggatctgttcccagggaacttttgaggttgtggagatgggtggccagcaa
ccctgggacatggctgatgcactgccatgtgactgaccatgtccatgctggca
tggagaccctcttcactgttttttctcgaacagaacacttaagccctctcacc
gtcatcaccaaagagactgaaaaagcagtgccccccagagacattgaagaagg
caatgtgaagatgctgggcatgcagatccccataaagaatgttgagatgctgg
cctctgttttgggttgccattagtgtcacccttctgctcgttgttctggctctt
gggtggagtggtttggtagcaacatcgacagagaaagctacgacgcaataggag
gtccatcctggatgacagcttcaagcttctgtctttcaaacagtaacatctgg
agcctggagatatcctcaggaagcacatctgtagtgcactcccagcaggccat
ggactagtactaaccacacactcaaaggggcatgggtgggtggagaagcagaa
ggagcaatcaagcttatctggatatttctttctttattttattttacatggaaa
taatattgatttcactttttcttttagtttctttgctctacgtgggcacctggca
ctaagggagtagcttattatcctacatcgcaaatttcaacagctacattatat
ttccttctgacacttggaaaggtattgaaatttctagaaatgtatccttctcac
aaagtagagaccaagagaaaaactcattgattgggtttctacttctttcaagg
actcaggaaatttctactttgaactgaggccaagtgcagctgttaagataacca
cacttaaactaaaggctaagaatataggcttgatgggaaattgaaggtaggct
gagtattgggaatccaaattgaattttgattctccttggcagtggaactacttt

Fig. 11 (Continuation)

gaagaagtgggtcaatggggttgctgcatgagcatgtacaacctctggagc
tagaagctcctcaggaaagccagttctccaagttcttaacctgtggcactgaa
aggaatggtgagttacctcttcatgttttagacagcaaaccctatccattaaa
gtacttggttagaacact

Fig. 12

MESGHLLWALLFMQSLWPQLTDGATRVYYLGI RDVQWNYAPKGRNVITNQPLD
SDIVASSFLKSDKNRIGGTYYKTIYKEYKDDSYTDEVAQPAWLGFLGPVLQAE
VGDVILIHLLKNFATRPYTIHPHGVFYEKDSEGSLYPDGSSGPLKADDSVPPGG
SHIYNWTIPEGHAPTDADPACLTWIYHSHVDAPRDIATGLIGPLITCKRGALD
GNSPPQRQDVDHDFLLFSVVDENLSWHLNENIATYCSDPASVDKEDETFQES
NRMHAINGFVFGNLPPELNMCAQKRVAWHLFGMGNEIDVHTAFFHGQMLTTRGH
HTDVANIFPATFVTAEMVPWEPGTWLI SCQVNSHFRDGMQALYKVKSCSMAPP
VDLLTGKVRQYFIEAHEIQWDYGPMGHDGSTGKNLREPGSISDKFFQKSSSRI
GGTYWKVRYEAFQDETFQEKMHLEEDRHLGILGPVIRAEVGDTIQVVFYNRAS
QPFSMQPHGVFYEKDYEGTVYNDGSSYPGLVAKPFEKVTYRWTVP PHAGPTAQ
DPACLTWMYFSAADPIRDTNSGLVGPLLVC RAGALGADGKQKGV DKEFFLLFT
VL DENKSWYSNANQAAAMLDFRLLEDIEGFQDSNRMHAINGFLFSNLPR LDM
CKGDTVAVHLLGLGTETDVHGV MFQGNTVQLQGM RKGAA MLFPHTFVMAIMQP
DNLGTFEIIYCQAGSHREAGMRAIYNVSQCPGHQATPRQRYQAARIYYIMAEV
EWDYCPDRSWEREWHNQSEKDSYGYIFLSNKDGLLGSRYKKAVFREYTDGTFR
IPRPRTGPEEHLGILGPLIKGEVGDILT VVFKNNASRPYSVHAHGVLESTTVW
PLAAEPGEVVTYQWNI PERSGPGPNDSACVSWIYYSAVDPIKDMYSGLVG PLA
ICQKGILEPHGGRSDMDREFALLFLIFDENKSWYLEENVATHGSQDPGSINLQ
DETFLESNMHAINGKLYANLRGLTMYQGERVAWYMLAMGQDVDLHTIHFHAE
SFLYRNGENYRADVVDLFPGTFEVVEMVASNPGTWLMHCHVTDH VHAGMETLF
TVFSRTEHLSPLTVITKETEKAVPPRDI EEGNVKMLGMQIPIKNVEMLASVLV
AISVTLLLVLALGGVVWYQHRQRKLRRNRRSILDDSFKLLSFKQ

- 19/28 -

Fig. 13

BLASTP alignment of SEQ ID NO:2 against swiss|P00450|CERU_HUMAN
 CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE) //:tr^{emb}l|M13699|HSCERP_1 gene:
 "CP"; Human ceruloplasmin (ferroxidase) mRNA, complete cds. //:gp|M13699|180256
 gene: "CP"; Human ceruloplasmin (ferroxidase) mRNA, complete cds.

This hit is scoring at : 0.0 (expectation value). Alignment length (overlap) : 810
 Identities : 49 % Scoring matrix : BLOSUM62 (used to infer consensus pattern)
 Database searched : nrdb

Q:1 MHAINGFVFGNLP^{EL}NMCAQKRV^{AW}HLFGMGNEIDVHTA^{FF}HGQMLT^{TR}GHHTDVANIFP
 M:::NG::FG:LP L:MCA::RV.W:LFGMGNE:DVH.AFFHGQ.LT:::..D.N:FP
 H:259 MYSVNGYTFGSLPGLSMCAEDRVK^WYLFGMGNEVDVHAA^{FF}HGQALT^{NK}NYRIDTINLFP

ATFVTAEMVPWEP**GTWLIS**CQVNS**HRDGM**QALYKVKSCSMAPPVDLLTGK-VRQYFIEA
 AT...A MV. .PG.W::SCQ :H.: G:QA::V:C::.. D :GK VR.Y:I.A
 ATLFDAYMVAQNPG^{EW}MLSCQNLNHLKAGLQ^{AF}FQVQ^{EC}NKSSKDNIRGKHVRHY^YIAA
 HEIQWDYGP^{MG}HDGSTGKNLREP^{GS}ISDKFFQKSSSRIGGTYW^{KV}RYEAFQDET^F-QEKM
 .EI W:Y.P.G D T :NL..PGS S FF:::RIGG:Y K: Y:::D:F ..K
 EEIWNYP^{SG}IDIFTKENLTAPGSDSAV^{FF}EQGTTRIGGSYK^{KL}VYREYTDASFTNRKE
 HLEEDRHLGILGPVIRAEVGD^{TI}QVVFYNRASQ^{PF}SMQPHGVFYEKDYEGTVN-----D
 .E:.HLGILGPVI AEVGD^{TI}:V.F:N:::P.S::P GV :.K: EGT.Y: .
 RGP^{EE}HLGILGPVIWAEVGD^{TI}RVTFHNKGAYPLSIEPIGVRFNKNNEGTY^{SP}NPYPQ
 GSSYPGLVAK--PFEKV^{TY}RWTVPPHAGPTAQDPA^{CL}TWMYFSAADPIRDTNSGLVGLL
 ..S.P :. P E..TY.WTVP...GPT .DP.CL MY:SA.DP.:D :GL:GP:
 SRSVPPSASHVAPTETFTYEWTVPKEVGPTNADPVCLAKMY^{YS}AVDPTKIDIFTGLIGPMK

Fig. 13 (Continuation)

VCRAGALGADGKQKGVDDKEFFLLFTVLVDENKSWYNNANQAAAAMLDFFRLLEDIEGFQDSN
 :C:.G:L A:G:QK.VDKEF:L. TV.DEN:S . . N : : : E.FQ:SN
 ICKKGS LHANGRQKDVDDKEFFYLFP TVFDENESLLEDNIRMFTTAPDQVDKEDEDFQESN

 RMHAINGFLFSNLPRLDMCKGDTVAVHLLGLGTETDVHGVMFQNTVQLQGMKGAAMLF
 :MH::NGF::N P L.MCKGD:V.W:L...G.E.DVHG:.F.GNT. :G R:...A LF
 KMHSMNGFMYGNGPGLTMCKGDSVWVYLF.SAGNEADVHGIYFSGNTYLWRGERRDTANLF

 PHTFVMAIMQPDNLGTFEICYCAGSHREAGMRAIYNVSQCPGHQATPRQRYQAARIYYIM
 P.T :... M PD. GTF.: C . . H . . GM:.. Y.V:QC .Q:..... Y .R.YYI.
 PQTSLTLHMWPDTEGTFNFVECLTTDHYTGGMKQKYTVNQCR-RQSEDSTFYLGERTYYIA

 AEEVEWDYCPDRSWEREWNHQSEKDSYGIYFLSNKDGLLGSRYKKAFFREYTDGTFRIPR
 A EVEWDY.P.R.WE:E H: .E:: . . FL.. : :GS:YKK.V:R:YTD.TFR:P
 AVEVEWDYSPQREWEKELHHLQEQN-VSNAFLDKGEFYIGSKYKVVYRQYTDSTFRVPV

 PRTGPEEHLGILGPLIKGEVGDILT VVFKNNASRPYSVHAHGV-LESTTVWPLAAEPGEV
 .R...EEHLGILGP :...VGD :...:FKN A:RPYS:HAHGV .ES:TV P... PGE.
 ERKAEHEHLGILGPQLHADVGDKVKIIFKNMATRPYSIHAHGVQTESSTVTPTL--PGET

 VTYQWNIPERSGPGPNDSDACVSWIYYSAVDPIKDMYSGLVGPLAICQKGIIEPHGGRSDM
 :TY W.IPERSG.G..DSAC:.W.YYS.VD.:KD:YSGL:GPL.:C:: .L: . . R...
 LTYVWKIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCRRRPYLKVFNP RRKL

 DREFALLFLIFDENKSWYLEENVATHGSQDPGSINLQDETFFLESNKMHAINGKLYANLRG
 EFALLFL:FDEN:SWYL:N:T: S..P .:N .DE.F:ESNKMHAING:::NL:G
 --EFALLFLVFDENESWYLDNIIKTY-SDHPEKVNKDDEEFIESNKMHAINGRMFGNLQG

Fig. 13 (Continuation)

LTMVQGERVAWYMLAMGQVDLHTIHFAESFLYRNGENYRADVVDLFP**GTFEVVEMVAS**
 LTM: G:.V WY:.MG.:DLHT:HFH..SF Y:: Y.:DV.D:FPGT:::EM...
 LTMHVGDEVNWYLMGMGNEIDLHTVHFHGHGSFQYKHRGVYSSDVDFIFPGTYQTLEMFP
 NPGTWLMHCHVTDHVAHAGMETLFTVFSRTE 800
 .PG.WL:HCHVTDH:HAGMET.:TV....:
 TPGIWLLHCHVTDHVAHAGMETTYTVLQNE 1061

prosite MULTICOPPER_OXIDASE1 region in bold

BLOCKS for Multicopper oxidases proteins in bold, underlined

Fig. 14

Prositate search result			
PS000079	74->95	MULTICOPPER_OXIDASE1	PDOC000076
PS000079	426->447	MULTICOPPER_OXIDASE1	PDOC000076
PS000079	773->794	MULTICOPPER_OXIDASE1	PDOC000076
PS000080	778->790	MULTICOPPER_OXIDASE2	PDOC000076

