

**(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE**

(11) Application No. AU 2006212813 B2

(54) Title
Compositions and methods related to soluble G-protein coupled receptors (sGPCRs)

(51) International Patent Classification(s)
C07K 14/435 (2006.01)

(21) Application No: **2006212813** (22) Date of Filing: **2006.02.08**

(87) WIPO No: **WO06/086402**

(30) Priority Data

(31) Number
60/650,866 (32) Date
2005.02.08 (33) Country
US

(43) Publication Date: **2006.08.17**
(44) Accepted Journal Date: **2012.02.02**

(71) Applicant(s)
Research Development Foundation

(72) Inventor(s)
Vale, Wylie;Perrin, Marilyn;Chen, Alon

(74) Agent / Attorney
Pizzeys, Level 2, Woden Plaza Offices Woden Town Square Woden, Canberra, ACT, 2606

(56) Related Art
Grace, C. R. R. et al., PNAS, 2004, Vol. 101, No. 35, pages 12836-12841
US 5,786,203 A
Perrin, M. H. et al., The Journal of Biological Chemistry, 2003, Vol. 278, No. 18, pages 15595-15600

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 August 2006 (17.08.2006)

PCT

(10) International Publication Number
WO 2006/086402 A3

(51) International Patent Classification:
C07K 14/435 (2006.01)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US (patent), UZ, VC, VN, YU, ZA, ZM, ZW.

(21) International Application Number:
PCT/US2006/004321

(22) International Filing Date: 8 February 2006 (08.02.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/650,866 8 February 2005 (08.02.2005) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:

US 60/650,866 (CIP)
Filed on 8 February 2005 (08.02.2005)

(71) Applicant (for all designated States except US): RESEARCH DEVELOPMENT FOUNDATION [US/US]; 402 North Division Street, Carson City, Nevada 89703 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): CHEN, Alon [IL/IL], PERRIN, Marilyn [US/US]; 8844 Robinhood Lane, La Jolla, California 92037 (US). VALE, Wylie [US/US]; 1643 Valdes Drive, La Jolla, California 92037 (US).

(74) Agent: LANDRUM, Charles, P.; FULBRIGHT & JAWORSKI L.L.P., Suite 2400, 600 Congress Avenue, Austin, TX 78701 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
14 December 2006

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



A3

WO 2006/086402 A3

(54) Title: COMPOSITIONS AND METHODS RELATED TO SOLUBLE G-PROTEIN COUPLED RECEPTORS(SGPCRS)

(57) Abstract: The present invention is directed to compositions and methods related to soluble G-protein coupled receptors (sG-PCR). In certain aspects the invention includes compositions and methods related to a soluble corticotropin releasing factor receptor related protein, sCRFR2, as well as its effects on CRFR signaling and interaction between CRF family ligand and CRFR receptors, including but not limited to CRFR2, CRFR1 and functional or signaling capable variants thereof.

COMPOSITIONS AND METHODS RELATED TO SOLUBLE G-PROTEIN COUPLED RECEPTORS (sGPCRs)

This application claims priority to provisiona United States Patent Application number 60/650,866, filed February 8, 2005, which is incorporated herein by reference in 5 its entirety.

The United States Government owns rights in present invention pursuant to grant number DK 26741 from the NIDDK.

I. TECHNICAL FIELD

The present invention is directed generally to method and compositions related to 10 molecular biology, neurology, and endocrinology. In certain aspects it is directed to compositions comprising and methods of using soluble G-protein coupled receptors (sGPCRs) as modulators of GPCR activity and/or modulators of the pharmacologic effects of the ligands that bind such soluble GPCRs.

II. BACKGROUND OF THE INVENTION

15 Receptors, in general, are molecular structures located in the cell membrane or within a cell that form a weak, reversible bond with an agent such as an antigen, hormone, or neurotransmitter. Each receptor is designed to bind with a specific agent(s). A specific family of receptors is the seven transmembrane ("7TM") or G-Protein-Coupled Receptor ("GPCR"). These receptors link with a Guanine Nucleotide-Binding 20 G-protein ("G-protein") in order to signal when an appropriate agent has bound the receptor. When the G-protein binds with Guanine DiPhosphate ("GDP"), the G-protein is inactive, or in an "off position." Likewise, when the G-protein binds with Guanine TriPhosphate ("GTP"), the G-protein is active, or in an "on position" whereby activation of a biological response in a cell is mediated.

25 GPCRs share a common structural motif. All these receptors have seven sequences of between 22 to 24 hydrophobic amino acids that form seven alpha helices, each of which spans the membrane (*i.e.*, transmembrane-1 (TM-1), transmebrane-2 (TM-2), etc.). The transmembrane helices are joined by strands of amino acids between transmembrane-2 and transmembrane-3, transmembrane-4 and transmembrane-5, and 30 transmembrane-6 and transmembrane-7 on the exterior, or "extracellular" side, of the cell membrane (these are referred to as "extracellular loops" or "extracellular" regions). The transmembrane helices are also joined by strands of amino acids between

transmembrane-1 and transmembrane-2, transmembrane-3 and transmembrane-4, and transmembrane-5 and transmembrane-6 on the interior, or "intracellular" side, of the cell membrane (these are referred to as "intracellular loops" or "intracellular" regions). The "carboxy" ("C") terminus of the receptor lies in the intracellular space within the cell, 5 and the "amino" ("N") terminus of the receptor lies in the extracellular space outside of the cell.

Generally, when a ligand binds with the receptor and "activates" the receptor, there is a change in the conformation of the intracellular region that allows for coupling 10 between the intracellular region and an intracellular "G-protein." It has been reported that GPCRs are "promiscuous" with respect to G-proteins, *i.e.*, that a GPCR can interact with more than one G-protein (Kenakin, 1988). Although other G-proteins exist, currently, Gq, Gs, Gi, and Go are G-proteins that have been identified. Ligand-activated GPCR coupling with the G-protein begins a signaling cascade process or signal transduction. Under normal conditions, signal transduction ultimately results in cellular 15 activation or cellular inhibition. It is thought that the third intracellular loop (IC-3) as well as the carboxy terminus of the receptor interact with the G-protein.

In general, the activity of almost every cell in the body is regulated by extracellular signals. A number of physiological events in humans as well as with a wide range of organisms use protein mediated transmembrane signaling via GPCRs. 20 Signals from a specific GPCR cause activation of a G-protein in the cell. The majority of signals are transmitted by means of GPCRs into the cell interior. There are many varying aspects of this signaling process involving multiple receptor subtypes for GPCRs and their G-protein linked counterparts as well as a variety of linked intracellular secondary messengers. The signal transduction may result in an overall or partial 25 activation or inactivation of an intracellular process or processes depending upon the proteins that are involved. Important signaling molecules or neurotransmitters which bind to GPCRs include, but are not limited to corticotropin releasing factor, parathyroid hormone, morphine, dopamine, histamine, 5-hydroxytryptamine, adenosine, calcitonin, gastric inhibitory peptide (GIP), glucagon, growth hormone-releasing hormone (GHRH), 30 parathyroid hormone (PTH), PACAP, secretin, vasoactive intestinal polypeptide (VIP), diuretic hormone, EMR1, latrophilin, brain-specific angiogenesis inhibitor (BAI), cadherin, EGF, LAG, (CELSR), and other similar proteins or molecules.

5 GPCRs constitute a superfamily of proteins. There are currently over 2000 GPCRs reported in literature, which are divided into at least three families: rhodopsin-like family (family A), the calcitonin receptors (family B), and metabotropic glutamate family (family C) (Ji *et al.*, 1998). The reported GPCRs include both characterized receptors and orphan receptors for which ligands have not yet been identified. (Wilson *et al.*, 1999; Wilson *et al.*, 1998; Marchese *et al.*, 1999). Despite the large number of GPCRs, generally, each GPCR share a similar molecular structure. Each GPCR comprises a string of amino acid residues of various lengths. GPCRs lie within the cell membrane in seven distinct coils called transmembranes. The amino terminus of the 10 GPCR lies outside the cell with the extracellular loops, while the carboxy-terminus lies inside the cell with the intracellular loops.

15 The ligands for GPCRs comprise small molecules as well as peptides and small proteins. The interactions between these ligands and their receptors vary from system to system but they may require the interaction with residues in several of the four extracellular domains and the N-terminus. In some instances the N-terminus alone may 20 maintain an ability to interact with or bind ligands. GPCRs with known ligands have been associated with many diseases including multiple sclerosis, diabetes, rheumatoid arthritis, asthma, allergies, inflammatory bowel disease, several cancers, thyroid disorders, heart disease, retinitis pigmentosa, obesity, neurological disorders, osteoporosis, Human Immunodeficiency Virus ("HIV") infection and Acquired Immune Deficiency Syndrome ("AIDS") (Murphy *et al.*, 2000; Mannstadt *et al.*, 1999; Berger *et al.*, 1999; Jacobson *et al.*, 1997; Meij, 1996;).

25 Accordingly, there is a need in the art for methods of producing modulators of GPCRs and the ligands that bind GPCRs for use as therapeutics. These therapeutics may be used to prevent or treat GPCR associated diseases and/or disorders.

BRIEF SUMMARY OF THE INVENTION

The present invention is directed to compositions and methods related to a sGPCR ligand binding domains, as well as effects of the sGPCR on GPCR signaling and interaction between GPCR ligands and their GPCRs.

30 An embodiment of the invention includes an isolated soluble G-protein coupled receptor (sGPCR) ligand binding domain. A sGPCR comprises all or part of a GPCR extracellular domain. In one aspect of the invention the sGPCR is an soluble form of a

5 GPCR family B member. In a further aspect the sGPCR is a GPCR subfamily B1 member. In still further aspects, a sGPCR is a soluble secretin receptor, VPAC₁ receptor, VPAC₂ receptor, PAC₁ receptor, glucagon receptor, growth hormone releasing hormone (GHRH) receptor, glucagon-related peptide 1 (GLP-1) receptor, glucagon-
related peptide 2 (GLP-2) receptor, gastric inhibitory polypeptide (GIP) receptor, 10 corticotropin releasing factor 1 (CRF1) receptor, corticotropin releasing factor 2 (CRF2) receptor, parathyroid hormone 1 (PTH1) receptor, parathyroid hormone 2 (PTH2) receptor, calcitonin receptor-like receptor, or calcitonin receptor. The sGPCR can be a soluble PTH1 receptor or PTH2 receptor. An embodiment of the invention also includes
15 a sGPCR that is a soluble form of the corticotropin releasing factor receptor type 2α (sCRFR2α). The amino acid sequence of a sCRFR2α may comprise an amino acid sequence encoded by exons 3, 4, and 5 of the CRFR2α gene or does not contain exon 6 or greater. A recombinant sGPCR of the invention may include 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 130, 135, 140, 150, 155, 160, 180, 200 or more amino acids,
15 including all ranges there between, of an GPCR extracellular domain(s), including all or part of the amino terminal extracellular domain. In certain aspects, a sGPCR ligand binding domain may comprise an amino acid sequence at least 70, 75, 80, 85, 90, 95, or 98% similar to 50, 75, 100, 125, 150 or more amino acids of SEQ ID NO:4 (sCRFR2α), SEQ ID NO:8 (sCRFR2β), SEQ ID NO:12 (sCRFR2γ), or SEQ ID NO:15 (mCRFR2α).
20 In a further aspect, a sCRFR comprises the amino acid sequence of SEQ ID NO:4, 8, 12, 15 or a combination thereof. In a still further aspect, the invention includes an isolated sGPCR further comprising an affinity tag, a label, a detectable or therapeutic chemical moiety, a biotin/avidin label, a radionuclide, a detectable or therapeutic enzyme, a fluorescent marker, a chemiluminescent marker, an immunoglobulin domain or any
25 combination thereof. In one aspect, the GPCR comprises an immunoglobulin domain, in particular an Fc domain. The sGPCR can be conjugated to a polymer, which includes, but is not limited to polyethylene glycol (PEG).

30 Embodiments of the invention include polynucleotides encoding sGPCR of the invention. Polynucleotide may further comprise a promoter operably coupled to the polynucleotide encoding the sGPCR. The sGPCR encoding sequence can be included in an expression cassette. The expression cassette may be comprised in an expression vector. The expression vector may include, but is not limited to a linear nucleic acid, a plasmid expression vector, or a viral expression vector. In certain aspects, an expression

vector is comprised in a delivery vector, which may include, but is not limited to a liposome, a polypeptide, a polycation, a lipid, a bacterium, or a virus.