Fig. 15

BLOCKS search results			
AC#	Description	Strength	Score
BL000079D	Multicopper oxidase protein	1315	1552

Fig. 16

HMMPFAM - alignment of 123_protein against pfam|hmm|Cu-oxidase

Multicopper oxidase

This hit is scoring at : 42.4 E=1e-08

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```

Q: 1 -----MH-AINGFVFGN--LPELNM
H: 1 llfvldenksyleakthvfkleinnvdkedeeefgapdgvinsrngmgfnsGslPgghpi
      CAQKRVAWHLFGMGNEID----VHTAFFHGQMLTTRG-----HHTDVANIFP--
      .. W L. :.N .D V :..FHG . ..D.. I P
      hlhghdfwlllinnltdntMlrvtsihfhGlfghtntvlalwadgPprrdtcpipPgg
      -ATFVTAEMVPW-EPGTWLI SCQ---VNSHFRDGMQALYKVKSC 99
      :.T .....PGTW ..C. :.H:DG::: V..
      AGssfty rFtaddnpGtWwyHCHPGM iswHyedGlrgvfivdpp 164

```

Fig. 17

HMPFAM - alignment of SEQ ID NO:2 against pfam|hmm|Cu-oxidase
Multicopper oxidase

This hit is scoring at : 86.1 E=7e-22 Scoring matrix : BLOSUM62 (used to infer
consensus pattern)

Q:313 LLFTVLDENKSWYSNANQAAAMLDFRLLEDIEGFQDSNRMH-AINGFLFSN--LPRLDM

LLF.VLDENKS Y .A... .L:... :::: E FQ .: : NG. F:: LP .:

H:1 llfvldenksyleakthvfkleinnvdkedeeefqapdgvinrngmgfnsGslPgghpi

CKGDTVAWHLLGLGTETD-----VHGVMFQGN TVQLQG-----MRKGAAMLFPH-

. W LL :... TD V : F.G . . . R:... : P

hlhghdfwllllinnltdntMlrvtshfhGlfqhgtntvlalwadgPprrdtcpipPg99

--TFVMAIMQPD-NLGTFEICYQ---AGSHREAGMRAIYNVSQC 451

: D N GT: :C. .. H E G:R::: V..

AgssftyrFtaddnpGtWwyHCHPGMIsWHyedGlrgrvfivdpp 164

Fig. 18

HMPFAM - alignment of SEQ ID NO:2 against pfam|hmm|Cu-oxidase
Multicopper oxidase

This hit is scoring at : 125.7 E=8.7e-34 Scoring matrix : BLOSUM62 (used to
infer consensus pattern)

```

Q: 656 LLFLIFDENKSWYLEENVATHGSGDPGSINLQDETFLSNKMH-AINGKLYAN--LRGLT
    LLF::DENKS YLE... : : : :DE.F : : : : NG. : : L G .
H: 1 llfvldenksyleakthvfk1.einnvdkedeeefgapdgvinrngmgfnsGslPgghp

    MYQGERVAWYMLAMGQDVD-----LHTIHFAESFLYRN-----GENYRADVVDLFP-
    :: . W.:L:...D : :IHFH. .N . R.D...: P
    ihlhghdfwvlllinnltdntMlrvtsihfhGlqhgntntvlalwadgPprrdtcpiPg

    --GTFEVVEMVAS-NPGTWLMHCH--VTDHVVHAGMETLFTVFSR 798
    G: .....A. NPGTW .HCH : : H.. G.:F.V .
    gAGssftyFtadnpGtWwyHCHPGMIsWHyedGlrGVfiVdpp 164
  
```

Fig. 19

BLASTP - alignment of 123 V2 against swiss|P00450|CERU HUMAN
 CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE) ./.:trembl|M13699|HSCERP_1 gene:
 "CP"; Human ceruloplasmin (ferroxidase) mRNA, complete cds. //:gp|M13699|180256
 gene: "CP"; Human ceruloplasmin (ferroxidase) mRNA, complete cds.

- This hit is scoring at : 0.0 (expectation value)
- Alignment length (overlap) : 1052
- Identities : 49 %
- Scoring matrix : BLOSUM62 (used to infer consensus pattern)
- Database searched : nrdb_1;

Q: 28 YYLGIRDVQWNYAPKG--RNVITNQPLSDIVASSFLKSDKNRIGGTYKKTIYKEYKDDSD
 YY:GI :..W:YA.. :.:I: .. .S:I :L::.:RIG .YKK.:Y :Y.D::

H: 24 YYIGIETTDYASDHGEKKLISVDTEHSNI---YLQNGPDRIGRLYKKALYLQYTDDET

YTDEVAQPAWLGFGLPVLQAEVGDVILHLKFNATRPYTIHPHGVFYEKDSGSLYPDGS
 :...:..P.WLGFGLP:::AE.GD :.:HLKN.A:RPYT.H.HG: Y K.:EG::YPD::
 FRTIEKPVWLGFGLPPIKAETGDKVYVHLKNLASRPYTFHSHGITYYKEHEGAIYPDNT

SGPLKADDSVPPGGSHIYNWTIPEGHAPTDADPACLTWYHSHVDAPRDIATGLIGPLIT
 :. :ADD.V PG :.:Y :..E :.P :.D C:T IYHSH:DAP:DIA:GLIGPLI.
 TDFQRADDKVYPGEQYTYMILLATEEQSPGEGDGNVCVTRIYHSHIDAPKDIASGLIGPLII

CKRGALDGNPPQRQDVHDHDFLLFSVVDENL SWHLNENIATYCSDPASVDKEDETFQES
 CK:::LD . :...:D::F::FSVVDEN.SW:L.:NI.TYCS:P..VDK:::E.FQES
 CKKDSLDKE--KEKHIDREFVVMFSVVDENFSWYLEDNIKTYCSEPEKVKDKDNEDFFQES

NRMHAINGFVFGNLPFLNMCAQKRVAWHLFGMGNEIDVHTAFFHGQMLTTRGHHTDVANI
 NRM:::NG:.FG:LP L:MCA:.RV.W:LFGMGNE:DVH.AFFHGQ.LT::...D..N:
 NRMYSVNGYTFGSLPGLSMCAEDRVKWLFLFGMGNEVDVHAFFHGQALTNKNYRIDTINL

- 27/28 -

Fig. 19 (Continuation)

FPATFVTAEMVPWPGTWLISCQVNSHFRDGMQALYKVKSCSMAPPVDLLTGK-VRQYFI
 FPAT...A MV. .PG.W::SCQ :H.: G:QA::V:C:... D :.GK VR.Y:I
 FPATLFDAYMVAQNPGEWMLSCQNLNHLKAGLQAFFQVQECNKSSSKDNIRGKHVRHYI
 EAHEIQWDYGPMGHGSTGKNLREPGSISDKFFQKSSSRIGGTYWKVRYEAFQDETTF-QE
 .A.EI W:Y.P.G D T :NL..PGS S FF:::RIGG:Y K: Y...D.:F ..
 AAEIWNYPGIDIFTKENLTAPGSDSAVFFEQGTTRIGGSYKKLVYREYTDASFTNR
 KMHLEEDRHLGILGPVIRAEVGDITQVVFYNRASQPFMSQPHGVFYEKDYEGTVYN---
 K . .E:.HLGILGPVI AEVGDTI:V.F:N:::P.S::P GV :.K: EGT.Y:
 KERGPHEEHLGILGPVIWAEVGDITRVTFNKGAYPPLSIEPIGVRFNKNNEGTYSPNYN
 -DGSSYPGLVA--KPFKEVTYRWTVPPHAGPTAQDPACLTWMYFSAADPIRDTNSGLVGP
 ...S.P .: .P E..TY.WTVP...GPT .DP.CL. MY:SA.DP.:D. :GL:GP
 PQRSVPPSASHVAPTETFTYEWTVPKVEGPTNADPVCIAKMYSAVDPTKDIFTGLIGP
 LLVCRAGALGADGKQKGVDKKEFFLLFTVL DENKSWYSNANQAAAMLDFRLLEDIEGFQD
 : :C:.G:L A:G:QK.VDKEF:L. TV.DEN:S . . N :... E.FQ:
 MKICKKGS LHANGRQKDV DKEFYLFPTVFDENESLLEDNIRMFTTAPDQVDKEDDFQE
 SNRMHAINGFLFSNLPRLLDMCKGDTVAWHLLGLGTETDVHGVMFQGN TVQLQGMKGAAM
 SN:MH::NGF::N P L.MCKGD:V.W:L...G.E.DVHG:.F.GNT. :G R:...A
 SNKMHSNMNGFMYGNQPGLTMCCKGDSVWVYLF SAGNEADVHGIYFSGNTYLWRGERDTAN
 LFPHTFVMAIMQPDNLGTFEICYQAGSHREAGMRAIYNVSQCPGHQATPRQYQAARIYY
 LFP.T :.. M PD. GTF:: C . .H .GM:: Y.V:QC .Q:.... Y .R.YY
 LFPQTSLLTHMWPDTEGTFNVECLTTDHYTGMMKQKYTVNQ-C-RRQSEDSTFFYLGETYY

- 28/28 -

Fig. 19 (Continuation)

IMAEEVEWDYCPDRSWEREWHNQSEKDSYGYIFLSNKDGLLGSRYKKAVFREYTDGTFRI
 I.A EVEWDY.P.R.WE:E H: .E: :... .FL... : :GS:YKK.V:R:YTD.TFR:
 IAAVEVEWDYSPQREWEKELHHLQEQ-NVSN AFLDKGEFYIGSKYKKVVYRQYTDSTFRV

 PRPRTGPEEHLGILGPLIKGEVGDILT VVFKNNASRPSVHAHGV-LESTTVWPLAAEPG
 P .R...EEHLGILGP :...VGD :...:FKN A:RPYS:HAHGV .ES:TV P... PG
 PVERKAE EHLGILGPQLHADVGDKVKIIFKNMATRPYSIHAHGVQTESSTVTPTL--PG

 EVVTYQWNIPERSGPGPND SACVSWIYYSAVDPIKMYSGLVGLAICQKGILEPHGGRS
 E.:TY W.IPERSG.G..DSAC:.W.YYS.VD.:KD:YSGL:GPL.:C.: .L: ..R.
 ETLTYVWKIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCRRPYLKVFNPRR

 DMDREFALLFLIFDENKSWYLEENVATHGSDPGSINLQDETTFLESNKMHAINGKLYANL
 .: EFALLFL:FDEN:SWYL:N:T: S..P .:N .DE.F:ESNKMHAING:::NL
 KL--EFALLFLVFDENESWYLD DNIKT Y-SDHPEKVNKDDDEFFIESNKMHAINGRMFGNL

 RGLTMYQGERVAWYMLAMGQD VDLHTIHFHAE SFLYRNGENYRADVVDLFPGTFEVVEMV
 :GLTM: G:.V WY::MG::DLHT:HFH..SF Y:: Y:.DV.D:FPGT:::EM.
 QGLTMHVGDENVWYLMGMGNEIDLHTVHFHGH SFQYKHRGVYSSDVFDIFPGTYQTLEMF

 ASNPGTWMHCHVT DHHVHAGMETLFTVFSRTE 1067
 ...PG.WL:HCHVT DH:HAGMET.:TV.....
 PRTPGIWLLHCHVT DHHHAGMETTYTVLQNE 1061

- 1 -

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN CERULOPLASMIN-LIKE PROTEIN

<130> LIO188 Foreign Countries

<150> US 60/242,426

<151> 2000-10-24

<150> US 60/313,796

<151> 2001-08-22

<160> 11

<170> PatentIn version 3.1

<210> 1

<211> 4227

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (839) .. (3514)

<223>

<400> 1

ggaaaagagg gcacccagcc cttccccctc cctcatcctc ccatcccagt aaaccctgcc

60

- 2 -

aaattggaat cctggactta atttaggaga aaggccctgt aaccaagata ctgactgaac	120
atggctggcg gactcaggct ggggtctgca gtgcagcatt aatgggcccgc tgacatgaat	180
atggagtagt tttctctagc aaagagtggc ttccagcttc ttaaagtctg acaagaaccg	240
gataggggga acctacaaga agaccatcta taaagaatac aaggatgact catacacaga	300
tgaagtggcc cagcctgcct ggttgggctt cctggggcca gtgttcgagg ctgaagtggg	360
ggatgtcatt cttattcacc tgaagaattt tgccactcgt ccctatacca tccaccctca	420
tgggtgtcttc tacgagaagg actctgaagg ttccctatac ccagatggct cctctggggc	480
actgaaagct gatgactctg ttccccggg gggcagccat atctacaact ggaccattcc	540
agaaggccat gcaccaccg atgctgaccc agcgtgcctc acctggatct accattctca	600
tgtagatgct ccacgagaca ttgcaactgg cctaattggg cctctcatca cctgtaaaag	660
aggagccctg gatgggaact cccctcctca acgccaggat gtagaccatg atttcttct	720
cctcttcagt gtggtagatg agaacctcag ctggcatctc aatgagaaca ttgccactta	780
ctgctcagat cctgcttcag tggacaaaga agatgagaca tttcaggaga gcaatagg	838
atg cat gca atc aat ggc ttt gtt ttt ggg aat tta cct gag ctg aac	886
Met His Ala Ile Asn Gly Phe Val Phe Gly Asn Leu Pro Glu Leu Asn	
1 5 10 15	
atg tgt gca cag aaa cgt gtg gcc tgg cac ttg ttt ggc atg ggc aat	934
Met Cys Ala Gln Lys Arg Val Ala Trp His Leu Phe Gly Met Gly Asn	
20 25 30	
gaa att gat gtc cac aca gca ttt ttc cat gga cag atg ctg act acc	982
Glu Ile Asp Val His Thr Ala Phe His Gly Gln Met Leu Thr Thr	
35 40 45	
cgt gga cac cac act gat gtg gct aac atc ttt cca gcc acc ttt gtg	1030
Arg Gly His His Thr Asp Val Ala Asn Ile Phe Pro Ala Thr Phe Val	
50 55 60	
act gct gag atg gtg ccc tgg gaa cct ggt acc tgg tta att agc tgc	1078
Thr Ala Glu Met Val Pro Trp Glu Pro Gly Thr Trp Leu Ile Ser Cys	
65 70 75 80	
caa gtg aac agt cac ttt cga gat ggc atg cag gca ctc tac aag gtc	1126
Gln Val Asn Ser His Phe Arg Asp Gly Met Gln Ala Leu Tyr Lys Val	
85 90 95	
aag tot tgc tcc atg gcc cct cct gtg gac ctg ctc aca ggc aaa gtt	1174
Lys Ser Cys Ser Met Ala Pro Pro Val Asp Leu Leu Thr Gly Lys Val	
100 105 110	
cga cag tac ttc att gag gcc cat gag att caa tgg gac tat ggc ccg	1222
Arg Gln Tyr Phe Ile Glu Ala His Glu Ile Gln Trp Asp Tyr Gly Pro	
115 120 125	
atg ggg cat gat ggg agt act ggg aag aat ttg aga gag cca ggc agt	1270

- 3 -

Met Gly His Asp Gly Ser Thr Gly Lys Asn Leu Arg Glu Pro Gly Ser	
130 135 140	
atc tca gat aag ttt ttc cag aag agc tcc agc cga att ggg ggc act	1318
Ile Ser Asp Lys Phe Phe Gln Lys Ser Ser Arg Ile Gly Gly Thr	
145 150 155 160	
tac tgg aaa gtg cga tat gaa gcc ttt caa gat gag aca ttc caa gag	1366
Tyr Trp Lys Val Arg Tyr Glu Ala Phe Gln Asp Glu Thr Phe Gln Glu	
165 170 175	
aag atg cat ttg gag gaa gat agg cat ctt gga atc ctg ggg cca gtg	1414
Lys Met His Leu Glu Glu Asp Arg His Leu Gly Ile Leu Gly Pro Val	
180 185 190	
atc cgg gct gag gtg ggt gac acc att cag gtg gtc ttc tac aac cgt	1462
Ile Arg Ala Glu Val Gly Asp Thr Ile Gln Val Val Phe Tyr Asn Arg	
195 200 205	
gcc tcc cag cca ttc agc atg cag ccc cat ggg gtc ttt tat gag aaa	1510
Ala Ser Gln Pro Phe Ser Met Gln Pro His Gly Val Phe Tyr Glu Lys	
210 215 220	
gac tat gaa ggc act gtg tac aat gat ggc tca tct tac cct ggc ttg	1558
Asp Tyr Glu Gly Thr Val Tyr Asn Asp Gly Ser Ser Tyr Pro Gly Leu	
225 230 235 240	
gtt gcc aag ccc ttt gag aaa gta aca tac cgc tgg aca gtc ccc cct	1606
Val Ala Lys Pro Phe Glu Lys Val Thr Tyr Arg Trp Thr Val Pro Pro	
245 250 255	
cat gcc ggt ccc act gct cag gat cct gct tgt ctc act tgg atg tac	1654
His Ala Gly Pro Thr Ala Gln Asp Pro Ala Cys Leu Thr Trp Met Tyr	
260 265 270	
ttc tct gct gca gat ccc ata aga gac aca aat tct ggc ctg gtg ggc	1702
Phe Ser Ala Ala Asp Pro Ile Arg Asp Thr Asn Ser Gly Leu Val Gly	
275 280 285	
ccg ctg ctg gtg tgc agg gct ggt gcc ttg ggt gca gat ggc aag cag	1750
Pro Leu Leu Val Cys Arg Ala Gly Ala Leu Gly Ala Asp Gly Lys Gln	
290 295 300	
aaa ggg gtg gat aaa gaa ttc ttt ctt ctc ttc act gtg ttg gat gag	1798
Lys Gly Val Asp Lys Glu Phe Phe Leu Leu Phe Thr Val Leu Asp Glu	
305 310 315 320	
aac aag agc tgg tac agc aat gcc aat caa gca gct gct atg ttg gat	1846
Asn Lys Ser Trp Tyr Ser Asn Ala Asn Gln Ala Ala Ala Met Leu Asp	
325 330 335	
ttc cga ctg ctt tca gag gat att gag ggc ttc caa gac tcc aat cgg	1894
Phe Arg Leu Leu Ser Glu Asp Ile Glu Gly Phe Gln Asp Ser Asn Arg	
340 345 350	
atg cat gcc att aat ggg ttt ctg ttc tot aac ctg ccc agg ctg gac	1942
Met His Ala Ile Asn Gly Phe Leu Phe Ser Asn Leu Pro Arg Leu Asp	
355 360 365	

- 4 -

atg tgc aag ggt gac aca gtg gcc tgg cac ctg ctc ggc ctg ggc aca Met Cys Lys Gly Asp Thr Val Ala Trp His Leu Leu Gly Leu Gly Thr 370 375 380	1990
gag act gat gtg cat gga gtc atg ttc cag ggc aac act gtg cag ctt Glu Thr Asp Val His Gly Val Met Phe Gln Gly Asn Thr Val Gln Leu 385 390 395 400	2038
cag ggc atg agg aag ggt gca gct atg ctc ttt cct cat acc ttt gtc Gln Gly Met Arg Lys Gly Ala Ala Met Leu Phe Pro His Thr Phe Val 405 410 415	2086
atg gcc atc atg cag cct gac aac ctt ggg aca ttt gag att tat tgc Met Ala Ile Met Gln Pro Asp Asn Leu Gly Thr Phe Glu Ile Tyr Cys 420 425 430	2134
cag gca ggc agc cat cga gaa gca ggg atg agg gca atc tat aat gtc Gln Ala Gly Ser His Arg Glu Ala Gly Met Arg Ala Ile Tyr Asn Val 435 440 445	2182
tcc cag tgt cct ggc cac caa gcc acc cct cgc caa cgc tac caa gct Ser Gln Cys Pro Gly His Gln Ala Thr Pro Arg Gln Arg Tyr Gln Ala 450 455 460	2230
gca aga atc tac tat atc atg gca gaa gaa gta gag tgg gac tat tgc Ala Arg Ile Tyr Tyr Ile Met Ala Glu Glu Val Glu Trp Asp Tyr Cys 465 470 475 480	2278