Still further embodiments of the invention include methods of modulating the activity of G-protein coupled receptor (GPCR) comprising: a) contacting a target tissue 5 with a sGPCR; and b) binding a GPCR ligand in the vicinity of the target tissue, wherein the activity of the GPCR in the tissue is modulated. The ligand can be a GPCR family B ligand, a GPCR subfamily B1 ligand. In certain aspects the ligand is a corticotropin releasing factor (CRF), urocortin 1, urocortin 2, usorcortin 3, stresscortin, parathyroid hormone, PTH-related hormone, TIP39, calcitonin, amylin, CGRP (CALCA 10 and CALCB), adrenomedullin, secretin, VIP, PACAP, glucagon, GHRH, GLP-1, GLP-2, GIP or any combination thereof. The methods may also include contacting a target tissue comprising the steps of: a) preparing sGPCR ligand binding domain in an appropriate pharmaceutical solution; and b) administering the pharmaceutical solution to an animal, human, subject, and/or patient in an amount to affect binding of a target 15 ligand in the target tissue of the animal. Administration can be, but is not limited to ingestion, injection, endoscopy or perfusion. Injection includes, but is not limited to intravenous, intramuscular, subcutaneous, intradermal, intracranial or intraperitoneal injection. Disorders that may be treated, ameliorated, modulation, reduced in severity, include disorders resulting from hyperactivation of a GPCR or hypersecretion of GPCR 20 ligand. In certain aspects the disorder is insulin sensitivity or type II diabetes. The disorder may also include an anxiety-related disorder; a mood disorder; a post-traumatic stress disorder; supranuclear palsy; immune suppression; drug or alcohol withdrawal symptoms; inflammatory disorders; pain; asthma; psoriasis and allergies; phobias; sleep disorders induced by stress; fibromyalgia; dysthemia; bipolar disorders; cyclothymia; 25 fatigue syndrome; stress-induced headache; cancer; human immunodeficiency virus infections; neurodegenerative diseases; gastrointestinal diseases; eating disorders; hemorrhagic stress; stress-induced psychotic episodes; euthyroid sick syndrome; syndrome of inappropriate antidiarrhetic hormone; obesity; infertility; head traumas; spinal cord trauma; ischemic neuronal damage; excitotoxic neuronal damage; epilepsy; 30 cardiovascular and heart related disorders; immune dysfunctions; muscular spasms; urinary incontinence; senile dementia of the Alzheimer's type; multiinfarct dementia; amyotrophic lateral sclerosis; chemical dependencies and addictions; psychosocial dwarfism, insulin hypersensitivity or hyposensitivity, hypoglycemia, skin disorders; or

hair loss. In certain aspects the disorder is an anxiety-related disorder; a mood disorder; bipolar disorder; post-traumatic stress disorder; inflammatory disorder; chemical dependency and addiction; gastrointestinal disorder; or skin disorder. In a further aspect the anxiety-related disorder is generalized anxiety or the mood disorder is depression. In 5 still a further aspect the gastrointestinal disorder is irritable bowel syndrome.

Other embodiments of the invention include methods of detecting a GPCR ligand comprising: a) contacting a sample suspected of containing a GPCR ligand with a sGPCR polypeptide; and b) assessing the presence or absence of sGPCR polypeptide bound ligand. The methods may further comprise characterizing the bound ligand. 10 Characterizing a bound ligand includes, but is not limited to various chromatographies, mass spectrometry, peptide sequencing and the like. The sGPCR polypeptide may or may not be operably coupled to a substrate or surface. The method can further comprise: c) administering a radiolabeled GPCR ligand; and d) assessing binding or competition for binding of the radiolabeled GPCR ligand to the sGPCR. The GPCR ligand may include, 15 but is not limited to corticotropin releasing factor (CRF), urocortin 1, urocortin 2, usorcortin 3, parathyroid hormone, PTH-related hormone, TIP39, calcitonin, amylin, CGRP (CALCA and CALCB), adrenomedullin, secretin, VIP, PACAP, glucagon, GHRH, GLP-1, GLP-2, or GIP.

Still other embodiments include methods of detecting a sGPCR comprising: a) 20 contacting a sample suspected of containing a sGPCR with a ligand that binds the sGPCR or a related surface bound GPCR; and b) assessing binding of GPCR ligand with components of the sample. The method can further comprise characterizing the bound sGPCR, which can include chromatography, mass spectrometry, protein fragmentation and sequencing, and the like. A GPCR ligand may be operably coupled to a substrate or surface. The methods can further comprise: c) administering a radiolabeled sGPCR; and d) 25 assessing binding or competition for binding of the radiolabeled sGPCR to the GPCR ligand in the presence and absence of the sample being tested. Exemplary ligands include corticotropin releasing factor (CRF), urocortin 1, urocortin 2, usorcortin 3, parathyroid hormone, PTH-related hormone, TIP39, calcitonin, amylin, CGRP (CALCA and CALCB), adrenomedullin, secretin, VIP, PACAP, glucagon, GHRH, GLP-1, GLP-30 2, GIP or other known GPCR ligands.

In yet still another embodiment of the invention includes antibodies that specifically bind a sGPCR. In certain aspects an antibody may bind the amino terminus

or carboxy terminus of the sGPCR. Aspects of the invention include an antibody that binds a carboxy terminal 5, 10, 15, 20 or more amino acid sequence, which may be derived from an alternative reading frame of a nucleotide sequence that encodes a transmembrane region of a GPCR (typically the result of alternative splicing and may be 5 engineered into a recombinant polynucleotide of the invention).

Embodiments of the invention include methods of detecting the expression of a sGPCR, either using protein, nucleic acid or both protein and nucleic acid evaluation or assessment. Aspects of the invention include methods of detecting a sGPCR mRNA comprising: a) obtaining a nucleic acid sample to be analyzed; and b) assessing the 10 presence of a sGPCR nucleic acid comprising a splice junction resulting in a sGPCR. The method may include assessing the presence of a particular species of mRNA by nucleic hybridization, nucleic acid amplification or other methods of analyzing nucleic acids. In a particular aspect a sGPCR is a soluble B type GPCR, a soluble B1 type GPCR, a soluble CRFR, a sCRFR1, a sCRFR2, or a sCRFR2 α . A polynucleotide can 15 include an exon/exon junction that includes the amino terminal amino acids of a GPCR and none or part of an exon encoding a portion of a transmembrane domain. In a particular aspect the splice junction of a sCRFR2 α is an exon 5/exon 7 junction, wherein exon designation is based on the genomic designation of CRFR2 exons. Based on the CRFR2a transcript the exons would be designated 3 and 5, respectively (see FIG. 1 for 20 an example).

A "soluble" GPCR (sGPCR) means a GPCR that comprises all or part of an extracellular domain of a receptor, but lacks all or part of one or more transmembrane domains which normally retains the full length receptor in the cell membrane, the soluble form is not integrated into the cell membrane. Thus, for example, when such a 25 soluble receptor is produced recombinantly in a mammalian cell, it can be secreted from the recombinant host cell through the plasma membrane, rather than remaining at the surface of the cell. In general, a soluble receptor of the invention is soluble in an aqueous solution. However, under certain conditions, the receptor can be in the form of an inclusion body, which is readily solubilized by standard procedures. Such sGPCR 30 may be derived from an engineered nucleic acid, a processed protein (e.g., protealized protein), a synthesized protein, or an isolated splice variant. A polynucleotide encoding such a sGPCR may be isolated or engineered.

As used herein, the terms "isolated" and "purified" are used interchangeably to refer to nucleic acids or polypeptides or biologically active portions thereof that are substantially or essentially free from components that normally accompany or interact with the nucleic acid or polypeptide as found in its naturally occurring environment.

5 Thus, an isolated or purified nucleic acid or polypeptide is substantially free of other cellular material or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

An "isolated" nucleic acid is free of sequences (preferably protein-encoding sequences) that naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acids can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acids in genomic DNA of the cell from which the nucleic acid is derived.

As used herein, the term "isolated" or "purified" as it is used to refer to a polypeptide of the invention means that the isolated protein is substantially free of cellular material and includes preparations of protein having less than about 30%, 20%, 10%, 5% or less (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, preferably culture medium represents less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one."

The use of the term "or" in the claims is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternative are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or."

30 Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific

embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

5 So that the matter in which the above-recited features, advantages, and objects of the invention as well as others which will become clear are attained and can be understood in detail, more particular descriptions and certain embodiments of the invention briefly summarized above are illustrated in the appended drawings. These drawings form a part of the specification. It is to be noted, however, that the appended
10 drawings illustrate certain embodiments of the invention and therefore are not to be considered limiting in their scope.

FIGs. 1A-1B. Illustrate an exemplary nucleotide and translated amino acid sequence of a soluble GPCR, the CRF receptor type 2 α (sCRFR2 α) (FIG. 1A). Underlined amino acids indicate the unique C-terminal tail. Boxed residues indicate
15 putative N-linked glycosylation sites. Schematic representation of the structure of the mouse CRFR2 gene (upper panel), the two known functional transcripts in mouse, α and β (middle panels) and the novel sCRFR2 α splice variant (lower panel) (FIG. 1B). The locations of the translation start sites (ATG) are indicated. Exons coding for the N-terminal extracellular domain (ECD), the seven transmembrane domains (7TM), and the
20 C-terminal cytoplasmic domain (CD) are indicated. 5' and 3' -UTRs are indicated by hatched boxes. Black boxes represent coding regions and open boxes represent exons downstream to the stop codon.

FIGs. 2A-2C. Show expression of CRFR2 α and sCRFR2 α mRNA in mouse brain and pituitary. FIG. 2A is a schematic representation and the oligonucleotide
25 primer locations of the amplified portion of mouse CRFR2 α (upper panel) and sCRFR2 α (lower panel) transcripts. The locations of the oligonucleotide primers, at exons three and seven, which result in the amplification of two products of 418 and 309 corresponding to CRFR2 α and sCRFR2 α , respectively, are indicated. FIG. 2B is a
30 representative image of electrophoretic analysis of the semiquantitative RT-PCR for mCRFR2 α and sCRFR2 α mRNA and the ribosomal protein S16 mRNA (upper panels). Southern blot hybridization of amplified mCRFR2 α and sCRFR2 α cDNA and the

ribosomal protein S16 cDNA fragments were also performed (lower panels). The radioactive bands were quantified by PhosphorImager and normalized values (relative to the S16 expression) are presented as relative densitometry units (FIG. 2C).

FIGs. 3A-3C. A highly specific antiserum raised in rabbit using a synthetic peptide fragment encoding the unique C-terminal tail of mouse sCRFR2 α protein (aa 113-143) was used to develop a sCRFR2 α radioimmunoassay, used for immunoblot analysis and for immunocytochemistry. FIG. 3A is a western immunoblot of mouse sCRFR2 α isolated from the medium of COS-M6 cells transiently transfected with sCRFR2 α FLAG construct reacted with anti-sCRFR2 α -(113-143) serum (left panel) or monoclonal M2 anti-FLAG (right panel). Lanes 1, 2, and 3 correspond to 0.1, 1.0, and 10 μ l of sCRFR2 α - FLAG extract, respectively. FIG. 3B, Displacement of [125 I]Tyr¹¹³ sCRFR2 α (aa 113-143) binding to rabbit anti-sCRFR2 α (aa113-143) by synthetic sCRFR2 α (aa 113-143) and by purified COS-M6 expressed sCRFR2 α (aa 113-143)-FLAG. FIG. 3C, Immunofluorescence staining of COS-M6 cells transiently transfected with mouse sCRFR2 α construct visualized with the anti-sCRFR2 α (aa 113-143) serum followed by a Cy3-conjugated secondary antibody (FIG. 3C(b)). The slides were counterstained with DAPI to visualize both transfected and non transfected cells (FIG. 3C(a)). Cells incubated with normal rabbit serum (NRS), as negative control, followed by a Cy3-conjugated secondary antibody did not show any staining (FIG. 3C(c)).

FIGs. 4A-4G. Illustrates the presence of sCRFR2 α -like immunoreactivity (ir) in the mouse brain using immunohistochemistry and radioimmunoassay (RIA). FIGs. 4A-4F show immunoperoxidase staining for sCRFR2 α in select mouse brain regions. Major sites of cellular expression included the principal output neurons of the olfactory bulb (FIG. 4A); the medial septal nucleus (FIG. 4B); and the basolateral (BLA), but not the central (CeA) nucleus of the amygdala (FIG. 4C); cerebral cortex, where stained cells were localized mainly in layers 5 and 2/3 (FIG. 4D); and red nucleus (FIG. 4E). In each of these sites, the pattern of cellular labeling was similar, though not necessarily identical, to that of CRFR1 mRNA expression. Immunolabeled fibers and varicosities were restricted to a handful of cell groups, including the paraventricular nucleus of the hypothalamus (PVH; FIG. 4F). FIG. 4G, sCRFR2 α -like immunoreactivity in acid-extracted and partially purified tissue isolated from mouse brain was measured by radioimmunoassay. Tissue extracts were tested at 5-7 dose levels and displaced [125 I]-

labeled Tyr¹¹³ sCRFR2 α (aa 113-143) binding to rabbit anti-sCRFR2 α (aa 113-143) in a dose-dependent manner.