cct gac cgg agc tgg gaa cgg gaa tgg cac aac cag tct gag aag gac Pro Asp Arg Ser Trp Glu Arg Glu Trp His Asn Gln Ser Glu Lys Asp 485 490 495	2326
agt tat ggt tac att ttc ctg agc aac aag gat ggg ctc ctg ggt tcc Ser Tyr Gly Tyr Ile Phe Leu Ser Asn Lys Asp Gly Leu Leu Gly Ser 500 505 510	2374
aga tac aag aaa gct gta ttc agg gaa tac act gat ggt aca ttc agg Arg Tyr Lys Lys Ala Val Phe Arg Glu Tyr Thr Asp Gly Thr Phe Arg 515 520 525	2422
atc cct cgg cca agg act gga cca gaa gaa cac ttg gga atc ttg ggt Ile Pro Arg Pro Arg Thr Gly Pro Glu Glu His Leu Gly Ile Leu Gly 530 535 540	2470
cca ctt atc aaa ggt gaa gtt ggt gat atc ctg act gtg gta ttc aag Pro Leu Ile Lys Gly Glu Val Gly Asp Ile Leu Thr Val Val Phe Lys 545 550 555 560	2518
aat aat gcc agc cgc ccc tac tct gtg cat gct cat gga gtg cta gaa Asn Asn Ala Ser Arg Pro Tyr Ser Val His Ala His Gly Val Leu Glu 565 570 575	2566
tct act act gtc tgg cca ctg gct gct gag cct ggt gag gtg gtc act Ser Thr Thr Val Trp Pro Leu Ala Ala Glu Pro Gly Glu Val Val Thr 580 585 590	2614
tat cag tgg aac atc cca gag agg tct ggc cct ggg ccc aat gac tct Tyr Gln Trp Asn Ile Pro Glu Arg Ser Gly Pro Gly Pro Asn Asp Ser 595 600 605	2662

- 5 -

gct tgt gtt tcc tgg atc tat tat tct gca gtg gat ccc atc aag gac Ala Cys Val Ser Trp Ile Tyr Tyr Ser Ala Val Asp Pro Ile Lys Asp 610 615 620	2710
atg tat agt ggc ctg gtg ggg ccc ttg gct atc tgc caa aag ggc atc Met Tyr Ser Gly Leu Val Gly Pro Leu Ala Ile Cys Gln Lys Gly Ile 625 630 635 640	2758
ctg gag ccc cat gga gga cgg agt gac atg gat cgg gaa ttt gca ttg Leu Glu Pro His Gly Gly Arg Ser Asp Met Asp Arg Glu Phe Ala Leu 645 650 655	2806
ttg ttc ttg att ttt gat gaa aat aag tct tgg tat ttg gag gaa aat Leu Phe Leu Ile Phe Asp Glu Asn Lys Ser Trp Tyr Leu Glu Glu Asn 660 665 670	2854
gtg gca acc cat ggg tcc cag gat cca ggc agt att aac cta cag gat Val Ala Thr His Gly Ser Gln Asp Pro Gly Ser Ile Asn Leu Gln Asp 675 680 685	2902
gaa act ttc ttg gag agc aat aaa atg cat gca atc aat ggg aaa ctc Glu Thr Phe Leu Glu Ser Asn Lys Met His Ala Ile Asn Gly Lys Leu 690 695 700	2950
tat gcc aac ctt agg ggt ctt acc atg tac caa gga gaa cga gtg gcc Tyr Ala Asn Leu Arg Gly Leu Thr Met Tyr Gln Gly Glu Arg Val Ala 705 710 715 720	2998
tgg tac atg ctg gcc atg ggc caa gat gtg gat cta cac acc atc cac Trp Tyr Met Leu Ala Met Gly Gln Asp Val Asp Leu His Thr Ile His 725 730 735	3046
ttt cat gca gag agc ttc ctc tat cgg aat ggc gag aac tac cgg gca Phe His Ala Glu Ser Phe Leu Tyr Arg Asn Gly Glu Asn Tyr Arg Ala 740 745 750	3094
gat gtg gtg gat ctg ttc oca ggg act ttt gag gtt gtg gag atg gtg Asp Val Val Asp Leu Phe Pro Gly Thr Phe Glu Val Val Glu Met Val 755 760 765	3142
gcc agc aac cct ggg aca tgg ctg atg cac tgc cat gtg act gac cat Ala Ser Asn Pro Gly Thr Trp Leu Met His Cys His Val Thr Asp His 770 775 780	3190
gtc cat gct ggc atg gag acc ctc ttc act gtt ttt tct cga aca gaa Val His Ala Gly Met Glu Thr Leu Phe Thr Val Phe Ser Arg Thr Glu 785 790 795 800	3238
cac tta agc cct ctc acc gtc atc acc aaa gag act gaa aaa gca gtg His Leu Ser Pro Leu Thr Val Ile Thr Lys Glu Thr Glu Lys Ala Val 805 810 815	3286
ccc ccc aga gac att gaa gaa ggc aat gtg aag atg ctg ggc atg cag Pro Pro Arg Asp Ile Glu Glu Gly Asn Val Lys Met Leu Gly Met Gln 820 825 830	3334
atc ccc ata aag aat gtt gag atg ctg gcc tct gtt ttg gtt gcc att Ile Pro Ile Lys Asn Val Glu Met Leu Ala Ser Val Leu Val Ala Ile	3382

- 6 -

835	840	845	
agt gtc acc ctt ctg ctc gtt gtt ctg gct ctt ggt gga gtg gtt tgg			3430
Ser Val Thr Leu Leu Leu Val Val Leu Ala Leu Gly Gly Val Val Trp			
850	855	860	
tac caa cat cga cag aga aag cta cga cgc aat agg agg tcc atc ctg			3478
Tyr Gln His Arg Gln Arg Lys Leu Arg Arg Asn Arg Arg Ser Ile Leu			
865	870	875	880
gat gac agc ttc aag ctt ctg tct ttc aaa cag taa catctggagc			3524
Asp Asp Ser Phe Lys Leu Leu Ser Phe Lys Gln			
885	890		
ctggagatat cctcaggaag cacatctgta gtgcactccc agcaggccat ggactagtca			3584
ctaaccocac actcaaaggg gcatgggtgg tggagaagca gaaggagcaa tcaagcttat			3644
ctggatatattt ctttctttat ttattttaca tggaaataat atgatttcac tttttcttta			3704
gtttctttgc tctacgtggg cacctggcac taaggagta ccttattatc ctacatcgca			3764
aatttcaaca gotacattat atttccttct gacacttggg aggtattgaa atttctagaa			3824
atgtatcctt ctcacaaagt agagaccaag agaaaaactc attgattggg tttctacttc			3884
tttcaaggac tcaggaaatt tcactttgaa ctgaggccaa gtgagctgtt aagataaccc			3944
acacttaaac taaaggctaa gaatataggc ttgatgggaa attgaaggta ggctgagtat			4004
tgggaatoca aattgaattt tgattctcct tggcagttaa ctactttgaa gaagtgggtca			4064
atgggttggt gctgccatga gcatgtacaa cctctggagc tagaagctcc tcaggaaaagc			4124
cagttctoca agttcttaac ctgtggcact gaaaggaatg ttgagttacc tcttcatggt			4184
ttagacagca aacctatcc attaaagtac ttgttagaac act			4227

<210> 2

<211> 891

<212> PRT

<213> Homo sapiens

<400> 2

Met	His	Ala	Ile	Asn	Gly	Phe	Val	Phe	Gly	Asn	Leu	Pro	Glu	Leu	Asn
1				5					10					15	

Met	Cys	Ala	Gln	Lys	Arg	Val	Ala	Trp	His	Leu	Phe	Gly	Met	Gly	Asn
		20						25					30		

- 7 -

Glu Ile Asp Val His Thr Ala Phe Phe His Gly Gln Met Leu Thr Thr
 35 40 45

Arg Gly His His Thr Asp Val Ala Asn Ile Phe Pro Ala Thr Phe Val
 50 55 60

Thr Ala Glu Met Val Pro Trp Glu Pro Gly Thr Trp Leu Ile Ser Cys
 65 70 75 80

Gln Val Asn Ser His Phe Arg Asp Gly Met Gln Ala Leu Tyr Lys Val
 85 90 95

Lys Ser Cys Ser Met Ala Pro Pro Val Asp Leu Leu Thr Gly Lys Val
 100 105 110

Arg Gln Tyr Phe Ile Glu Ala His Glu Ile Gln Trp Asp Tyr Gly Pro
 115 120 125

Met Gly His Asp Gly Ser Thr Gly Lys Asn Leu Arg Glu Pro Gly Ser
 130 135 140

Ile Ser Asp Lys Phe Phe Gln Lys Ser Ser Ser Arg Ile Gly Gly Thr
 145 150 155 160

Tyr Trp Lys Val Arg Tyr Glu Ala Phe Gln Asp Glu Thr Phe Gln Glu
 165 170 175

Lys Met His Leu Glu Glu Asp Arg His Leu Gly Ile Leu Gly Pro Val
 180 185 190

Ile Arg Ala Glu Val Gly Asp Thr Ile Gln Val Val Phe Tyr Asn Arg
 195 200 205

Ala Ser Gln Pro Phe Ser Met Gln Pro His Gly Val Phe Tyr Glu Lys
 210 215 220

Asp Tyr Glu Gly Thr Val Tyr Asn Asp Gly Ser Ser Tyr Pro Gly Leu
 225 230 235 240

Val Ala Lys Pro Phe Glu Lys Val Thr Tyr Arg Trp Thr Val Pro Pro
 245 250 255

His Ala Gly Pro Thr Ala Gln Asp Pro Ala Cys Leu Thr Trp Met Tyr
 260 265 270

- 8 -

Phe Ser Ala Ala Asp Pro Ile Arg Asp Thr Asn Ser Gly Leu Val Gly
 275 280 285

Pro Leu Leu Val Cys Arg Ala Gly Ala Leu Gly Ala Asp Gly Lys Gln
 290 295 300

Lys Gly Val Asp Lys Glu Phe Phe Leu Leu Phe Thr Val Leu Asp Glu
 305 310 315 320

Asn Lys Ser Trp Tyr Ser Asn Ala Asn Gln Ala Ala Ala Met Leu Asp
 325 330 335

Phe Arg Leu Leu Ser Glu Asp Ile Glu Gly Phe Gln Asp Ser Asn Arg
 340 345 350

Met His Ala Ile Asn Gly Phe Leu Phe Ser Asn Leu Pro Arg Leu Asp
 355 360 365

Met Cys Lys Gly Asp Thr Val Ala Trp His Leu Leu Gly Leu Gly Thr
 370 375 380

Glu Thr Asp Val His Gly Val Met Phe Gln Gly Asn Thr Val Gln Leu
 385 390 395 400

Gln Gly Met Arg Lys Gly Ala Ala Met Leu Phe Pro His Thr Phe Val
 405 410 415

Met Ala Ile Met Gln Pro Asp Asn Leu Gly Thr Phe Glu Ile Tyr Cys
 420 425 430

Gln Ala Gly Ser His Arg Glu Ala Gly Met Arg Ala Ile Tyr Asn Val
 435 440 445

Ser Gln Cys Pro Gly His Gln Ala Thr Pro Arg Gln Arg Tyr Gln Ala
 450 455 460

Ala Arg Ile Tyr Tyr Ile Met Ala Glu Glu Val Glu Trp Asp Tyr Cys
 465 470 475 480

Pro Asp Arg Ser Trp Glu Arg Glu Trp His Asn Gln Ser Glu Lys Asp
 485 490 495

Ser Tyr Gly Tyr Ile Phe Leu Ser Asn Lys Asp Gly Leu Leu Gly Ser
 500 505 510

- 9 -

Arg Tyr Lys Lys Ala Val Phe Arg Glu Tyr Thr Asp Gly Thr Phe Arg
 515 520 525

Ile Pro Arg Pro Arg Thr Gly Pro Glu Glu His Leu Gly Ile Leu Gly
 530 535 540

Pro Leu Ile Lys Gly Glu Val Gly Asp Ile Leu Thr Val Val Phe Lys
 545 550 555 560

Asn Asn Ala Ser Arg Pro Tyr Ser Val His Ala His Gly Val Leu Glu
 565 570 575

Ser Thr Thr Val Trp Pro Leu Ala Ala Glu Pro Gly Glu Val Val Thr
 580 585 590

Tyr Gln Trp Asn Ile Pro Glu Arg Ser Gly Pro Gly Pro Asn Asp Ser
 595 600 605

Ala Cys Val Ser Trp Ile Tyr Tyr Ser Ala Val Asp Pro Ile Lys Asp
 610 615 620

Met Tyr Ser Gly Leu Val Gly Pro Leu Ala Ile Cys Gln Lys Gly Ile
 625 630 635 640

Leu Glu Pro His Gly Gly Arg Ser Asp Met Asp Arg Glu Phe Ala Leu
 645 650 655

Leu Phe Leu Ile Phe Asp Glu Asn Lys Ser Trp Tyr Leu Glu Glu Asn
 660 665 670

Val Ala Thr His Gly Ser Gln Asp Pro Gly Ser Ile Asn Leu Gln Asp
 675 680 685

Glu Thr Phe Leu Glu Ser Asn Lys Met His Ala Ile Asn Gly Lys Leu
 690 695 700

Tyr Ala Asn Leu Arg Gly Leu Thr Met Tyr Gln Gly Glu Arg Val Ala
 705 710 715 720

Trp Tyr Met Leu Ala Met Gly Gln Asp Val Asp Leu His Thr Ile His
 725 730 735

Phe His Ala Glu Ser Phe Leu Tyr Arg Asn Gly Glu Asn Tyr Arg Ala

- 10 -

740					745					750					
Asp	Val	Val	Asp	Leu	Phe	Pro	Gly	Thr	Phe	Glu	Val	Val	Glu	Met	Val
		755					760					765			
Ala	Ser	Asn	Pro	Gly	Thr	Trp	Leu	Met	His	Cys	His	Val	Thr	Asp	His
		770					775					780			
Val	His	Ala	Gly	Met	Glu	Thr	Leu	Phe	Thr	Val	Phe	Ser	Arg	Thr	Glu
						790						795			800
His	Leu	Ser	Pro	Leu	Thr	Val	Ile	Thr	Lys	Glu	Thr	Glu	Lys	Ala	Val
				805					810					815	
Pro	Pro	Arg	Asp	Ile	Glu	Glu	Gly	Asn	Val	Lys	Met	Leu	Gly	Met	Gln
			820					825					830		
Ile	Pro	Ile	Lys	Asn	Val	Glu	Met	Leu	Ala	Ser	Val	Leu	Val	Ala	Ile
			835					840					845		
Ser	Val	Thr	Leu	Leu	Leu	Val	Val	Leu	Ala	Leu	Gly	Gly	Val	Val	Trp
			850				855					860			
Tyr	Gln	His	Arg	Gln	Arg	Lys	Leu	Arg	Arg	Asn	Arg	Arg	Ser	Ile	Leu
						870					875				880
Asp	Asp	Ser	Phe	Lys	Leu	Leu	Ser	Phe	Lys	Gln					
				885					890						

<210> 3

<211> 567

<212> DNA

<213> Homo sapiens

<400> 3

gaactttctt ggagagtaat ataatcatgc aatcaatggg aaactctatg ccaaccttag	60
gggtcttacc atgtaccaag gagaacgagt ggcctggtac atgctggcca tgggccaaga	120
tgtggatcta cacaccatcc actttcatgc agagagcttc ctctatcgga atggcgagaa	180
ctaccgggca gatgtggtgg atctgttccc agggactttt gaggttgtgg agatggtggc	240
cagcaaccct gggacatggc tgatgcaactg ccatgtgact gaccatgtcc atgctggcat	300

- 11 -