FIGs. 5A-5B. sCRFR2 α protein interferes with the induction of cAMP and MAPK signaling mediated by Ucn 1 or CRF. FIG. 5A shows activation of CRE-luciferase reporter by Ucn 1 or CRF, with or without sCRFR2 α preincubation, in 293T cells transiently transfected with mouse CRFR2 α . Luciferase reporter containing a fragment of the CRE promoter of the EVX1 gene was cotransfected into 293T cells with CRFR2 α expression vectors. Luciferase activity was measured following treatment (4h) with 0.0001-100 nM Ucn 1 or CRF, in the presence or absence of 0.1 nM sCRFR2 α . Assays were normalized to cotransfected β -galactosidase activity. The representative means of six replicates from one experiment is shown in the graph. FIG. 5B, Equilibrated CATH.a cells were treated with Ucn 1 (10 nM) with or without sCRFR2 α (0.4 or 4 nM). After 5 min of receptor stimulation, cell lysates were harvested and subjected to SDS-PAGE immunoblot analysis using phospho-ERK1/2-p42,44 antibody and ERK2-p44 antibody. The ERK activation was calculated by normalizing the levels of phosphorylated ERK1/2-p42, 44 to total ERK2-p44. The representative of means of triplicates from one experiment is shown in the graph. *, P < 0.05 vs. vehicle treatment, #, P < 0.05 vs. Ucn 1 treatment, UD = undetected.

DETAILED DESCRIPTION OF THE INVENTION

Useful therapeutic approaches for the treatment of diseases associated with GPCRs and associated signalling pathways include the inhibition or modulation of the activation or inhibition of the GPCR. One approach is the development of small molecule inhibitors, which are costly to develop and bring to market. A drawback of the treatment with small molecule inhibitors or antagonists of GPCRs is the risk of toxicity, particularly with repeated application. Also, many GPCRs have no small molecule receptor antagonists. The development of a GPCR antagonist that is less costly and/or less toxic than small molecule inhibitors is worthwhile. Embodiments of the invention are directed to compositions and methods related to soluble GPCR (sGPCR) ligand binding domains, as well as its effects on GPCR signaling and interaction between GPCR ligands and their GPCRs. sGPCRs may be used to antagonize the activation or inhibition of GPCRs *in vitro* and/or *in vivo*.

I. G-PROTEIN COUPLED RECEPTORS (GPCRs)

5 GPCRs constitute a superfamily of proteins, which are divided into three families: rhodopsin-like family (family A), the calcitonin receptors (family B), and metabotropic glutamate family (family C) (Ji *et al.*, 1998), each of which may further be divided into subfamilies. The reported GPCRs include both characterized receptors and orphan receptors, those for which ligands have not yet been identified (Wilson *et al.*, 1999; Wilson *et al.*, 1998; Marchese *et al.*, 1999). Despite the large number of GPCRs, generally, each GPCR share a similar molecular structure. Each GPCR comprises a string of amino acid residues of various lengths. GPCRs lie within the cell membrane in 10 seven distinct coils called transmembranes. The amino terminus of the GPCR is outside the cell as are the extracellular loops, while the carboxy-terminus is inside the cell with the intracellular loops.

15 GPCR family A (Rhodopsin like) includes, but is not limited to amine, peptide, hormone protein, rhodopsin, olfactory, prostanoid, nucleotide-like, cannabinoid, platelet activating factor, gonadotropin-releasing hormone, thyrotropin-releasing hormone and secretagogue, melatonin, viral, lysosphingolipid and LPA (EDG), leukotriene B4 receptor and other similar receptor proteins.

20 GPCR family B (Secretin like) includes, but is not limited to receptors for calcitonin, corticotropin releasing factor (CRF), gastric inhibitory peptide (GIP), glucagon, growth hormone-releasing hormone (GHRH), parathyroid hormone (PTH), pituitary adenylate cyclase-activating polypeptide (PACAP), secretin, vasoactive intestinal polypeptide (VIP), diuretic hormone, EMR1, latrophilin, brain-specific angiogenesis inhibitor (BAI), methuselah-like proteins (MTH), cadherin/EGF/LAG (CELSR), and other similar ligands. Harmar (2001) describes three subfamilies of 25 GPCR family B, subfamily B1, B2 and B3.

30 Subfamily B1 -- Subfamily B1 includes, but is not limited to the classical hormone receptors, which are encoded by at least 15 genes in humans, with at least five putative members in *Drosophila* and three in *C. elegans*. The ligands for receptors in this family are polypeptide hormones of approximately 27-141 amino-acid residues; at least nine of the mammalian receptors respond to ligands that are structurally related to one another (glucagon, glucagon-like peptides (GLP-1, GLP-2), glucose-dependent insulinotropic polypeptide, secretin, vasoactive intestinal peptide (VIP), PACAP, and

growth-hormone-releasing hormone (GHRH). All members of this subfamily have been shown to be capable of regulating intracellular concentrations of cAMP by coupling to adenylate cyclase through a stimulatory G protein (Gs). Some members of the subfamily are capable of signaling through additional G-protein-coupled signaling pathways, for example through activation of phospholipase C.

Subfamily B2 -- Subfamily B2 consists of a large number of family-B GPCRs with long extracellular amino termini, containing diverse structural elements linked to the core 7TM motif. The prototype members of this subfamily were an EGF-module-containing, mucin-like hormone receptor (EMR1) isolated from a human 10 neuroectodermal cDNA library (Baud *et al.*, 1995) and the leukocyte cell-surface antigen CD97 (Hamann *et al.*, 1995). Subfamily B2 also includes the calcium-independent receptors for α -latrotoxin. Three genes encoding calcium-independent latrotoxin receptors (CL-1 CL-2 and CL-3) have been identified. Secondly, the brain-specific angiogenesis inhibitors 1, 2 and 3 (BAI1, BAI2, BAI3), a group of proteins that have 15 been implicated in the vascularization of glioblastomas are also included in this subfamily. Thirdly, the protein encoded by the *Drosophila* gene flamingo and its orthologs in humans (the cadherin EGF LAG seven-pass G-type receptors Celsr1, Celsr2 and Celsr3) and in *C. elegans* (F15B9.7) is also included in the B2 subfamily. Finally, the subfamily includes a fourth, diverse group of receptors that contain some motifs 20 common to receptors in subfamily B2 but are otherwise structurally unrelated (human epididymis 6 (HE6), EGF-TM7-latrophilin-related protein (ETL), the immunoglobulin-repeat-containing receptor Ig hepta, G-protein-coupled receptor 56 (GPR56) and very large G-protein-coupled receptor 1 (VLGR1)). Analysis of the sequenced human genome (1 April 2001, UCSC Human Genome Project Working Draft 25 (genome.ucsc.edu)) indicates that there are at least 18 human genes encoding members of subfamily B2, and there are at least four in *Drosophila* and three in *C. elegans*. The structure and functions of members of subfamily B2 have been reviewed recently by Stacey *et al.* (2000).

Subfamily B3 -- The prototype of a third group (subfamily B3) of family-B GPCRs is methuselah (mth), a gene isolated in a screen for single-gene mutations that 30 extended average lifespan in *D. melanogaster* (Lin *et al.*, 1998). The gene encodes a polypeptide that displays sequence similarity to other family-B GPCRs solely within the

TM7 region. At least eight paralogs of methuselah are encoded within the *Drosophila* genome sequence.

The characteristic feature of all family-B GPCRs is the 7TM motif, which is distantly related to comparable regions of some other GPCR families but much more 5 highly conserved within family B. Conserved cysteine residues within extracellular loops EC1 and EC2 probably form a disulphide bridge, by analogy with family-A GPCRs in which this feature is also conserved (Palczewski *et al.*, 2000). In contrast to family-A GPCRs, however, many of which appear to rely on internal hydrophobic sequences for targeting to the plasma membrane, most family-B GPCRs appear to have 10 an amino-terminal signal peptide. Studies using site-directed mutagenesis and the construction of chimeras between hormone receptors in family B have shown that the amino-terminal extracellular domain is essential for ligand binding but that the transmembrane domains and associated extracellular loop regions of the receptors provide information necessary for specific interaction with ligands. All of the hormone 15 receptors in family B contain a conserved region within the amino-terminal extracellular domain close to TM1 that may play a role in ligand binding. Splice variation in this region of the PAC1 receptor has been shown to influence ligand-binding specificity and affinity (Dautzenberg *et al.*, 1999).

Receptors in subfamily B2 contain a variety of additional structural motifs in 20 their large amino-terminal extracellular domains that suggest a role for this domain in cell-cell adhesion and signaling. These include EGF domains (in Celsr1, Celsr2, Celsr3, EMR1, EMR2, EMR3, CD97 and Flamingo), laminin and cadherin repeats (in Flamingo and its human orthologs Celsr1, Celsr2 and Celsr3), olfactomedin-like domains (in the latrotoxin receptors), thrombospondin type 1 repeats (in BAI1, BAI2 and BAI3) and, in 25 Ig hepta, an immunoglobulin C-2-type domain also found in fibroblast growth factor (FGF) receptor 2 and in the neural cell adhesion molecule L1. VLGR1 has two copies of a motif (Calx-beta) present in Na⁺-Ca²⁺-exchangers and integrin subunit β4.

Family C (Metabotropic glutamate/pheromone) GPCR includes Metabotropic glutamate, calcium-sensing like, putative pheromone receptors, GABA-B, orphan 30 GPRC5, orphan GPCR6, bride of sevenless proteins (BOSS), taste receptors (T1R) and other similar proteins.

In certain embodiments, the sGPCRs of the invention are class B receptors. In aspect, the sGPCRs of the invention are subfamily B1 receptors, and in a further aspect, the sGPCRs are CRFR1 and CRFR2, and parathyroid hormone receptor. Table 1 includes a non-limiting set of exemplary members of the GPCR family, accession numbers and associated UNIGENE and OMIM entries are incorporated herein by reference as of the priority date and the date of filing of this application. Unigene entries can be accessed by internet links contained in OMIM webpage. Numerous other GPCRs and their accession numbers may be found at the website defined by the following address on the world wide web gpcr.org/7tm/htmls/entries.html.

5

Table 1. Exemplary GPCRs.

GPCR description	Protein acc. # / mRNA acc # / OMIM # (each of which are incorporated by reference)
BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 1 (BAI 1)	O14514 / AB005297 / 602682
BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 2 (BAI 2)	O60241 / CR623649 / 602683
BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 3 (KIAA0550a) (BAI 3)	O60242 / AB005299 / 602684
CALCITONIN RECEPTOR (CT-R, CALCR)	P30988 / NM_001742 / 114131
LEUCOCYTE ANTIGEN CD97	P48960 / X84700 / 601211
CALCITONIN GENE-RELATED PEPTIDE TYPE 1 RECEPTOR, CGRP TYPE 1 RECEPTOR, CALCRL, CGRPR	Q16602 / NM_005795 / 114190
CORTICOTROPIN RELEASING FACTOR RECEPTOR 1 (CRF-R, CRF1, CRHR1, CRHR, CRFR)	P34998 / NM_004382 / 122561
CORTICOTROPIN RELEASING FACTOR RECEPTOR 2 (CRF-R, CRF2, CRHR2, CRF2R, CRH2R)	Q13324 / NM_001883 / 602034
CELL SURFACE GLYCOPROTEIN EMR1 (EMR1 HORMONE RECEPTOR)	Q14246 / X81479 / 600493
EGF-LIKE MODULE EMR2	AAF21974 / AF114491 / 606100
EGF-LIKE MODULE-CONTAINING MUCIN-LIKE RECEPTOR EMR3	AAK15076 / AF239764 / 606101
GASTRIC INHIBITORY POLYPEPTIDE RECEPTOR (GIP-R, GLUCOSE-DEPENDENT INSULINOTROPIC POLYPEPTIDE RECEPTOR)	P48546 / NM_000164 / 137241
GLUCAGON-LIKE PEPTIDE 1 RECEPTOR (GLP-1 RECEPTOR, GLP-1-R, GLP1R)	P43220 / NM_002062 / 138032
GLUCAGON RECEPTOR (GL-R, GCGR)	P47871 / NM_000160 / 138033
GLUCAGON-LIKE PEPTIDE 2 RECEPTOR (GLP-2 RECEPTOR, GLP-2-R, GLP-2R, GLP2R)	O95838 / NM_004246 / 603659
G PROTEIN-COUPLED RECEPTOR 56	AAD30545 / NM_005682 / 604110
GROWTH HORMONE-RELEASING HORMONE RECEPTOR (GHRH RECEPTOR, GRF RECEPTOR, GRFR, GHRHR)	Q02643 / NM_000823 / 139191
PITUITARY ADENYLYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE I RECEPTOR (PACAP TYPE I RECEPTOR, PACAP-R-1, ADCYAP1R1)	P41586 / NM_001118 / 102981
PARATHYROID HORMONE RECEPTOR (PTH2)	P49190 / NM_005048 / 601469
PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR (PTH1R)	Q03431 / NM_000316 / 168468
SECREtin RECEPTOR (SCT-R, SCTR)	P47872 / NM_002980 / 182098
VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 (VIPR1, VPAC1)	P32241 / NM_004624 / 192321
VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 2 (VIPR2, VIP2R, VPAC2, etc.)	P41587 / NM_003382 / 601970

A. Corticotropin releasing factor (CRF) and its receptors

As an example of GPCRs contemplated by the invention, the CRF receptors are described in detail. One of skill in the art would be able to adapt these specific teachings to other members of the GPCR family, particularly type B and more particularly to 5 subfamily B1 receptors. In certain aspects the invention includes, but is not limited to the sGPCR derived from the soluble corticotropin releasing factor receptors (sCRFR), in particular sCRFR2 α . The hypothalamic hypophysiotropic peptide corticotropin releasing factor (CRF), originally isolated from the hypothalamus (Vale *et al.*, 1981), plays an important role in the regulation of the hypothalamo-pituitary-adrenal (HPA) 10 axis under basal and stress conditions (River and Vale, 1983; Muglia *et al.*, 1995). Further, CRF acts to integrate endocrine, autonomic, and behavioral responses to stressors (River and Vale, 1983; Muglia *et al.*, 1995; Koob and Heinrichs, 1999). The 15 mammalian CRF peptide family comprises urocortin 1 (Ucn 1) (Vaughan *et al.*, 1995) and the peptides, urocortin 2 (Ucn 2) and urocortin 3 (Ucn 3) also known as stresscopin-related peptide (Reyes *et al.*, 2001; Hsu and Hsueh, 2001), and stresscopin (Hsu and Hsueh, 2001; Lewis *et al.*, 2001), respectively.

The effects of CRF-related peptides are mediated through activation of two high affinity membrane receptors, CRFR1 (Chen *et al.*, 1993; Vita *et al.*, 1993; Chang *et al.*, 1993) and CRFR2 (Perrin *et al.*, 1995; Stenzel *et al.*, 1995; Kishimoto *et al.*, 1995; 20 Lovenberg *et al.*, 1995; Chen *et al.*, 2005), which belong to the B1 subfamily of seven-transmembrane domain (7TMD) receptors that signal by coupling to G-proteins. One functional variant of the CRFR1 gene is expressed both in humans and rodents, along with several non-functional variants, which are produced by differential splicing of various exons (Pisarchik and Slominski, 2004; Grammatopoulos *et al.*, 1999). The 25 CRFR2 has three functional splice variants in human (α , β , and γ) and two rodent variants (α and β) that are produced by the use of alternate 5' exons (Perrin *et al.*, 1995; Stenzel *et al.*, 1995; Kishimoto *et al.*, 1995; Lovenberg *et al.*, 1995; Chen *et al.*, 2005; Grammatopoulos *et al.*, 1999; Kostich *et al.*, 1998). CRFR1 mRNA is widely expressed 30 in mammalian brain and pituitary, with high levels found in the anterior pituitary, cerebral cortex, cerebellum, amygdala, hippocampus, and olfactory bulb (Van Pett *et al.*, 2000). In the periphery, CRFR1 is expressed in testes, ovary, skin, and spleen. CRFR2 mRNA is expressed in a discrete pattern in the brain with highest densities in the lateral

septal nucleus (LS), bed nucleus of stria terminalis (BNST), ventromedial hypothalamic nucleus (VMH), olfactory bulb, and mesencephalic raphe nuclei (Van Pett *et al.*, 2000). The CRFR2 α is the major splice variant expressed in the rodent brain (Lovenberg *et al.* 1995) while CRFR2 β is expressed in peripheral tissues, with highest levels in the 5 skeletal muscle, heart, and skin (Perrin *et al.*, 1995).

The distributions of CRFR1 and CRFR2 are distinct and imply diverse physiological functions, as demonstrated by the divergent phenotypes of the CRFR1 or CRFR2 null mice. Mice deficient for CRFR1 display decreased anxiety-like behavior and have an impaired stress response (Smith *et al.*, 1998; Timpl *et al.*, 1998), while the 10 CRFR2-null mice have increased anxiety-like behaviors and an exaggerated HPA response to stress (Zhu *et al.*, 1999; Valerio *et al.*, 2001; Khan *et al.*, 1993). However, the responses to administration of CRFR2 agonists and antagonists into specific brain regions reveal both anxiolytic and anxiogenic roles for CRFR2 (Bale and Vale, 2004).

15 Radioreceptor and functional assays have demonstrated that CRFR1 and CRFR2 differ pharmacologically: Ucn 1 has equal affinities for both receptors and is more potent than CRF on CRFR2, whereas Ucn 2 and Ucn 3 appear to be selective for CRFR2 (Vaughan *et al.*, 1995; Reyes *et al.*, 2001; Lewis *et al.*, 2001). The activation of specific CRFRs in distinct tissues or cell types by receptor-selective CRF peptides initiates a variety of signaling pathways, including coupling to different G-proteins, stimulation of 20 PKB, PKC, intracellular calcium, and mitogen-activated protein kinase (MAPK) (for reviews see Bale and Vale, 2004; Perrin and Vale, 1999; Brar *et al.*, 2002).

CRFR1 and CRFR2 both exist as multiple splice variants. The inventors have identified a cDNA from mouse brain encoding an exemplary splice variant of sCRFR2 α in which exon six is deleted from the nucleic acid encoding CRFR2 α . Translation of 25 this isoform produces a predicted 143 amino acid soluble protein. The translated protein includes a majority of the first extracellular domain (ECD1) of the CRFR2 α followed by a unique 38 amino acid hydrophilic C-terminus resulting from a frame shift produced by deletion of exon six. Studies have demonstrated high levels of expression of sCRFR2 α in the olfactory bulb, cortex, and midbrain regions. A protein corresponding to 30 sCRFR2 α , expressed and purified from either mammalian or bacterial cell systems, binds several CRF family ligands with low nanomolar affinities. Further, the purified sCRFR2 α protein inhibits cellular responses to CRF and urocortin 1. Thus, a sCRFR2 α

protein can be a biological modulator of CRF family ligands. The modulation of CRF family ligands is not limited to brain and may be used in any tissue that is exposed to one or more members of the CRF family of ligands.

Aspects of the invention generally relate to compositions and methods of 5 achieving a therapeutic effect, including the modulation of GPCR ligand activity, such as CRF family ligands, using a soluble GPCR ligand binding polypeptide, such as CRF binding polypeptide, as an antagonist either alone or together with one or more other hormone antagonist (e.g., small molecule antagonist), including but not limited to antagonist of ligand(s) of the CRF family.

10 One manner in which to antagonize the action of a ligand is to subject the ligand to a decoy or soluble receptor so as to limit the local concentration of ligand(s) that bind the decoy and modulate the ligands ability to signal via its cell surface receptor. Soluble proteins related to membrane receptors can be generated by enzymatic truncation of membrane bound receptors as suggested for the GHRH receptor (Rekaski *et al.*, 2000),
15 dopamine D3 receptor (Liu *et al.*, 1994), and calcitonin receptor (Seck *et al.*, 2003), or by alternative splicing in the case of the glutamate receptors (Malherbe *et al.*, 1999; Zhu *et al.*, 1999; Valerio *et al.*, 2001). Splice variants containing only the extracellular region of GPCRs have been reported (Pisarchik and Slominski, 2004; Grammatopoulos *et al.*, 1999; Kostich *et al.*, 1998; Malherbe *et al.*, 1999; Zhu *et al.*, 1999; Valerio *et al.*,
20 2001; Khan *et al.*, 1993; Graves *et al.*, 1992; You *et al.*, 2000; Schwarz *et al.*, 2000). In the majority of cases, these proteins act as binding, non-signaling molecules also referred to as decoy receptors. Two partial cDNA fragments (CRFR1e and CRFR1h), comprising deletion of exon 3 and 4, and addition of a cryptic exon in CRFR1 were identified in human skin and predicted to exist as a soluble proteins (Pisarchik and Slominski, 2004).
25 One of these fragments, CRFR1e, exhibited dominant negative effects when co-transfected with the wildtype CRFR1.

30 Kehne and Lombaert (2002) discuss non-peptidic CRF receptor antagonists for the treatment of anxiety, depression, and stress disorders. CRF is implicated in psychiatric disorders, such as anxiety and depression. Since the identification of corticotropin releasing factor (CRF) an extensive research effort has solidified the importance of this 41 amino acid peptide and its related family members in mediating the body's behavioral, endocrine, and autonomic responses to stress.

Preclinical and clinical evidence implicate CRF, in general, and CRF receptors, in particular, in anxiety and depression. Clinical studies have demonstrated a dysfunctional hypothalamic-pituitary-adrenal (HPA) axis and/or elevated CRF levels in depression and in some anxiety disorders. Preclinical data utilizing correlational methods, genetic models, and exogenous CRF administration techniques in rodents and non-human primates supports a link between hyperactive CRF pathways and anxiogenic and depressive-like symptoms. Studies employing the use of receptor knockouts and selective, non-peptidic antagonists of the CRFR1 have demonstrated anxiolytic and antidepressant effects under certain types of laboratory conditions. A Phase II, open-label, clinical trial in major depressive disorder has reported that a CRFR1 antagonist was safe and effective in reducing symptoms of anxiety and depression.

Various nonlimiting activities of CRF antagonists are described by Owens *et al.* (1991). CRF antagonists are described as being effective in the treatment of stress-related illnesses; mood disorders such as depression, major depressive disorder, single 15 episode depression, recurrent depression, child abuse induced depression, postpartum depression, dysthemia, bipolar disorders, and cyclothymia; chronic fatigue syndrome; eating disorders such as anorexia and bulimia nervosa; generalized anxiety disorder; panic disorder; phobias; obsessive-compulsive disorder; post-traumatic stress disorder; pain perception such as fibromyalgia; headache; gastrointestinal diseases; hemorrhagic 20 stress; ulcers; stress-induced psychotic episodes; fever; diarrhea; post-operative ileus; colonic hypersensitivity; irritable bowel syndrome; Crohn's disease; spastic colon; inflammatory disorders such as rheumatoid arthritis and osteoarthritis; pain; asthma; psoriasis; allergies; osteoporosis; premature birth; hypertension, congestive heart failure; sleep disorders; neurodegenerative diseases such as Alzheimer's disease, senile dementia 25 of the Alzheimer's type, multiinfarct dementia, Parkinson's disease, and Huntington's disease; head trauma; ischemic neuronal damage; excitotoxic neuronal damage; epilepsy; stroke; spinal cord trauma; psychosocial dwarfism; euthyroid sick syndrome; syndrome of inappropriate antidiuretic hormone; obesity; chemical dependencies and addictions; drug and alcohol withdrawal symptoms; infertility; cancer; muscular spasms; 30 urinary incontinence; hypoglycemia and immune dysfunctions including stress induced immune dysfunctions, immune suppression, and human immunodeficiency virus infections; and stress-induced infections in humans and animals. These and other conditions amenable to CRF modulation are set out in the literature, that includes

Lovenberg *et al.* (1995); Chalmers *et al.* (1996); and U.S. Patent 5,063,245, each of which is incorporated in its entirety by reference.

II. POLYPEPTIDES

Polypeptides of the invention include soluble forms of GPCRs or soluble receptors. Soluble receptors of the invention may comprise subunits which have been changed from a membrane bound to a soluble form. Thus, soluble peptides may be produced by truncating the polypeptide to remove, for example, the 7 transmembrane regions and/or the cytoplasmic tail. Alternatively, the transmembrane domains may be abolished by deletion, or by substitutions of the normally hydrophobic amino acid residues which comprise a transmembrane domain with hydrophilic ones. In either case, a substantially hydrophilic or soluble polypeptide is created which will reduce lipid affinity and improve aqueous solubility. Deletion of the transmembrane domains is preferred over substitution with hydrophilic amino acid residues because it avoids introducing potentially immunogenic epitopes. Soluble receptors of the invention may include any number of well-known leader sequences at the N-terminus. Such a sequence would allow the peptides to be expressed and targeted to the secretion pathway in a eukaryotic system.

A. Fusion Proteins

Receptors are powerful tools to elucidate biological pathways and to treat various disease states via their easy conversion to immunoglobulin fusion proteins. These dimeric soluble receptor forms are good inhibitors of events mediated by either secreted or surface bound ligands. By binding to these ligands they prevent the ligand from interacting with cell associated receptors. Not only are these receptor-Ig fusion proteins useful in an experimental sense, but they have been successfully used clinically in the case of TNF-R-Ig to treat inflammatory bowel disease, rheumatoid arthritis, and the acute clinical syndrome accompanying OKT3 administration (Eason *et al.*, 1996; van Dullemen *et al.*, 1995). The inventors contemplate that manipulation of the many events mediated by signaling through the GPCRs will have wide application in the treatment of GPCR associated diseases.

Preferably, stable plasma proteins -- which typically have a half-life greater than hours in the circulation of a mammal -- can be used to construct the receptor fusion proteins. Such plasma proteins include but are not limited to: immunoglobulins, serum

albumin, lipoproteins, apolipoproteins and transferrin. Sequences that can target the soluble receptors to a particular cell or tissue type may also be attached to the receptor ligand binding domain to create a specifically localized soluble receptor fusion protein.