```

ggagaccctc ttcactgttt tttctcgaac agaacactta agccctotca ccgtcatcac   360
caaagagact gaaaaagcag tgccccccag agacattgaa gaaggcaatg tgaagatgct   420
gggcatgcag atccccataa agaatgttga gatgctggcc tctgttttgg ttgccattag   480
tgtcaccctt ctgctcgttg ttctggctct tggaggagtg gtttggtacc aacatcgaca   540
gagaaagcta cgacgcaata ggaggtc                                     567

```

<210> 4

<211> 363

<212> DNA

<213> Homo sapiens

```

<400> 4
gaacagaaca cttaagccct ctcaccgtca tcaccaaaga gactgaaaaa gtgcccccca   60
gagacattga agaaggcaat gtgaagatgc tgggcatgca gatccccata aagaatgttg   120
agatgctggc ctctgttttg gttgccatta gtgtcaccct tctgctcgtt gttctggctc   180
ttggtggagt ggtttggtac caacatcgac agagaaagct acgacgcaat aggagggtcca   240
tcctggatga cagcttcaag cttctgtott tcaacagtac atctggagcc tggagatata   300
ctcaggaagc acatctgtag tgcaactcca gcaggccatg gactagtcac taaccacaca   360
ctc                                     363

```

<210> 5

<211> 251

<212> DNA

<213> Homo sapiens

```

<400> 5
gcgagcatgt accaggccac togtgctoot tggtagatag taagaccctt aagggttgga   60
tagagtttcc cattgattgc atgcatttta ttgctctcca agaaagtttc atcctgtagg   120
ttaatactgc ctggatcctg ggacccatgg gttgccacat tttcctccaa ataccaagac   180
ttattttcat caaaaatcaa gaacaacaat gcaaattccc gatccatgtc actccgtcct   240
ccatgggcca t                                     251

```

- 12 -

<210> 6

<211> 240

<212> DNA

<213> Homo sapiens

<400> 6

tcgaacagaa cacttaagcc ctctgaccgt catcaccaaa gagactgaaa aagcagtggc	60
ccccagagac attgaagaag gcaatgtgaa gatgctgggc atgcagatcc ccataaagaa	120
tgttgagatg ctggcctctg ttttggttgc cattagtgtc acccttctgc tcgttgttct	180
ggctcttggg ggagttgttt ggtaccaaca tcgacagaga aagctacgac gcaataggag	240

<210> 7

<211> 3321

<212> DNA

<213> Homo sapiens

<400> 7

atgaagattt tgatacttgg tatttttctg tttttatgta gtacccacgc ctgggcgaaa	60
gaaaagcatt attacattgg aattattgaa acgacttggg attatgootc tgaccatggg	120
gaaaagaaac ttatttctgt tgacacggaa cattccaata totatottca aaatggccca	180
gatagaattg ggagactata taagaaggcc ctttatcttc agtacacaga tgaaaccttt	240
aggacaacta tagaaaaacc ggtctggctt gggtttttag gccctattat caaagctgaa	300
actggagata aagtttatgt aacttaaaa aaccttgcct ctaggcocct cacctttcat	360
tcacatggaa taacttacta taaggaacat gagggggcca totacootga taacaccaca	420
gattttcaaa gagcagatga caaagtatat ccaggagagc agtatacata catgttgctt	480
gccactgaag aacaaagtcc tggggaagga gatggcaatt gtgtgactag gatttaccat	540
tcccacattg atgctccaaa agatattgcc tcaggactca tcggaccttt aataatctgt	600
aaaaaagatt ctctagataa agaaaaagaa aaacatattg accgagaatt tgtggtgatg	660
ttttctgtgg tggatgaaaa tttcagctgg tacctagaag acaacattaa aacctactgc	720
tcagaaccag agaaagttga caaagacaac gaagacttcc aggagagtaa cagaatgtat	780
tctgtgaatg gatacacttt tggaagtctc ccaggactct ccatgtgtgc tgaagacaga	840

- 13 -

gtaaaatggt accttttttg tatgggtaat gaagttgatg tgcacgcagc tttctttcac	900
gggcaagcac tgactaaca gaactaccgt attgacacaa tcaacctott toctgctacc	960
ctgtttgatg cttatatggt ggcccagaac cctggagaat ggatgctcag ctgtcagaat	1020
ctaaaccato tgaaagccgg tttgcaagcc tttttccagg tccaggagt taacaagtct	1080
tcatcaaagg ataatatccg tgggaagcat gttagacact actacattgc cgctgaggaa	1140
atcatctgga actatgctcc ctctggtata gacatcttca ctaaagaaaa cttaacagca	1200
cctggaagtg actcagcggg gttttttgaa caaggtacca caagaattgg aggcctttat	1260
aaaaagctgg tttatcgtga gtacacagat gcctccttca caaatcgaaa ggagagaggc	1320
cctgaagaag agcatcttgg catcctgggt cctgtcattt gggcagagggt gggagacacc	1380
atcagagtaa ccttccataa caaaggagca tatcccctca gtattgagcc gattgggggtg	1440
agattcaata agaacaacga gggcacatac tattcccaa attacaacc ccagagcaga	1500
agtgtgcctc cttcagcctc ccatgtggca cccacagaaa cattcaccta tgaatggact	1560
gtcccccagg aagtaggacc cactaatgca gatcctgtgt gtctagctaa gatgtattat	1620
tctgctgtgg atccactaa agatatattc actgggctta ttgggccaat gaaaatatgc	1680
aagaaaggaa gtttcatgac aaatgggaga cagaaagatg tagacaagga attctatttg	1740
tttctacag tatttgatga gaatgagagt ttactcctgg aagataatat tagaatgttt	1800
acaactgcac ctgatcaggt ggataaggaa gatgaagact ttcagggaatc taataaaatg	1860
cactocatga atggattcat gtatgggaat cagccgggtc tcaactatgtg caaaggagat	1920
toggctgtgt ggtacttatt cagcgccgga aatgaggccg atgtacatgg aatatacttt	1980
tcaggaaaca catatctgtg gagaggagaa cggagagaca cagcaaacct cttccctcaa	2040
acaagtotta cgctccacat gtggcctgac acagagggga cttttaatgt tgaatgcctt	2100
acaactgac attacacagg cggcatgaag caaaaatata ctgtgaacca atgcaggcgg	2160
cagtctgagg attccacott ctacctggga gagaggacat actatatcgc agcagtggag	2220
gtggaatggg attattcccc acaaaggag tgggaaaagg agctgcatca tttacaagag	2280
cagaatgttt caaatgcatt tttagataag ggagagtttt acataggctc aaagtacaag	2340
aaagttgtgt atoggcagta tactgatagc acattccgtg ttccagtgga gagaaaagct	2400
gaagaagaac atctgggaat tctaggtcca caacttcatg cagatgttgg agacaaagtc	2460
aaaattatct ttaaaaacat ggccacaagg ccctactcaa tacatgccca tgggggtacaa	2520
acagagagtt ctacagttac tccaacatta ccaggtgaaa ctctcactta cgtatggaaa	2580

- 14 -

```

atcccagaaa gatctggagc tggaacagag gattctgctt gtattccatg ggcttattat 2640
tcaactgtgg atcaagttaa ggacctctac agtggattaa ttggccccct gattgtttgt 2700
cgaagacctt acttgaaagt attcaatccc agaaggaagc tggaatttgc ccttctgttt 2760
ctagtttttg atgagaatga atcttggtac ttagatgaca acatcaaaac atactctgat 2820
caccocgaga aagtaaacia agatgatgag gaattcatag aaagcaataa aatgcatgct 2880
attaatggaa gaatgtttgg aaacctacia ggcctcacia tgcacgtggg agatgaagtc 2940
aactggtatc tgatgggaat gggcaatgaa atagacttac aactgtaca ttttcacggc 3000
catagcttcc aatacaagca caggggagtt tatagttctg atgtctttga ctttttccct 3060
ggaacatacc aaaccctaga aatgtttcca agaacacctg gaatttggtt actocactgc 3120
catgtgaocg accacattca tgctggaatg gaaaccactt acaccgttct acaaaatgaa 3180
gacaccaaat ctggctgaat gaaataaatt ggtgataagt ggaaaaaaga gaaaaaccaa 3240
tgattcataa caatgtatgt gaaagtgtaa aatagaatgt tactttggaa tgactataaa 3300
cattaaaaga gactggagca t 3321

```

<210> 8

<211> 1065

<212> PRT

<213> Homo sapiens

<400> 8

```

Met Lys Ile Leu Ile Leu Gly Ile Phe Leu Phe Leu Cys Ser Thr Pro
1           5           10           15

```

```

Ala Trp Ala Lys Glu Lys His Tyr Tyr Ile Gly Ile Ile Glu Thr Thr
20           25           30

```

```

Trp Asp Tyr Ala Ser Asp His Gly Glu Lys Lys Leu Ile Ser Val Asp
35           40           45

```

```

Thr Glu His Ser Asn Ile Tyr Leu Gln Asn Gly Pro Asp Arg Ile Gly
50           55           60

```

```

Arg Leu Tyr Lys Lys Ala Leu Tyr Leu Gln Tyr Thr Asp Glu Thr Phe
65           70           75           80

```

- 15 -

Arg Thr Thr Ile Glu Lys Pro Val Trp Leu Gly Phe Leu Gly Pro Ile
 85 90 95

Ile Lys Ala Glu Thr Gly Asp Lys Val Tyr Val His Leu Lys Asn Leu
 100 105 110

Ala Ser Arg Pro Tyr Thr Phe His Ser His Gly Ile Thr Tyr Tyr Lys
 115 120 125

Glu His Glu Gly Ala Ile Tyr Pro Asp Asn Thr Thr Asp Phe Gln Arg
 130 135 140

Ala Asp Asp Lys Val Tyr Pro Gly Glu Gln Tyr Thr Tyr Met Leu Leu
 145 150 155 160

Ala Thr Glu Glu Gln Ser Pro Gly Glu Gly Asp Gly Asn Cys Val Thr
 165 170 175

Arg Ile Tyr His Ser His Ile Asp Ala Pro Lys Asp Ile Ala Ser Gly
 180 185 190

Leu Ile Gly Pro Leu Ile Ile Cys Lys Lys Asp Ser Leu Asp Lys Glu
 195 200 205

Lys Glu Lys His Ile Asp Arg Glu Phe Val Val Met Phe Ser Val Val
 210 215 220

Asp Glu Asn Phe Ser Trp Tyr Leu Glu Asp Asn Ile Lys Thr Tyr Cys
 225 230 235 240

Ser Glu Pro Glu Lys Val Asp Lys Asp Asn Glu Asp Phe Gln Glu Ser
 245 250 255

Asn Arg Met Tyr Ser Val Asn Gly Tyr Thr Phe Gly Ser Leu Pro Gly
 260 265 270

Leu Ser Met Cys Ala Glu Asp Arg Val Lys Trp Tyr Leu Phe Gly Met
 275 280 285

Gly Asn Glu Val Asp Val His Ala Ala Phe Phe His Gly Gln Ala Leu
 290 295 300

Thr Asn Lys Asn Tyr Arg Ile Asp Thr Ile Asn Leu Phe Pro Ala Thr
 305 310 315 320

- 16 -

Leu Phe Asp Ala Tyr Met Val Ala Gln Asn Pro Gly Glu Trp Met Leu
 325 330 335

Ser Cys Gln Asn Leu Asn His Leu Lys Ala Gly Leu Gln Ala Phe Phe
 340 345 350

Gln Val Gln Glu Cys Asn Lys Ser Ser Ser Lys Asp Asn Ile Arg Gly
 355 360 365

Lys His Val Arg His Tyr Tyr Ile Ala Ala Glu Glu Ile Ile Trp Asn
 370 375 380

Tyr Ala Pro Ser Gly Ile Asp Ile Phe Thr Lys Glu Asn Leu Thr Ala
 385 390 395 400

Pro Gly Ser Asp Ser Ala Val Phe Phe Glu Gln Gly Thr Thr Arg Ile
 405 410 415

Gly Gly Ser Tyr Lys Lys Leu Val Tyr Arg Glu Tyr Thr Asp Ala Ser
 420 425 430

Phe Thr Asn Arg Lys Glu Arg Gly Pro Glu Glu Glu His Leu Gly Ile
 435 440 445

Leu Gly Pro Val Ile Trp Ala Glu Val Gly Asp Thr Ile Arg Val Thr
 450 455 460

Phe His Asn Lys Gly Ala Tyr Pro Leu Ser Ile Glu Pro Ile Gly Val
 465 470 475 480

Arg Phe Asn Lys Asn Asn Glu Gly Thr Tyr Tyr Ser Pro Asn Tyr Asn
 485 490 495

Pro Gln Ser Arg Ser Val Pro Pro Ser Ala Ser His Val Ala Pro Thr
 500 505 510

Glu Thr Phe Thr Tyr Glu Trp Thr Val Pro Lys Glu Val Gly Pro Thr
 515 520 525

Asn Ala Asp Pro Val Cys Leu Ala Lys Met Tyr Tyr Ser Ala Val Asp
 530 535 540

Pro Thr Lys Asp Ile Phe Thr Gly Leu Ile Gly Pro Met Lys Ile Cys
 545 550 555 560

- 17 -

Lys Lys Gly Ser Leu His Ala Asn Gly Arg Gln Lys Asp Val Asp Lys
 565 570 575

Glu Phe Tyr Leu Phe Pro Thr Val Phe Asp Glu Asn Glu Ser Leu Leu
 580 585 590

Leu Glu Asp Asn Ile Arg Met Phe Thr Thr Ala Pro Asp Gln Val Asp
 595 600 605

Lys Glu Asp Glu Asp Phe Gln Glu Ser Asn Lys Met His Ser Met Asn
 610 615 620

Gly Phe Met Tyr Gly Asn Gln Pro Gly Leu Thr Met Cys Lys Gly Asp
 625 630 635 640

Ser Val Val Trp Tyr Leu Phe Ser Ala Gly Asn Glu Ala Asp Val His
 645 650 655

Gly Ile Tyr Phe Ser Gly Asn Thr Tyr Leu Trp Arg Gly Glu Arg Arg
 660 665 670

Asp Thr Ala Asn Leu Phe Pro Gln Thr Ser Leu Thr Leu His Met Trp
 675 680 685

Pro Asp Thr Glu Gly Thr Phe Asn Val Glu Cys Leu Thr Thr Asp His
 690 695 700

Tyr Thr Gly Gly Met Lys Gln Lys Tyr Thr Val Asn Gln Cys Arg Arg
 705 710 715 720

Gln Ser Glu Asp Ser Thr Phe Tyr Leu Gly Glu Arg Thr Tyr Tyr Ile
 725 730 735

Ala Ala Val Glu Val Glu Trp Asp Tyr Ser Pro Gln Arg Glu Trp Glu
 740 745 750

Lys Glu Leu His His Leu Gln Glu Gln Asn Val Ser Asn Ala Phe Leu
 755 760 765

Asp Lys Gly Glu Phe Tyr Ile Gly Ser Lys Tyr Lys Lys Val Val Tyr
 770 775 780

Arg Gln Tyr Thr Asp Ser Thr Phe Arg Val Pro Val Glu Arg Lys Ala

- 18 -

785		790		795		800									
Glu	Glu	Glu	His	Leu	Gly	Ile	Leu	Gly	Pro	Gln	Leu	His	Ala	Asp	Val
				805					810					815	
Gly	Asp	Lys	Val	Lys	Ile	Ile	Phe	Lys	Asn	Met	Ala	Thr	Arg	Pro	Tyr
			820					825					830		
Ser	Ile	His	Ala	His	Gly	Val	Gln	Thr	Glu	Ser	Ser	Thr	Val	Thr	Pro
		835					840					845			
Thr	Leu	Pro	Gly	Glu	Thr	Leu	Thr	Tyr	Val	Trp	Lys	Ile	Pro	Glu	Arg
	850					855					860				
Ser	Gly	Ala	Gly	Thr	Glu	Asp	Ser	Ala	Cys	Ile	Pro	Trp	Ala	Tyr	Tyr
865					870					875					880
Ser	Thr	Val	Asp	Gln	Val	Lys	Asp	Leu	Tyr	Ser	Gly	Leu	Ile	Gly	Pro
				885					890					895	
Leu	Ile	Val	Cys	Arg	Arg	Pro	Tyr	Leu	Lys	Val	Phe	Asn	Pro	Arg	Arg
			900					905					910		
Lys	Leu	Glu	Phe	Ala	Leu	Leu	Phe	Leu	Val	Phe	Asp	Glu	Asn	Glu	Ser
		915					920					925			
Trp	Tyr	Leu	Asp	Asp	Asn	Ile	Lys	Thr	Tyr	Ser	Asp	His	Pro	Glu	Lys
	930					935					940				
Val	Asn	Lys	Asp	Asp	Glu	Glu	Phe	Ile	Glu	Ser	Asn	Lys	Met	His	Ala
945					950					955					960
Ile	Asn	Gly	Arg	Met	Phe	Gly	Asn	Leu	Gln	Gly	Leu	Thr	Met	His	Val
				965					970					975	
Gly	Asp	Glu	Val	Asn	Trp	Tyr	Leu	Met	Gly	Met	Gly	Asn	Glu	Ile	Asp
			980					985					990		
Leu	His	Thr	Val	His	Phe	His	Gly	His	Ser	Phe	Gln	Tyr	Lys	His	Arg
		995					1000					1005			
Gly	Val	Tyr	Ser	Ser	Asp	Val	Phe	Asp	Ile	Phe	Pro	Gly	Thr	Tyr	
	1010					1015					1020				

- 19 -

Gln Thr Leu Glu Met Phe Pro Arg Thr Pro Gly Ile Trp Leu Leu
 1025 1030 1035

His Cys His Val Thr Asp His Ile His Ala Gly Met Glu Thr Thr
 1040 1045 1050

Tyr Thr Val Leu Gln Asn Glu Asp Thr Lys Ser Gly
 1055 1060 1065

<210> 9

<211> 3321

<212> DNA

<213> Homo sapiens

<400> 9