All or a functional fragment of GPCR extracellular region comprising the GPCR

5 ligand binding domain may be fused to an immunoglobulin constant region like the Fc domain of a human IgG1 heavy chain. Soluble receptor-IgG fusions proteins are common immunological reagents and methods for their construction are well known in the art (see, for example U.S. Patent 5,225,538, which is incorporated herein in its entirety by reference).

10 A functional GPCR ligand binding domain may be fused to an immunoglobulin (Ig) Fc domain. The Ig Fc may be derived from an immunoglobulin class or subclass including but not limited to IgG1. The Fc domains of antibodies belonging to different Ig classes or subclasses can activate diverse secondary effector functions. Activation occurs when the Fc domain is bound by a cognate Fc receptor. Secondary effector

15 functions include the ability to activate the complement system or to cross the placenta. The properties of the different classes and subclasses of immunoglobulins are described in the art.

One skilled in the art will appreciate that different amino acid residues forming the junction point of the receptor-Ig fusion protein may alter the structure, stability and

20 ultimate biological activity of the sGPCR fusion protein. One or more amino acids may be added to the C- terminus of the selected sGPCR fragment to modify the junction point with the selected fusion domain.

The N-terminus of the sGPCR fusion protein may also be varied by changing the position at which the selected sGPCR DNA fragment is cleaved at its 5' end for insertion

25 into the recombinant expression vector. The stability and activity of each sGPCR fusion protein may be tested and optimized using routine experimentation, including but not limited to assays for ligand binding.

Using sGPCR ligand binding domain sequences within the extracellular domain as shown herein, amino acid sequence variants may also be constructed to modify the

30 affinity of the sGPCR molecules for their ligands. The soluble molecules of this invention can compete for binding with endogenous receptors. It is envisioned that any soluble molecule comprising a GPCR ligand binding domain that can compete with

native receptors for ligand binding is a receptor blocking agent or ligand trapping agent that falls within the scope of the present invention.

B. Protein Conjugates

With respect to the protein's half-life, one way to increase the circulation half-life of a protein is to ensure a reduction in the clearance of the protein, in particular via renal clearance and receptor-mediated clearance. This may be achieved by conjugating the protein to a chemical moiety which is capable of increasing the apparent size, thereby reducing renal clearance and increasing the *in vivo* half-life. Furthermore, attachment of a chemical moiety to the protein may effectively block proteolytic enzymes from physical contact with the protein, thus preventing degradation by non-specific proteolysis. Polyethylene glycol (PEG) is one such chemical moiety that has been used in the preparation of therapeutic protein products. Recently, G-CSF molecule modified with a single, N-terminally linked 20 kDa PEG group (Neulastam) was approved for sale in the United States. This PEGylated G-CSF molecule has been shown to have an increased half-life compared to non-PEGylated G-CSF and thus may be administered less frequently than current G-CSF products, but it does not reduce the duration of neutropenia significantly compared to non-PEGylated G-CSF.

Polyethylene glycol (PEG) modification is important for pharmaceutical and biotechnological applications. PEGylation (the covalent attachment of PEG) leads for example to shielding of antigenic or immunogenic epitopes. Moreover, it reduces receptor-mediated uptake by the reticuloendothelial system or prevents recognition and degradation by proteolytic enzymes. PEGylation of proteins has been shown to increase their bioavailability by reducing the renal filtration.

The term "conjugate" is intended to indicate a heterogeneous molecule formed by the covalent attachment of one or more polypeptides, typically a single polypeptide, to one or more non-polypeptide moieties such as polymer molecules, lipophilic compounds, carbohydrate moieties or organic derivatizing agents. The term covalent attachment means that the polypeptide and the non-polypeptide moiety are either directly covalently joined to one another, or else are indirectly covalently joined to one another through an intervening moiety or moieties, such as a bridge, spacer, or linkage moiety or moieties. Preferably, the conjugate is soluble at relevant concentrations and conditions, *i.e.*, soluble in physiological fluids such as blood. Compositions and methods for

preparing a conjugate of the invention are described in U.S. Patent 6,831,158, which is incorporated herein by reference in its entirety. The methods described in U.S. Patent 6,831,158 are directed to conjugation of G-CSF, but can be readily adapted to conjugation of the sGPCRs of the present invention.

5 The "polymer molecule" is a molecule formed by covalent linkage of two or more monomers. The term "polymer" may be used interchangeably with the term "polymer molecule". The term is intended to cover carbohydrate molecules including carbohydrate molecules attached to the polypeptide by *in vivo* N- or O-glycosylation, such molecule is also referred to as "an oligosaccharide moiety". Except where the
10 number of polymer molecule(s) is expressly indicated every reference to "a polymer", "a polymer molecule", "the polymer" or "the polymer molecule" contained in a polypeptide of the invention or otherwise used in the present invention shall be a reference to one or more polymer molecule(s).

15 The term "attachment group" is intended to indicate an amino acid residue group of the polypeptide capable of coupling to the relevant non-polypeptide moiety. For instance, for polymer conjugation, in particular to PEG, a frequently used attachment group is the ϵ -amino group of lysine or the N-terminal amino group. Other polymer attachment groups include a free carboxylic acid group (e.g., that of the C-terminal amino acid residue or of an aspartic acid or glutamic acid residue), suitably activated
20 carbonyl groups, oxidized carbohydrate moieties and mercapto groups. Useful attachment groups and their matching non-peptide moieties are exemplified in Table 2.

Table 2

Attachment Group	Amino Acid	Example of non-peptide moiety	Conjugation method / activated PEG	Reference
-NH ₂	N-terminal Lys, Arg, His	Polymer, e.g., PEG with amide or imine group	mPEG-SPA Tresylated mPEG	Shearwater Corp. Delgado <i>et al.</i> , 1992.
-COOH	C-terminal Asp and Glu	Polymer, e.g., PEG with ester or amide group Oligosaccharide moiety	mPEG-Hz in vitro coupling	Shearwater Corp.
-SH	Cys	Polymer, e.g. PEG, with disulfide, maleimide or vinyl sulfone group Oligosaccharide moiety	PEG vinylsulfone PEG-maleimide In vitro coupling	Shearwater Corp. Delgado <i>et al.</i> , 1992
-OH	Ser, Thr, -OH, lys	Oligosaccharide moiety PAG with ester, ether, carbamate, carbonate	In vivo O-linked glycosylation	
-CONH ₂	Asn as part of an N-glycosylation site	Oligosaccharide moiety Polymer, e.g. PEG	In vivo N-glycosylation	
Aromatic – CONH ₂	Phe, Tyr, Trp, Gln	Oligosaccharide moiety	In vitro coupling	Yan and Wold, 1984
Aldehyde Ketone	Oxidized oligosaccharide	Polymer, e.g. PEG PEG hydroxide	PEGylation	Andresz <i>et al.</i> , 1978 WO 92/16655 WO 00/23114
Guanidino	Arg	Oligosaccharide moiety	In vitro coupling	Lunblad and Noyes, Chemical reagents for protein modification, CRC Press
Imidazole ring	His	Oligosaccharide moiety	In vitro coupling	Lunblad and Noyes, Chemical reagents for protein modification, CRC Press

C. Site-Specific Mutagenesis

In one embodiment, amino acid sequence variants of a polypeptide can be prepared. These may, for instance, be minor sequence variants of polypeptides that arise due to natural variation within the population or they may be homologs found in other species. They also may be sequences that do not occur naturally but that are sufficiently similar that they function similarly and/or elicit an immune response that cross-reacts with natural forms of the polypeptide. Sequence variants can be prepared by standard methods of site-directed mutagenesis such as those described below.

Amino acid sequence variants of the polypeptide can be substitutional, insertional, or deletion variants. Deletion variants lack one or more residues of the native protein which are not essential for function or immunogenic activity, and are exemplified by the variants of a receptor lacking a transmembrane sequence.

5 Substitutional variants typically contain the exchange of one amino acid for another at one or more sites within the protein, and may be designed to modulate one or more properties of the polypeptide such as stability against proteolytic cleavage or immunogenicity. Substitutions preferably are conservative, that is, one amino acid is replaced with one of similar shape and charge. Conservative substitutions are well known in the art and include, for example, the changes of: alanine to serine; arginine to lysine; asparagine to glutamine or histidine; aspartate to glutamate; cysteine to serine; glutamine to asparagine; glutamate to aspartate; glycine to proline; histidine to asparagine or glutamine; isoleucine to leucine or valine; leucine to valine or isoleucine; lysine to arginine; methionine to leucine or isoleucine; phenylalanine to tyrosine, leucine or methionine; serine to threonine; threonine to serine; tryptophan to tyrosine; tyrosine to tryptophan or phenylalanine; and valine to isoleucine or leucine.

10

15

Insertional variants include fusion proteins such as those used to allow rapid purification of the polypeptide and also can include hybrid proteins containing sequences from other proteins and polypeptides. For example, an insertional variant could include portions of the amino acid sequence of a polypeptide from one species, together with portions of the homologous polypeptide from another species. Other insertional variants can include those in which additional amino acids are introduced within the coding sequence of the polypeptide, for example a protease cleavage site(s) may be introduced.

20 Modification and changes may be made in the structure of a polynucleotide and still obtain a functional molecule that encodes a protein or polypeptide with desirable characteristics. The following is a discussion based upon changing the amino acids of a protein to create an equivalent, or even an improved, second-generation molecule. The amino acid changes may be achieved by changing the codons of the DNA sequence, according to the following data.

25 30 For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate

molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological activity certain amino acid substitutions can be made in a protein sequence still obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the DNA sequences of genes, mRNA or 5 polynucleotides without appreciable loss of their biological utility or activity.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte & Doolittle, 1982).

10 **Table 3**

Amino Acids			Codons			
Alanine	Ala	A	GCA	GCC	GCG	GCU
Cysteine	Cys	C	UGC	UGU		
Aspartic acid	Asp	D	GAC	GAU		
Glutamic acid	Glu	E	GAA	GAG		
Phenylalanine	Phe	F	UUC	UUU		
Glycine	Gly	G	GGA	GGC	GGG	GGU
Histidine	His	H	CAC	CAU		
Isoleucine	Ile	I	AUA	AUC	AUU	
Lysine	Lys	K	AAA	AAG		
Leucine	Leu	L	UUA	UUG	CUA	CUC
Methionine	Met	M	AUG		CUG	CUU
Asparagine	Asn	N	AAC	AAU		
Proline	Pro	P	CCA	CCC	CCG	CCU
Glutamine	Gln	Q	CAA	CAG		
Arginine	Arg	R	AGA	AGG	CGA	CGC
Serine	Ser	S	AGC	AGU	UCA	UCC
Threonine	Thr	T	ACA	ACC	ACG	ACU
Valine	Val	V	GUU	GUC	GUG	GUU
Tryptophan	Trp	W	UGG			
Tyrosine	Tyr	Y	UAC	UAU		

It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of 15 the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those 20 which are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological 5 property of the protein.

It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent and immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those that are within ± 1 are particularly 10 preferred, and those within ± 0.5 are even more particularly preferred.

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent proteins or peptides, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, incorporating one or more of the foregoing 15 considerations, by introducing one or more nucleotide sequence changes into the DNA. In general, the technique of site-specific mutagenesis is well known in the art. The technique typically employs a bacteriophage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage vectors are commercially available and 20 their use is generally well known to those skilled in the art. Double stranded plasmids are also routinely employed in site-directed mutagenesis, which eliminates the step of transferring the gene of interest from a phage to a plasmid.

The preparation of sequence variants of a GPCR, including but not limited to sCRFR2 α , polynucleotide using site-directed mutagenesis is provided as a means of 25 producing potentially useful species, *i.e.*, species with altered ligand binding properties that include an increased affinity for a particular ligand, and is not meant to be limiting, as there are other ways in which sequence variants of nucleic acids may be obtained. For example, recombinant vectors encoding the desired gene may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants.

30 **D. Expression and Purification of Polypeptides**

The polynucleotides of the invention, in particular 100, 150, 200 250, 300, 400, 450, 500, 550 or more contiguous nucleotides of the DNA encoding a GPCR, a family B

5 GPCR, a family B1 GPCR, or a polynucleotide that is 70, 75, 80, 85, 90, 95, 98, or 100% identical to the sequence specified in the accompanying sequence listing, e.g., SEQ ID NO:1, 3, 5, 7, 9, 11, 13, or 14 can be expressed as encoded peptides or proteins. In a particular aspect the DNA encodes all or part of a GPCR extracellular domain and in particular an amino terminal extracellular domain. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. It is believed that virtually any expression system may be employed in the expression of the claimed nucleic acid sequences.

10 In certain embodiments, the present invention concerns novel compositions comprising at least one proteinaceous molecule, such as sGPCR, asCRFR, or a sCRFR2. As used herein, a “proteinaceous molecule,” “proteinaceous composition,” “proteinaceous compound,” “proteinaceous chain” or “proteinaceous material” generally refers, but is not limited to, a protein of greater than about 200 amino acids or the full 15 length endogenous sequence translated from a gene; a polypeptide of greater than about 100 amino acids; and/or a peptide of from about 3 to about 100 amino acids. All the “proteinaceous” terms described above may be used interchangeably herein. Furthermore, these terms may be applied to fusion proteins as well.

20 In certain embodiments the size of the at least one proteinaceous molecule may comprise, but is not limited to, about or at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1750, 2000, 2250, 2500 or greater amino molecule residues, and any range derivable therein, particularly 50, 60, 70, 80, 90, 100, 110, 120, 25 130, 140, 150, 160, 170, 180, 190, or more contiguous amino acid sequences of such lengths of a GPCR, a family B GPCR, a family B1 GPCR, or SEQ ID NO:2, 4, 6, 8, 10, 12 or 15, including the full length of SEQ ID NO:4, 8, 12, or 15. Both cDNA and genomic sequences are suitable for eukaryotic expression, as the host cell will generally process the genomic transcripts to yield functional mRNA for translation into protein.

30 As used herein, the terms “engineered” and “recombinant” cells are intended to refer to a cell into which an exogenous DNA segment or polynucleotide, such as a cDNA or polynucleotide has been introduced. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly

introduced exogenous DNA segment or gene. Engineered cells are thus cells having a gene or genes introduced through the hand of man. Recombinant cells include those having an introduced cDNA or genomic DNA, and may also include genes positioned adjacent to a promoter not naturally associated with the particular introduced gene.

5 To express a recombinant protein or polypeptide, whether mutant or wild-type, in accordance with the present invention one would prepare an expression vector that comprises one of the claimed isolated nucleic acids under the control of one or more promoters. To bring a coding sequence “under the control of” a promoter, one positions the 5' end of the translational initiation site of the reading frame generally between about 10 1 and 50 nucleotides “downstream” of (*i.e.*, 3' of) the chosen promoter. The “upstream” promoter stimulates transcription of the inserted DNA and promotes expression of the encoded recombinant protein. This is the meaning of “recombinant expression” in the context used here.

15 Many standard techniques are available to construct expression vectors containing the appropriate nucleic acids and transcriptional/translational control sequences in order to achieve protein or peptide expression in a variety of host-expression systems. Cell types available for expression include, but are not limited to, bacteria, such as *E. coli*, *B. subtilis*, *E. coli* strain RR1, *E. coli* LE392, *E. coli* B, *E. coli* χ 1776 (ATCC No. 31537) as well as *E. coli* W3110 (F-, lambda-, prototrophic, ATCC No. 273325); bacilli such as *Bacillus subtilis*; and other enterobacteriaceae such as *Salmonella typhimurium*, *Serratia marcescens*, and various *Pseudomonas* species transformed with recombinant phage DNA, plasmid DNA or cosmid DNA expression vectors.

25 The polynucleotide or polynucleotide fragment encoding a polypeptide can be inserted into an expression vector by standard subcloning techniques. In one embodiment, an *E. coli* expression vector is used that produces the recombinant polypeptide as a fusion protein, allowing rapid affinity purification of the protein. Examples of such fusion protein expression systems are the glutathione *S*-transferase system (Pharmacia, Piscataway, NJ), the maltose binding-protein system (New England 30 Biolabs, Beverley, MA), the FLAG system (IBI, New Haven, CT), and the 6xHis system (Qiagen, Chatsworth, CA). Further useful vectors include pIN vectors (Inouye *et al.*, 1985); and pGEX vectors, for use in generating glutathione *S*-transferase (GST) soluble

fusion proteins. Other suitable fusion proteins are those with β -galactosidase, ubiquitin, or the like.

For expression in *Saccharomyces*, the plasmid YRp7, for example, is commonly used (Stinchcomb *et al.*, 1979; Kingsman *et al.*, 1979; Tschemper *et al.*, 1980). This 5 plasmid contains the *trpl* gene, which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1 (Jones, 1977). The presence of the *trpl* lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

10 Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, 1980) or other glycolytic enzymes (Hess *et al.*, 1968; Holland *et al.*, 1978), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable 15 expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination.

Other suitable promoters, which have the additional advantage of transcription 20 controlled by growth conditions, include the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization.

In addition to micro-organisms, cultures of cells derived from multicellular 25 organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture including mammalian and insect cells (e.g., U.S. Patent No. 4,215,051).

Examples of useful mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, WI38, BHK, COS-7, 293, HepG2, NIH3T3, 30 RIN and MDCK cell lines. In addition, a host cell may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the

specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the encoded protein.

Specific initiation signals may also be required for efficient translation of the claimed isolated nucleic acid coding sequences. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may additionally need to be provided. One of ordinary skill in the art would readily be capable of determining this need and providing the necessary signals. It is well known that the initiation codon must be in-frame (or in-phase) with the reading frame of the desired coding sequence to ensure translation of the entire insert. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements or transcription terminators (Bittner *et al.*, 1987).

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express constructs encoding G-proteins may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with vectors controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, *etc.*), and a selectable marker. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched medium, and then are switched to a selective medium. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci, which in turn can be cloned and expanded into cell lines.

A number of selection systems may be used, including, but not limited, to the herpes simplex virus thymidine kinase (Wigler *et al.*, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska *et al.*, 1962) and adenine phosphoribosyltransferase genes (Lowy *et al.*, 1980), in *tk*, *hgprt* or *aprt* cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for *dhfr*, which confers resistance to methotrexate (Wigler *et al.*, 1980; O'Hare *et al.*, 1981); *gpt*, which confers resistance to mycophenolic acid (Mulligan *et al.*, 1981); *neo*, which confers resistance to the aminoglycoside G-418 (Colbere-Garapin *et al.*, 1981); and *hygro*, which confers resistance to hygromycin.

Once the polynucleotide sequence coding a particular polypeptide has been determined or engineered, the polynucleotide can be inserted into an appropriate expression system. In this case, the inventors contemplate a polynucleotide encoding a sGPCR ligand binding domain polypeptide. The polynucleotide can be expressed in any 5 number of different recombinant DNA expression systems to generate large amounts of the polypeptide product, which can then be purified and/or isolated to be used as a therapeutic or to vaccinate animals to generate antisera, or in certain aspects of the invention as an antagonist of GPCR ligand and/or GPCR activation. In further aspects, sGPCRs of the invention can be used in methods to detect, screen, or identify ligands, 10 receptors, or agonist and/or antagonist of GPCRs. A polynucleotide of the invention may be expressed to obtain a GPCR ligand binding domain, a family B GPCR ligand binding domain, a family B1 GPCR ligand binding domain, a sCRFR ligand binding domain or a CRFR2 ligand binding domain polypeptide comprising an amino acid sequence including all or part of the amino acid sequence as set forth in the sequence 15 listing, *e.g.*, SEQ ID NO:2, 4, 6, 8, 10, 12, or 15.

As an alternative to recombinant polypeptides, synthetic peptides corresponding to the polypeptides of the invention can be prepared, including antigenic peptides. Such antigenic peptides are at least six amino acid residues long, and may contain up to approximately 35 residues. Automated peptide synthesis machines include those 20 available from Applied Biosystems (Foster City, CA). Use of such small peptides for vaccination typically requires conjugation of the peptide to an immunogenic carrier protein such as hepatitis B surface antigen, keyhole limpet hemocyanin or bovine serum albumin. Methods for performing this conjugation are well known in the art.

1. Purification of Expressed Proteins

25 Further aspects of the present invention concern the purification for isolation, and in particular embodiments, the substantial purification, of a protein or peptide comprising all or part of a sGPCR ligand binding domain. The term "purified or isolated protein or peptide" as used herein, is intended to refer to a composition, isolatable from other components, wherein the protein, polypeptide or peptide is purified to any degree 30 relative to its naturally-obtainable state, *i.e.*, in this case, relative to its purity within a organism or tissue. A purified or isolated protein or peptide therefore also refers to a protein or peptide, free from the environment in which it may naturally occur. A

purified or isolated protein or polypeptide may have a purity greater than or at least 70, 75, 80, 85, 90, 95, 98, or 99% purity.

Generally, “purified” will refer to a protein or peptide composition that has been subjected to fractionation to remove various other components, and which composition 5 substantially retains its expressed activity. Where the term “substantially purified” is used, this designation will refer to a composition in which the protein or peptide forms the major component of the composition, such as constituting about 50% or more of the proteins in the composition.

Various methods for quantifying the degree of purification of the protein or 10 peptide will be known to those of skill in the art in light of the present disclosure. These include, for example, determining the specific activity (e.g., binding affinity for GPCR ligand including, but not limited to CRF or a ligand of the CRF family) of an active fraction, or assessing the number of polypeptides within a fraction by SDS/PAGE analysis. A preferred method for assessing the purity of a fraction is to calculate the 15 specific activity of the fraction, to compare it to the specific activity of the initial extract, and to thus calculate the degree of purity, herein assessed by a “-fold purification number.” The actual units used to represent the amount of activity, which may include binding activity or affinity, will, of course, be dependent upon the particular assay technique chosen.

20 Various techniques suitable for use in protein purification will be well known to those of skill in the art. These include, for example, precipitation with ammonium sulphate, polyethylene glycol, antibodies and the like or by heat denaturation, followed by centrifugation; chromatography steps such as ion exchange, gel filtration, reverse phase, hydroxylapatite, and/or affinity chromatography; isoelectric focusing; gel 25 electrophoresis; and combinations of such and other techniques. As is generally known in the art, it is believed that the order of conducting the various purification steps may be changed, or that certain steps may be omitted, and still result in a suitable method for the preparation of a substantially purified protein or peptide.

There is no general requirement that the protein or peptide always be provided in 30 their most purified state. Indeed, it is contemplated that less substantially purified products will have utility in certain embodiments. Partial purification may be accomplished by using fewer purification steps in combination, or by utilizing different

forms of the same general purification scheme. For example, it is appreciated that a cation-exchange column chromatography performed utilizing an HPLC apparatus will generally result in a greater-fold purification than the same technique utilizing a low pressure chromatography system. Methods exhibiting a lower degree of relative 5 purification may have advantages in total recovery of protein product, or in maintaining the activity of an expressed protein.

E. Preparation of Antibodies Specific for sGPCRs

For some embodiments, it will be desired to produce antibodies that bind with high specificity to the protein product(s) of an isolated nucleic acid encoding for sGPCR, 10 including but not limited to sCRFR2 α . In certain aspects, an antibody preparation is contemplated that recognizes or binds the c-terminus of a GPCR, particularly a splice variant such as a sCRFR2 α splice variant and thus can be used to distinguish a sGPCR polypeptide from a membrane associated receptor. Such antibodies may be used in any of a variety of applications known to those of skill in the art, including but not limited to: 15 immunodetection methods, immunoprecipitation methods, ELISA assays, protein purification methods, *etc.* Means for preparing and characterizing antibodies are well known in the art (*See, e.g.*, Harlow and Lane, 1988, incorporated herein by reference).

Methods for generating polyclonal antibodies are well known in the art. Briefly, a polyclonal antibody is prepared by immunizing an animal with an antigenic 20 composition and collecting antisera from that immunized animal. A wide range of animal species can be used for the production of antisera. Typically the animal used for production of antisera is a rabbit, a mouse, a rat, a hamster, a guinea pig, a horse, or a goat. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies.

As is well known in the art, a given composition may vary in its 25 immunogenicity. It is often necessary therefore to boost the host immune system, as may be achieved by coupling a peptide or polypeptide immunogen to a carrier. Exemplary and preferred carriers are keyhole limpet hemocyanin (KLH) and bovine serum albumin (BSA). Other albumins such as ovalbumin, mouse serum albumin or 30 rabbit serum albumin can also be used as carriers. Means for conjugating a polypeptide to a carrier protein are well known in the art and include glutaraldehyde,

m-maleimidobenzoyl-N-hydroxysuccinimide ester, carbodiimide and bis-biazotized benzidine.

As is also well known in the art, the immunogenicity of a particular immunogen composition can be enhanced by the use of non-specific stimulators of the immune response, known as adjuvants. Exemplary and preferred adjuvants include complete Freund's adjuvant (a non-specific stimulator of the immune response containing killed *Mycobacterium tuberculosis*), incomplete Freund's adjuvants and aluminum hydroxide adjuvant.

Monoclonal antibodies (MAbs) may be readily prepared through use of well-known techniques, such as those exemplified in U.S. Patent 4,196,265, incorporated herein by reference. Typically, this technique involves immunizing a suitable animal with a selected immunogen composition, *e.g.*, a purified or partially purified expressed protein, polypeptide or peptide. The immunizing composition is administered in a manner that effectively stimulates antibody producing cells.

The animals are injected with antigen as described above. Following immunization, somatic cells with the potential for producing antibodies, specifically B lymphocytes (B cells), are selected for use in the MAb generating protocol. Often, a panel of animals will have been immunized and the spleen of the animal with the highest antibody titer will be removed and the spleen lymphocytes obtained by homogenizing the spleen with a syringe.