```

atgaagattt tgatacttgg tatttttctg tttttatgta gtaccccagc ctgggcgaaa      60
gaaaagcatt attacattgg aattattgaa acgacttggg attatgocctc tgaccatggg      120
gaaaagaaac ttatttctgt tgacacggaa cattccaata totatcttca aaatggccca      180
gatagaattg ggagactata taagaaggcc ctttatcttc agtacacaga tgaaaccttt      240
aggacaaacta tagaaaaacc ggtctggcct gggtttttag gccctattat caaagctgaa      300
actggagata aagtttatgt aactttaaaa aaccttgcct ctaggcccta cacctttcat      360
tcacatggaa taacttacta taaggaacat gagggggcca tctaccctga taacaccaca      420
gattttcaaa gagcagatga caaagtatat ccaggagagc agtatacata catgttgctt      480
gccactgaag aacaaagtcc tggggaagga gatggcaatt gtgtgactag gatttaccat      540
tcccacattg atgctccaaa agatattgcc tcaggactca tcggaccttt aataatctgt      600
aaaaaagatt ctctagataa agaaaaagaa aaacatattg accgagaatt tgtggtgatg      660
ttttctgtgg tggatgaaaa tttcagctgg tacctagaag acaacattaa aacctactgc      720
tcagaaccag agaaagttga caaagacaac gaagacttcc aggagagtaa cagaatgtat      780
totgtgaatg gatacacttt tggaagtctc ccaggactct ccatgtgtgc tgaagacaga      840
gtaaaatggt accttttttg tatgggtaat gaagttgatg tgcacgcagc tttctttcac      900
gggcaagcac tgactaacia gaactaccgt attgacacaa tcaacctctt tcctgctacc      960
ctgtttgatg cttatatggt ggcccagaac cctggagaat ggatgctcag ctgtcagaat     1020
ctaaaccatc tgaaagccgg tttgcaagcc tttttccagg tccaggagtg taacaagtct     1080

```

- 20 -

tcatacaagg	ataatatccg	tgggaagcat	gtagacact	actacattgc	cgctgaggaa	1140
atcatctgga	actatgctcc	ctctggtata	gacatcttca	ctaaagaaaa	cttaacagca	1200
cctggaagtg	actcagcggg	gttttttgaa	caaggtagca	caagaattgg	aggctcttat	1260
aaaaagctgg	tttatcgtga	gtacacagat	gcctccttca	caaatcgaaa	ggagagaggc	1320
cctgaagaag	agcatcttgg	catcctgggt	cctgtcattt	gggcagaggt	gggagacacc	1380
atcagagtaa	ccttcataa	caaaggagca	tatccctca	gtattgagcc	gattgggggtg	1440
agattcaata	agaacaacga	gggcacatac	tattcccca	attacaaccc	ccagagcaga	1500
agtgtgcctc	cttcagcctc	ccatgtggca	cccacagaaa	cattcaccta	tgaatggact	1560
gtcccaaaag	aagtaggacc	cactaatgca	gatcctgtgt	gtctagctaa	gatgtattat	1620
tctgctgtgg	atccactaa	agatatattc	actgggctta	ttggggccaat	gaaaatatgc	1680
aagaaaggaa	gtttacatgc	aatgggaga	cagaaagatg	tagacaagga	attctatttg	1740
tttctacag	tatttgatga	gaatgagagt	ttactcctgg	aagataatat	tagaatgttt	1800
acaactgcac	ctgatcaggt	ggataaggaa	gatgaagact	ttcaggaatc	taataaaatg	1860
cactocatga	atggattcat	gtatgggaat	cagccgggtc	tcactatgtg	caaaggagat	1920
tgggtogtgt	ggtacttatt	cagcgccgga	aatgaggccg	atgtacatgg	aatatacttt	1980
tcaggaaaca	catatctgtg	gagaggagaa	cggagagaca	cagcaaacct	cttcctcaa	2040
acaagtctta	cgctccacat	gtggcctgac	acagagggga	cttttaatgt	tgaatgcott	2100
acaactgac	attacacagg	cggcatgaag	caaaaatata	ctgtgaacca	atgcaggogg	2160
cagtctgagg	attccacctt	ctacctggga	gagaggacat	actatatogc	agcagtggag	2220
gtggaatggg	attattcccc	acaaaggag	tgggaaaagg	agctgcatca	tttacaagag	2280
cagaatgttt	caaatgcatt	tttagataag	ggagagtgtt	acataggctc	aaagtacaag	2340
aaagtgtgt	atcggcagta	tactgatagc	acattccgtg	ttccagtggg	gagaaaagct	2400
gaagaagaac	atctgggaat	tctaggtcca	caacttcatt	cagatgttgg	agacaaagtc	2460
aaaattatct	ttaaaaacat	ggccacaagg	ccctactcaa	tacatgcccc	tggggtacaa	2520
acagagagtt	ctacagttac	tccaacatta	ccaggtgaaa	ctctcaactta	cgtatggaaa	2580
atoccagaaa	gatctggagc	tggaacagag	gattctgctt	gtattccatt	ggcttattat	2640
tcaactgtgg	atcaagttaa	ggacctctac	agtggattaa	ttggccccct	gattgtttgt	2700
cgaagacctt	acttgaaagt	attcaatccc	agaagggaagc	tggaaatttg	ccttctgttt	2760
ctagtttttg	atgagaatga	atcttggtac	ttagatgaca	acatcaaaac	atactctgat	2820
caccccgaga	aagtaaaca	agatgatgag	gaattcatag	aaagcaataa	aatgcatgct	2880

- 21 -

attaatggaa gaatgttttg aaacctacaa ggcctcacia tgcacgtggg agatgaagtc	2940
aactggtatc tgatgggaat gggcaatgaa atagacttac aactgtaca ttttcacggc	3000
catagcttcc aatacaagca caggggagtt tatagttctg atgtctttga cattttccct	3060
ggaacatacc aaacctaga aatgtttcca agaacacctg gaatttggtt actccactgc	3120
catgtgaccg accacattca tgctggaatg gaaaccactt acaccgttct acaaaatgaa	3180
gacaccaa atctggctgaat gaaataaatt ggtgataagt ggaaaaaaga gaaaaaccaa	3240
tgattcataa caatgtatgt gaaagtgtaa aatagaatgt tactttggaa tgactataaa	3300
cattaaaaga gactggagca t	3321

<210> 10

<211> 4215

<212> DNA

<213> Homo sapiens

<400> 10

cctgtttccc agagtaatgt gggccatgga gtcaggccac ctctctctgg ctctgctggt	60
catgcagtcc ttgtggcctc aactgactga tggagccact cgagtctact acctgggcat	120
ccgggatgtg cagtggaaact atgctcccaa ggaagaaat gtcacacga accagcctct	180
ggacagtgac atagtggctt ccagcttctt aaagtctgac aagaaccgga tagggggaac	240
ctacaagaag accatctata aagaatacaa ggatgactca tacacagatg aagtggccca	300
gcctgcctgg ttgggcttcc tggggccagt gttgcaggct gaagtggggg atgtcattct	360
tattcacctg aagaattttg ccactcgtcc ctataccatc caccctcatg gtgtcttcta	420
cgagaaggac tctgaagggt cctataccc agatggctcc tctgggccac tgaaagctga	480
tgactctggt ccccgggggg gcagccatat ctacaactgg accattccag aaggccatgc	540
accacccgat gctgaccocag cgtgcctcac ctggatctac cattctcatg tagatgctcc	600
acgagacatt gcaactggcc taattgggccc tctcatcacc tgtaaaagag gagccctgga	660
tgggaactcc cctcctcaac gccaggatgt agaccatgat ttcttctcc tcttcagtgt	720
ggtagatgag aacctcagct ggcattctca tgagaacatt gccacttact gctcagatcc	780
tgcttcagtg gacaaagaag atgagacatt tcaggagagc aataggatgc atgcaatcaa	840
tggtcttggt tttgggaatt tacctgagct gaacatgtgt gcacagaaac gtgtggcctg	900

- 22 -

gcacttgttt ggcattgggca atgaaattga tgtccacaca gcattttttcc atggacagat	960
gctgactacc cgtggacacc aactgatgt ggctaacatc tttccagcca ctttgtgac	1020
tgctgagatg gtgccctggg aacctggtac ctggttaatt agctgccaaag tgaacagtca	1080
ctttogagat ggcattgcagg cactctacaa ggtcaagtct tgctccatgg cccctcctgt	1140
ggacctgctc acaggcaaag ttcgacagta cttcattgag gcccatgaga ttcaatggga	1200
ctatggcccg atggggcatg atgggagtac tgggaagaat ttgagagagc caggcagtat	1260
ctcagataag tttttccaga agagctccag ccgaattggg ggcacttaact ggaaagtgcg	1320
atatgaagcc tttcaagatg agacattcca agagaagatg catttgaggg aagataggca	1380
tcttggaatc ctggggccag tgatccgggc tgaggtgggt gacaccatc aggtggtctt	1440
ctacaacogt gcctcccagc cattcagcat gcagcccat ggggtctttt atgagaaaga	1500
ctatgaaggc actgtgtaca atgatggctc atcttaccct ggcttggttg ccaagccctt	1560
tgagaaagta acatacogt ggacagtccc ccctcatgcc ggtccactg ctcaggatcc	1620
tgcttgtctc acttggatgt acttctctgc tgcagatccc ataagagaca caaattctgg	1680
cctggtgggc ccgctgctgg tgtgcagggc tggtgccctg ggtgcagatg gcaagcagaa	1740
aggggtggat aaagaattct ttcttctctt cactgtgttg gatgagaaca agagctggta	1800
cagcaatgcc aatcaagcag ctgctatgtt ggatttccga ctgctttcag aggatattga	1860
gggcttccaa gactccaatc ggatgcatgc cattaatggg tttctgttct ctaacctgcc	1920
caggctggac atgtgcaagg gtgacacagt ggcctggcac ctgctcggcc tgggcacaga	1980
gactgatgtg catggagtca tgttccaggg caacactgtg cagcttcagg gcatgaggaa	2040
gggtgcagct atgctotttc ctcatacctt tgtcatggcc atcatgcagc ctgacaacct	2100
tgggacattt gagatttatt gccaggcagg cagccatcga gaagcaggga tgagggcaat	2160
ctataatgtc tccagtgtc ctggccacca agccaccct cgccaacgct accaagctgc	2220
aagaatctac tatatcatgg cagaagaagt agagtgggac tattgccctg accggagctg	2280
ggaacgggaa tggcacaacc agtctgagaa ggacagttat ggttacattt toctgagcaa	2340
caaggatggg ctctggggtt ccagatacaa gaaagctgta ttcagggaat aactgatgg	2400
tacattcagg atccctoggc caaggactgg accagaagaa cacttgggaa tottgggtoc	2460
acttatcaaa ggtgaagttg gtgatatcct gactgtggta ttcaagaata atgccagccg	2520
cccctactct gtgcatgctc atggagtgtc agaatctact actgtctggc cactggctgc	2580
tgagcctggg gaggtggtca cttatcagtg gaacatccca gagaggtctg gccctgggac	2640
caatgactct gcttgtgttt cctggatcta ttattctgca gtggatccca tcaaggacat	2700

- 23 -