The antibody-producing B lymphocytes from the immunized animal are then fused with cells of an immortal myeloma cell, generally one of the same species as the animal that was immunized. Myeloma cell lines suited for use in hybridoma-producing fusion procedures preferably are non-antibody-producing, have high fusion efficiency, and have enzyme deficiencies that render them incapable of growing in certain selective media that support the growth of only the desired fused cells (hybridomas). Any one of a number of myeloma cells may be used, as are known to those of skill in the art (Goding, 1986). For example, where the immunized animal is a mouse, one may use P3-X63/Ag8, X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 and S194/5XX0 Bul; for rats, one may use R210.RCY3, Y3-Ag 1.2.3, IR983F and 4B210; and U-266, GM1500-GRG2, LICR-LON-HMy2 and UC729-6 are all useful in connection with human cell fusions.

One preferred murine myeloma cell is the NS-1 myeloma cell line (also termed P3-NS-1-Ag4-1), which is readily available from the NIGMS Human Genetic Mutant Cell Repository by requesting cell line repository number GM3573. Another mouse myeloma cell line that may be used is the 8-azaguanine-resistant mouse murine 5 myeloma SP2/0 non-producer cell line.

Large amounts of the monoclonal antibodies of the present invention may also be obtained by multiplying hybridoma cells *in vivo*. Cell clones are injected into mammals that are histocompatible with the parent cells, *e.g.*, syngeneic mice, to cause growth of antibody-producing tumors. Optionally, the animals are primed with a hydrocarbon, 10 especially oils such as pristane (tetramethylpentadecane) prior to injection.

In accordance with the present invention, fragments of a monoclonal antibody can be obtained by methods which include digestion with enzymes such as pepsin or papain and/or cleavage of disulfide bonds by chemical reduction. Alternatively, 15 monoclonal antibody fragments encompassed by the present invention can be synthesized using an automated peptide synthesizer, or by expression of full-length polynucleotide or of polynucleotide fragments encoding all or part of Mab.

Antibody conjugates may be prepared by methods known in the art, *e.g.*, by reacting an antibody with an enzyme in the presence of a coupling agent such as glutaraldehyde or periodate. Conjugates with fluorescein markers are prepared in the 20 presence of these coupling agents or by reaction with an isothiocyanate. Conjugates with metal chelates are similarly produced. Other moieties to which antibodies may be conjugated include radionuclides such as ^3H , ^{125}I , ^{131}I , ^{32}P , ^{35}S , ^{14}C , ^{51}Cr , ^{36}Cl , ^{57}Co , ^{58}Co , ^{59}Fe , ^{75}Se , ^{152}Eu , and $^{99\text{m}}\text{Tc}$. Radioactively labeled antibodies of the present invention can be produced according to well-known methods. For instance, antibodies can be 25 iodinated by contact with sodium or potassium iodide and a chemical oxidizing agent such as sodium hypochlorite, or an enzymatic oxidizing agent, such as lactoperoxidase. Antibodies according to the invention may be labeled with technetium- 99 by ligand exchange process, for example, by reducing pertechnate with stannous solution, chelating the reduced technetium onto a Sephadex column and applying the antibody to 30 this column or by direct labelling techniques, *e.g.*, by incubating pertechnate, a reducing agent such as SnCl_2 , a buffer solution such as sodium-potassium phthalate solution, and the antibody.

III. NUCLEIC ACIDS ENCODING sGPCR POLYPEPTIDES

The present invention includes nucleic acids that encode all or part of a sGPCR, such as but not limited to a GPCR, a family B GPCR, a family B1 GPCR, a CRFR, or a CRFR2 polypeptide, and may include various nucleic acid sequences needed for 5 delivery of the nucleic acid sequence as well as the transcription and/or translation of the nucleic acid sequence. Nucleic acid molecules of the invention may include various contiguous stretches of the nucleic acid, for example about 10, 15, 17, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1250, 1500, 1750, 2000, 2100, including all or part of the full 10 length nucleic acid sequences in the sequence listing, e.g., SEQ ID NO:1, 3, 5, 7, 9, 11, 13, or 14, or polynucleotides of those GPCRs referenced herein, fragments thereof, mRNAs, or cDNAs comprising sequences described or referenced herein, and mutants of each are contemplated. Also contemplated are molecules that are complementary to the above mentioned sequences and that bind to these sequences under high stringency 15 conditions. These probes will be useful in a variety of hybridization embodiments, such as Southern and northern blotting.

Various probes and primers can be designed around the disclosed nucleotide sequences. Primers may be of any length but, typically, are 10-20 bases in length. In particular aspects, the probe or primer can be used to identify or screen for the presence 20 of an alternatively spliced form of a GPCR, such as but not limited to the CRFR2 gene that includes an exon 5/exon 7 splice junction (may also be described as an exon 3/exon 5 junction as it relates to CRFR2 α transcription). These probes or primers may either hybridize unique sequence of the engineered nucleic acid or splice junction, or amplify a nucleic acid characteristic of the engineered nucleic acid or the splice junction. By 25 assigning numeric values to a sequence, for example, the first residue is 1, the second residue is 2, *etc.*, an algorithm defining all primers can be proposed:

n to n + y

where n is an integer from 1 to the last number of the sequence and y is the length of the primer minus one, where n + y does not exceed the last number of the sequence. Thus, 30 for a 10-mer, the probes correspond to bases 1 to 10, 2 to 11, 3 to 12 ... and so on. For a 15-mer, the probes correspond to bases 1 to 15, 2 to 16, 3 to 17 ... and so on. For a 20-mer, the probes correspond to bases 1 to 20, 2 to 21, 3 to 22 ... and so on.

In certain aspects the nucleic acid sequences of the invention may be used to encode various polypeptides described herein. In one embodiment of the present invention, the nucleic acid sequences may be used as hybridization probes or amplification primers. In certain embodiments, these probes and primers consist of 5 oligonucleotide fragments. Such fragments should be of sufficient length to provide specific hybridization to an RNA or DNA sample extracted from tissue. The sequences typically will be 10-20 nucleotides, but may be longer. Longer sequences, *e.g.*, 40, 50, 100, 500 and even up to full length, are preferred for certain embodiments.

The use of a hybridization probe of between 17 and 100 nucleotides in length 10 allows the formation of a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 20 bases in length are generally preferred, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of particular hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having stretches of 20 to 30 15 nucleotides, or even longer where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means or by introducing selected sequences into recombinant vectors for recombinant production. Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of genes, polynucleotides or 20 RNAs, or to provide primers for amplification of DNA or RNA from tissues. Depending on the application envisioned, one will desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence.

For applications requiring high selectivity, one will typically desire to employ 25 relatively stringent or high stringency conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.10 M NaCl at temperatures of about 50°C to about 70°C. Such high stringency conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating specific genes or detecting specific mRNA transcripts. It is generally appreciated that conditions can be 30 rendered more stringent by the addition of increasing amounts of formamide.

In certain embodiments, it will be advantageous to employ nucleic acid sequences of the present invention in combination with an appropriate means of detection, such as a fluorescent or radiolabel, for determining hybridization. A wide

variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of being detected.

For applications in which the nucleic acid segments of the present invention are 5 incorporated into expression vectors, such as plasmids, cosmids or viral polynucleotides, these segments may be combined with other DNA sequences, such as promoters, polyadenylation signals, restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is contemplated that a nucleic acid fragment of almost any length may be employed, with 10 the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

DNA segments encoding a specific polynucleotide may be introduced into recombinant host cells and employed for expressing a sGPCR, such as but not limited to a sCRFR2 α polypeptide. Alternatively, through the application of genetic engineering 15 techniques, subportions or derivatives of selected polynucleotides may be employed.

Throughout this application, the term “expression construct” is meant to include any type of genetic construct containing a nucleic acid having a sequence defining a product, such as but not limited to a product encoding a polypeptide, in which part or all of the nucleic acid sequence is capable of being transcribed. The transcript may be 20 translated into a protein, but it need not be. Thus, in certain embodiments, expression includes both transcription of a polynucleotide and translation of a RNA into a polypeptide product.

In preferred embodiments, the nucleic acid is under transcriptional control of a promoter. A “promoter” refers to a DNA sequence recognized by the synthetic 25 machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a polynucleotide. The phrase “under transcriptional control” means that the promoter is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the polynucleotide. The term promoter will be used here to refer to a group of transcriptional control modules that are 30 clustered around the initiation site for a RNA polymerase, in particular RNA polymerase II. In certain aspects, at least one module in each promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some

promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 late genes, a discrete element overlying the start site itself helps to fix the place of initiation.

5 The particular promoter that is employed to control the expression of a nucleic acid is not believed to be critical, so long as it is capable of expressing the nucleic acid in the targeted cell. Thus, where a human cell is targeted, it is preferable to position the nucleic acid coding region adjacent to and under the control of a promoter that is capable of being expressed in a human cell. Generally speaking, such a promoter might include either a human or viral promoter.

10 In various other embodiments, the human cytomegalovirus (CMV) immediate early gene promoter, the SV40 early promoter and the Rous sarcoma virus long terminal repeat can be used to obtain high-level expression of transgenes. The use of other viral or mammalian cellular or bacterial phage promoters which are well-known in the art to achieve expression of a transgene is contemplated as well, provided that the levels of expression are sufficient for a given purpose. Several elements/promoters, as described below, may be employed, in the context of the present invention, to regulate the expression of a polynucleotide, such as a transgene. This list is not intended to be exhaustive of all the possible elements involved in the promotion of transgene expression but, merely, to be exemplary thereof.

15 20 Any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of a polynucleotide. Use of a T3, T7 or SP6 cytoplasmic expression system is another possible embodiment. Eukaryotic cells can support cytoplasmic transcription from certain bacterial promoters if the appropriate bacterial polymerase is provided, either as part of the delivery complex or as an additional genetic expression construct. Use of the baculovirus system will involve high level expression from the powerful polyhedrin promoter.

25 Promoters include, but are not limited to Immunoglobulin Heavy Chain, Immunoglobulin Light Chain, T-Cell Receptor, HLA DQ α and DQ β , β -Interferon, Interleukin-2, Interleukin-2 Receptor, MHC Class II 5, MHC Class II HLA-DR α , β -Actin, Muscle Creatine Kinase, Prealbumin (Transthyretin), Elastase I, Metallothionein, Collagenase, Albumin Gene, α -Fetoprotein, α -Globin, β -Globin, c-fos, c-HA-ras, Insulin, Neural Cell Adhesion Molecule (NCAM), α_1 -Anti-trypsin, H2B (TH2B)

Histone, Mouse or Type I Collagen, Glucose-Regulated Proteins (GRP94 and GRP78), Rat Growth Hormone, Human Serum Amyloid A (SAA), Troponin I (TN I), Platelet-Derived Growth Factor, Duchenne Muscular Dystrophy, SV40, Polyoma, Retroviruses, Papilloma Virus, Hepatitis B Virus, Human Immunodeficiency Virus, Cytomegalovirus,
5 Gibbon Ape Leukemia Virus.

Various element (inducers) include, but are not limited to MT II (Phorbol Ester (TPA)Heavy metals); MMTV (Glucocorticoids, β -Interferon, poly(rI)X, poly(rc)); Adenovirus 5 E2 (Ela); c-jun (Phorbol Ester (TPA), H₂O₂); Collagenase (Phorbol Ester (TPA)); Stromelysin (Phorbol Ester (TPA), IL-1); SV40 (Phorbol Ester (TPA)); Murine
10 MX Gene (Interferon, Newcastle Disease Virus); GRP78 Gene (A23187); α -2-Macroglobulin (IL-6); Vimentin (Serum); MHC Class I Gene H-2kB (Interferon); HSP70 (Ela, SV40 Large T Antigen); Proliferin (Phorbol Ester-TPA); Tumor Necrosis Factor (FMA); and Thyroid Stimulating Hormone α Gene (Thyroid Hormone).

One will typically include a polyadenylation signal to effect proper
15 polyadenylation of the transcript. The nature of the polyadenylation signal is not believed to be crucial to the successful practice of the invention, and any such sequence may be employed. Preferred embodiments include the SV40 polyadenylation signal and the bovine growth hormone polyadenylation signal, convenient and known to function well in various target cells. Also contemplated is the inclusion of a terminator as an
20 element of an expression cassette. These elements can serve to enhance message levels and to minimize read through from the cassette into other sequences.

In various embodiments of the invention, an expression construct may comprise a virus or engineered construct derived from a viral genome. The ability of certain viruses to enter cells via receptor-mediated endocytosis and to integrate into the host cell
25 genome and express viral genes stably and efficiently have made them attractive candidates for the transfer of foreign genes into mammalian cells (Ridgeway, 1988; Nicolas and Rubenstein, 1988; Baichwal and Sugden, 1986; Temin, 1986). The first viruses used as vectors were DNA viruses including the papovaviruses (simian virus 40, bovine papilloma virus, and polyoma) (Ridgeway, 1988; Baichwal and Sugden, 1986)
30 and adenoviruses (Ridgeway, 1988; Baichwal and Sugden, 1986) and adeno-associated viruses. Retroviruses also are attractive gene transfer vehicles (Nicolas and Rubenstein, 1988; Temin, 1986) as are vaccinia virus (Ridgeway, 1988) and adeno-associated virus (Ridgeway, 1988). Such vectors may be used to (i) transform cell lines *in vitro* for the

purpose of expressinG-proteins of interest or (ii) to transform cells *in vitro* or *in vivo* to provide therapeutic polypeptides in a gene therapy scenario.