```

gtatagtggc ctggtggggc ccttggctat ctgccaaaag ggcatoctgg agcccatgg 2760
aggacggagt gacatggatc gggaatttgc attgttgttc ttgatttttg atgaaaataa 2820
gtcttgggtat ttggaggaaa atgtggcaac ccatgggtcc caggatccag gcagtattaa 2880
cctacaggat gaaactttct tggagagcaa taaaatgcat gcaatcaatg ggaaactcta 2940
tgccaacctt aggggtctta ccatgtacca aggagaacga gtggcctggc acatgctggc 3000
catgggccaa gatgtggatc tacacaccat ccactttcat gcagagagct tcctotatcg 3060
gaatggcgag aactaccggg cagatgtggc ggatctgttc ccagggactt ttgaggttgt 3120
ggagatggtg gccagcaacc ctgggacatg gctgatgcac tgccatgtga ctgacctgt 3180
ccatgctggc atggagaccc tcttctactgt tttttctcga acagaacact taagccctot 3240
cacgctcatc accaaagaga ctgaaaaagc agtgccccc agagacattg aagaaggcaa 3300
tgtgaagatg ctgggcatgc agatcccat aaagaatgtt gagatgctgg cctctgtttt 3360
ggttgccatt agtgtcacc ttctgctcgt tgttctggct cttggtggag tggtttggt 3420
ccaacatcga cagagaaagc tacgacgcaa taggaggtcc atcctggatg acagcttcaa 3480
gcttctgtct ttc aaacagt aacatctgga gcctggagat atcctcagga agcacatctg 3540
tagtgcactc ccagcaggcc atggactagt cactaaccac aactcaaag gggcatgggt 3600
ggtggagaag cagaaggagc aatcaagctt atctggatat ttctttcttt atttatttta 3660
catggaaata atatgatttc actttttctt tagtttcttt gctctacgtg ggcacctggc 3720
actaaggag taccttatta tcctacatcg caaatttcaa cagctacatt atatttctt 3780
ctgacacttg gaaggattg aaatttctag aaatgtatcc ttctcacaaa gtagagacca 3840
agagaaaaac tcattgattg ggtttctact tctttcaagg actcaggaaa tttcactttg 3900
aactgaggcc aagtgagctg ttaagataac ccacacttaa actaaaggct aagaatatag 3960
gcttgatggg aaattgaagg taggctgagt attgggaatc caaattgaat tttgattctc 4020
cttggcagtg aactactttg aagaagtggc caatgggttg ttgctgcoat gagcatgtac 4080
aacctctgga gctagaagct cctcaggaaa gccagttctc caagttctta acctgtggca 4140
ctgaaaggaa tgttgagtta cctcttcatg ttttagacag caaacctat ccattaaagt 4200
acttgttaga acact 4215

```

<210> 11

<211> 1158

- 24 -

<212> PRT

<213> Homo sapiens

<400> 11

Met Glu Ser Gly His Leu Leu Trp Ala Leu Leu Phe Met Gln Ser Leu
 1 5 10 15

Trp Pro Gln Leu Thr Asp Gly Ala Thr Arg Val Tyr Tyr Leu Gly Ile
 20 25 30

Arg Asp Val Gln Trp Asn Tyr Ala Pro Lys Gly Arg Asn Val Ile Thr
 35 40 45

Asn Gln Pro Leu Asp Ser Asp Ile Val Ala Ser Ser Phe Leu Lys Ser
 50 55 60

Asp Lys Asn Arg Ile Gly Gly Thr Tyr Lys Lys Thr Ile Tyr Lys Glu
 65 70 75 80

Tyr Lys Asp Asp Ser Tyr Thr Asp Glu Val Ala Gln Pro Ala Trp Leu
 85 90 95

Gly Phe Leu Gly Pro Val Leu Gln Ala Glu Val Gly Asp Val Ile Leu
 100 105 110

Ile His Leu Lys Asn Phe Ala Thr Arg Pro Tyr Thr Ile His Pro His
 115 120 125

Gly Val Phe Tyr Glu Lys Asp Ser Glu Gly Ser Leu Tyr Pro Asp Gly
 130 135 140

Ser Ser Gly Pro Leu Lys Ala Asp Asp Ser Val Pro Pro Gly Gly Ser
 145 150 155 160

His Ile Tyr Asn Trp Thr Ile Pro Glu Gly His Ala Pro Thr Asp Ala
 165 170 175

Asp Pro Ala Cys Leu Thr Trp Ile Tyr His Ser His Val Asp Ala Pro
 180 185 190

Arg Asp Ile Ala Thr Gly Leu Ile Gly Pro Leu Ile Thr Cys Lys Arg
 195 200 205

- 25 -

Gly	Ala	Leu	Asp	Gly	Asn	Ser	Pro	Pro	Gln	Arg	Gln	Asp	Val	Asp	His
210						215					220				
Asp	Phe	Phe	Leu	Leu	Phe	Ser	Val	Val	Asp	Glu	Asn	Leu	Ser	Trp	His
225					230					235					240
Leu	Asn	Glu	Asn	Ile	Ala	Thr	Tyr	Cys	Ser	Asp	Pro	Ala	Ser	Val	Asp
				245					250					255	
Lys	Glu	Asp	Glu	Thr	Phe	Gln	Glu	Ser	Asn	Arg	Met	His	Ala	Ile	Asn
			260					265						270	
Gly	Phe	Val	Phe	Gly	Asn	Leu	Pro	Glu	Leu	Asn	Met	Cys	Ala	Gln	Lys
		275					280					285			
Arg	Val	Ala	Trp	His	Leu	Phe	Gly	Met	Gly	Asn	Glu	Ile	Asp	Val	His
	290					295					300				
Thr	Ala	Phe	Phe	His	Gly	Gln	Met	Leu	Thr	Thr	Arg	Gly	His	His	Thr
305					310					315					320
Asp	Val	Ala	Asn	Ile	Phe	Pro	Ala	Thr	Phe	Val	Thr	Ala	Glu	Met	Val
				325						330				335	
Pro	Trp	Glu	Pro	Gly	Thr	Trp	Leu	Ile	Ser	Cys	Gln	Val	Asn	Ser	His
			340					345						350	
Phe	Arg	Asp	Gly	Met	Gln	Ala	Leu	Tyr	Lys	Val	Lys	Ser	Cys	Ser	Met
		355					360						365		
Ala	Pro	Pro	Val	Asp	Leu	Leu	Thr	Gly	Lys	Val	Arg	Gln	Tyr	Phe	Ile
	370					375					380				
Glu	Ala	His	Glu	Ile	Gln	Trp	Asp	Tyr	Gly	Pro	Met	Gly	His	Asp	Gly
385					390					395					400
Ser	Thr	Gly	Lys	Asn	Leu	Arg	Glu	Pro	Gly	Ser	Ile	Ser	Asp	Lys	Phe
				405					410					415	
Phe	Gln	Lys	Ser	Ser	Ser	Arg	Ile	Gly	Gly	Thr	Tyr	Trp	Lys	Val	Arg
			420					425					430		
Tyr	Glu	Ala	Phe	Gln	Asp	Glu	Thr	Phe	Gln	Glu	Lys	Met	His	Leu	Glu
		435					440					445			

- 26 -

Glu Asp Arg His Leu Gly Ile Leu Gly Pro Val Ile Arg Ala Glu Val
 450 455 460

Gly Asp Thr Ile Gln Val Val Phe Tyr Asn Arg Ala Ser Gln Pro Phe
 465 470 475 480

Ser Met Gln Pro His Gly Val Phe Tyr Glu Lys Asp Tyr Glu Gly Thr
 485 490 495

Val Tyr Asn Asp Gly Ser Ser Tyr Pro Gly Leu Val Ala Lys Pro Phe
 500 505 510

Glu Lys Val Thr Tyr Arg Trp Thr Val Pro Pro His Ala Gly Pro Thr
 515 520 525

Ala Gln Asp Pro Ala Cys Leu Thr Trp Met Tyr Phe Ser Ala Ala Asp
 530 535 540

Pro Ile Arg Asp Thr Asn Ser Gly Leu Val Gly Pro Leu Leu Val Cys
 545 550 555 560

Arg Ala Gly Ala Leu Gly Ala Asp Gly Lys Gln Lys Gly Val Asp Lys
 565 570 575

Glu Phe Phe Leu Leu Phe Thr Val Leu Asp Glu Asn Lys Ser Trp Tyr
 580 585 590

Ser Asn Ala Asn Gln Ala Ala Ala Met Leu Asp Phe Arg Leu Leu Ser
 595 600 605

Glu Asp Ile Glu Gly Phe Gln Asp Ser Asn Arg Met His Ala Ile Asn
 610 615 620

Gly Phe Leu Phe Ser Asn Leu Pro Arg Leu Asp Met Cys Lys Gly Asp
 625 630 635 640

Thr Val Ala Trp His Leu Leu Gly Leu Gly Thr Glu Thr Asp Val His
 645 650 655

Gly Val Met Phe Gln Gly Asn Thr Val Gln Leu Gln Gly Met Arg Lys
 660 665 670

Gly Ala Ala Met Leu Phe Pro His Thr Phe Val Met Ala Ile Met Gln

- 27 -

675	680	685
Pro Asp Asn Leu Gly Thr Phe Glu Ile Tyr Cys Gln Ala Gly Ser His 690 695 700		
Arg Glu Ala Gly Met Arg Ala Ile Tyr Asn Val Ser Gln Cys Pro Gly 705 710 715 720		
His Gln Ala Thr Pro Arg Gln Arg Tyr Gln Ala Ala Arg Ile Tyr Tyr 725 730 735		
Ile Met Ala Glu Glu Val Glu Trp Asp Tyr Cys Pro Asp Arg Ser Trp 740 745 750		
Glu Arg Glu Trp His Asn Gln Ser Glu Lys Asp Ser Tyr Gly Tyr Ile 755 760 765		
Phe Leu Ser Asn Lys Asp Gly Leu Leu Gly Ser Arg Tyr Lys Lys Ala 770 775 780		
Val Phe Arg Glu Tyr Thr Asp Gly Thr Phe Arg Ile Pro Arg Pro Arg 785 790 795 800		
Thr Gly Pro Glu Glu His Leu Gly Ile Leu Gly Pro Leu Ile Lys Gly 805 810 815		
Glu Val Gly Asp Ile Leu Thr Val Val Phe Lys Asn Asn Ala Ser Arg 820 825 830		
Pro Tyr Ser Val His Ala His Gly Val Leu Glu Ser Thr Thr Val Trp 835 840 845		
Pro Leu Ala Ala Glu Pro Gly Glu Val Val Thr Tyr Gln Trp Asn Ile 850 855 860		
Pro Glu Arg Ser Gly Pro Gly Pro Asn Asp Ser Ala Cys Val Ser Trp 865 870 875 880		
Ile Tyr Tyr Ser Ala Val Asp Pro Ile Lys Asp Met Tyr Ser Gly Leu 885 890 895		
Val Gly Pro Leu Ala Ile Cys Gln Lys Gly Ile Leu Glu Pro His Gly 900 905 910		
Gly Arg Ser Asp Met Asp Arg Glu Phe Ala Leu Leu Phe Leu Ile Phe		

- 28 -

915	920	925
Asp Glu Asn Lys Ser Trp Tyr Leu Glu Glu Asn Val Ala Thr His Gly		
930	935	940
Ser Gln Asp Pro Gly Ser Ile Asn Leu Gln Asp Glu Thr Phe Leu Glu		
945	950	955 960
Ser Asn Lys Met His Ala Ile Asn Gly Lys Leu Tyr Ala Asn Leu Arg		
	965	970 975
Gly Leu Thr Met Tyr Gln Gly Glu Arg Val Ala Trp Tyr Met Leu Ala		
	980	985 990
Met Gly Gln Asp Val Asp Leu His Thr Ile His Phe His Ala Glu Ser		
	995	1000 1005
Phe Leu Tyr Arg Asn Gly Glu Asn Tyr Arg Ala Asp Val Val Asp		
1010	1015	1020
Leu Phe Pro Gly Thr Phe Glu Val Val Glu Met Val Ala Ser Asn		
1025	1030	1035
Pro Gly Thr Trp Leu Met His Cys His Val Thr Asp His Val His		
1040	1045	1050
Ala Gly Met Glu Thr Leu Phe Thr Val Phe Ser Arg Thr Glu His		
1055	1060	1065
Leu Ser Pro Leu Thr Val Ile Thr Lys Glu Thr Glu Lys Ala Val		
1070	1075	1080
Pro Pro Arg Asp Ile Glu Glu Gly Asn Val Lys Met Leu Gly Met		
1085	1090	1095
Gln Ile Pro Ile Lys Asn Val Glu Met Leu Ala Ser Val Leu Val		
1100	1105	1110
Ala Ile Ser Val Thr Leu Leu Leu Val Val Leu Ala Leu Gly Gly		
1115	1120	1125
Val Val Trp Tyr Gln His Arg Gln Arg Lys Leu Arg Arg Asn Arg		
1130	1135	1140
Arg Ser Ile Leu Asp Asp Ser Phe Lys Leu Leu Ser Phe Lys Gln		
1145	1150	1155