In an alternative embodiment, the sGPCR encoding nucleic acids employed may actually encode antisense constructs that hybridize, under intracellular conditions, to a 5 sGPCR or other encoding nucleic acid. The term “antisense construct” is intended to refer to nucleic acids, preferably oligonucleotides, complementary to the base sequences of a target DNA or RNA.

As used herein, the terms “complementary” means nucleic acid sequences that are substantially complementary over their entire length and have very few base 10 mismatches. For example, nucleic acid sequences of fifteen bases in length may be termed complementary when they have a complementary nucleotide at thirteen or fourteen positions with only a single mismatch. Naturally, nucleic acid sequences which are “completely complementary” will be nucleic acid sequences which are entirely complementary throughout their entire length and have no base mismatches.

15 **A. Detection and Quantitation of Nucleic Acids**

One embodiment of the instant invention comprises a method for identification of sGPCR nucleic acid, such as but not limited to CRFR2 α nucleic acids, in a biological sample by amplifying and detecting nucleic acids corresponding to sGPCR. The biological sample can be any tissue or fluid in which the polynucleotide might be 20 present. Nucleic acid used as a template for amplification is isolated from cells contained in the biological sample, according to standard methodologies (Sambrook *et al.*, 1989). The nucleic acid may be fractionated or whole cell RNA.

Pairs of primers that selectively hybridize to nucleic acids corresponding to sGPCR are contacted with the isolated nucleic acid under conditions that permit 25 selective hybridization. Once hybridized, the nucleic acid:primer complex is contacted with one or more enzymes that facilitate template-dependent nucleic acid synthesis. Multiple rounds of amplification, also referred to as “cycles,” are conducted until a sufficient amount of amplification product is produced. The amplification products may be detected. In certain applications, the detection may be performed by visual means. 30 Alternatively, the detection may involve indirect identification of the product via chemiluminescence, radioactive scintigraphy of incorporated radiolabel or fluorescent

label, or even via a system using electrical or thermal impulse signals (Affymax technology; Bellus, 1994).

A number of template dependent processes are available to amplify the marker sequences present in a given template sample. One of the best known amplification 5 methods is the polymerase chain reaction (referred to as PCR) which is described in detail in U.S. Patents 4,683,195, 4,683,202 and 4,800,159, and in Innis *et al.*, 1990, each of which is incorporated herein by reference in its entirety. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction ("LCR"), disclosed 10 in EPA No. 320 308, which is incorporated herein by reference in its entirety. U.S. Patent 4,883,750 describes a method similar to LCR for binding probe pairs to a target sequence. Also, Qbeta Replicase, described in PCT Application No. PCT/US87/00880, may be used as still another amplification method in the present invention. An isothermal amplification method, in which restriction endonucleases and ligases are used 15 to achieve the amplification of target molecules that contain nucleotide 5'-[alpha-thio]-triphosphates in one strand of a restriction site may also be useful in the amplification of nucleic acids in the present invention, Walker *et al.*, (1992), incorporated herein by reference in its entirety. Still further, Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves 20 multiple rounds of strand displacement and synthesis, *i.e.*, nick translation. A similar method, called Repair Chain Reaction (RCR), involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. Target specific sequences can also be detected using a cyclic probe reaction (CPR). Still another amplification method described in GB 25 Application No. 2 202 328, and in PCT Application No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. Other nucleic acid amplification procedures include transcription-based amplification systems (TAS), including nucleic acid sequence based amplification (NASBA) and 3SR (Kwoh *et al.*, 1989); PCT Application WO 88/10315, incorporated 30 herein by reference in their entirety).

Following amplification, it may be desirable to separate the amplification product from the template and the excess primer for the purpose of determining whether specific amplification has occurred. In one embodiment, amplification products are

separated by agarose, agarose-acrylamide or polyacrylamide gel electrophoresis using standard methods. See Sambrook *et al.*, 1989.

Chromatographic techniques may be employed to effect separation. There are many kinds of chromatography which may be used in the present invention: adsorption, 5 partition, ion-exchange and molecular sieve, and many specialized techniques for using them including column, paper, thin-layer and gas chromatography (Freifelder, 1982).

IV. METHODS FOR sGPCR GENE EXPRESSION

In one embodiment of the present invention, there are provided methods for increased sGPCR expression in a cell, such as but not limited to sCRFR2 α expression. 10 This is particularly useful where there is an aberration in the protein or protein expression is not sufficient for normal function. This will allow for the alleviation of symptoms of disease experienced as a result of deficiency of sGPCR, hyperactivation of GPCR or an abundance of GPCR ligand.

The general approach to increasing sGPCR is to contact or administer to a cell, 15 tissue, animal, or subject a sGPCR polypeptide. While it is preferred that the protein may be delivered directly, a conceivable embodiment involves providing a nucleic acid encoding a sGPCR polypeptide to the cell or neighboring cells. Following this provision, the sGPCR polypeptide is synthesized by the host cell's transcriptional and translational machinery, as well as any that may be provided by the expression construct. 20 *Cis*-acting regulatory elements necessary to support the expression of the sGPCR polynucleotide will be provided, in the form of an expression construct. It also is possible that expression of virally-encoded sGPCR could be stimulated or enhanced, or the expressed polypeptide be stabilized, thereby achieving the same or similar effect.

In order to effect expression of constructs encoding sGPCR polynucleotides, the 25 expression construct must be delivered by a delivery vector into a cell. One mechanism for delivery is via viral infection, where the expression construct is encapsidated in a viral particle which will deliver either a replicating or non-replicating nucleic acid.

The ability of certain viruses to enter cells via receptor-mediated endocytosis, to integrate into host cell genome and express viral genes stably and efficiently have made 30 them attractive candidates for the transfer of foreign genes into mammalian cells (Ridgeway, 1988; Nicolas and Rubenstein, 1988; Baichwal and Sugden, 1986; Temin, 1986). The first viruses used as gene vectors were DNA viruses including the

papovaviruses (simian virus 40, bovine papilloma virus, and polyoma) (Ridgeway, 1988; Baichwal and Sugden, 1986) and adenoviruses (Ridgeway, 1988; Baichwal and Sugden, 1986). These have a relatively low capacity for foreign DNA sequences and have a restricted host spectrum. Furthermore, their oncogenic potential and cytopathic effects 5 in permissive cells raise safety concerns. They can accommodate only up to 8 kb of foreign genetic material but can be readily introduced in a variety of cell lines and laboratory animals (Nicolas and Rubenstein, 1988; Temin, 1986).

The retroviruses are a group of single-stranded RNA viruses characterized by an 10 ability to convert their RNA to double-stranded DNA in infected cells; they can also be used as vectors. Other viral vectors may be employed as expression constructs in the present invention. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Baichwal and Sugden, 1986; Coupar *et al.*, 1988) adeno-associated virus (AAV) (Ridgeway, 1988; Baichwal and Sugden, 1986; Hermonat and Muzycska, 1984) and herpesviruses may be employed. They offer several attractive features for various 15 mammalian cells (Friedmann, 1989; Ridgeway, 1988; Baichwal and Sugden, 1986; Coupar *et al.*, 1988; Horwitz *et al.*, 1990).

Several non-viral methods for the transfer of expression constructs into cultured 20 mammalian cells also are contemplated by the present invention. These include calcium phosphate precipitation (Graham and Van Der Eb, 1973; Chen and Okayama, 1987; Rippe *et al.*, 1990) DEAE-dextran (Gopal, 1985), electroporation (Tur-Kaspa *et al.*, 1986; Potter *et al.*, 1984), direct microinjection (Harland and Weintraub, 1985), DNA-loaded liposomes (Nicolau and Sene, 1982; Fraley *et al.*, 1979) and lipofectamine-DNA 25 complexes, cell sonication (Fechheimer *et al.*, 1987), gene bombardment using high velocity microprojectiles (Yang *et. al.*, 1990), and receptor-mediated transfection (Wu and Wu, 1987; Wu and Wu, 1988). Some of these techniques may be successfully adapted for *in vivo* or *ex vivo* use, as discussed below.

In another embodiment of the invention, the expression construct may simply 30 consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro*, but it may be applied to *in vivo* use as well. Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA coated

microparticles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 5 1990). The microparticles used have consisted of biologically inert substances such as tungsten or gold beads.

In a further embodiment of the invention, the expression construct may be entrapped in a liposome. Liposomes are vesicular structures characterized by a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes 10 have multiple lipid layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers (Ghosh and Bachhawat, 1991). Also contemplated are lipofectamine-DNA complexes.

15 Other expression constructs which can be employed to deliver a nucleic acid encoding a sCRFR2 α polynucleotide into cells are receptor-mediated delivery vehicles. These take advantage of the selective uptake of macromolecules by receptor-mediated endocytosis in almost all eukaryotic cells. Because of the cell type-specific distribution of various receptors, the delivery can be highly specific (Wu and Wu, 1993).

20 **V. PHARMACEUTICALS AND METHODS FOR THE TREATMENT OF DISEASE**

In additional embodiments, the present invention concerns formulation of one or 25 more of the polynucleotide, polypeptide, and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell, tissue, animal, patient, or subject either alone, or in combination with one or more other modalities of therapy.

Aqueous pharmaceutical compositions of the present invention will have an effective amount of a sGPCR expression construct, an expression construct that encodes 30 a therapeutic gene along with sGPCR, or a sGPCR protein and/or compound that modulates GPCR ligand activity or sensitivity, or other endocrine function. Such compositions generally will be dissolved or dispersed in a pharmaceutically acceptable carrier or aqueous medium. An "effective amount," for the purposes of therapy, is

defined at that amount that causes a clinically measurable difference in the condition of the subject. This amount will vary depending on the substance, the condition of the patient, the type of treatment, *etc.*

The phrases "pharmaceutically or pharmacologically acceptable" refer to 5 molecular entities and compositions that do not produce a significant adverse, allergic or other untoward reaction when administered to an animal, or human. As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutically active substances is 10 well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredients, its use in the therapeutic compositions is contemplated.

In addition to the compounds formulated for parenteral administration, such as 15 those for intravenous or intramuscular injection, other pharmaceutically acceptable forms include, *e.g.*, tablets or other solids for oral administration; time release capsules; and any other form currently used, including creams, lotions, inhalants and the like.

The active compounds of the present invention will often be formulated for parenteral administration, *e.g.*, formulated for injection *via* the intravenous, intramuscular, subcutaneous, or even intraperitoneal routes. The preparation of an 20 aqueous composition that contains sGPCR alone or in combination with a conventional therapeutic agent as active ingredients will be known to those of skill in the art in light of the present disclosure. Typically, such compositions can be prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for using to prepare solutions or suspensions upon the addition of a liquid prior to injection can also be 25 prepared; and the preparations can also be emulsified.

The pharmaceutical forms suitable for injectable use include sterile aqueous 30 solutions or dispersions; formulations including sesame oil, peanut oil or aqueous propylene glycol; and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In many cases, the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi.

The carrier also can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin, by the 5 maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption 10 of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions 15 are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any 20 additional desired ingredient from a previously sterile-filtered solution thereof.

Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms, such as the type of injectable solutions described above, with even drug release capsules and the like being 25 employable.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this 30 connection, sterile aqueous media which can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage could be dissolved in 1 mL of isotonic NaCl solution and either added to 1000 mL of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's

Pharmaceutical Sciences" (1980)). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject.

5 In certain aspects of the methods of the invention, the route the therapeutic composition is administered may be by parenteral administration. The parenteral administration may be intravenous injection, subcutaneous injection, intramuscular injection, intramedullary injection, ingestion or a combination thereof. In certain aspects, the composition comprising sGPCR is administered from about 0.1 to about 10
10 microgram/kg/body weight per dose. In certain aspects, the composition comprising sGPCR is administered from about 1 to about 5 microgram/kg/body weight per dose. In certain aspects, the composition comprising sGPCR is administered from about 1.2 to about 3.6 microgram/kg/body weight per dose. In certain aspects, the composition comprising sGPCR is administered from about 1.2 to about 2.4 microgram/kg/body
15 weight per dose. In preferred aspects, the amount of sGPCR administered per dose may be about 0.1, about 0.2, about 0.3, about 0.4, about 0.5, about 0.6, about 0.7, about 0.8, about 0.9, about 1.0, about 1.1, about 1.2, about 1.3, about 1.4, about 1.5, about 1.6, about 1.7, about 1.8, about 1.9, about 2.0, about 2.1, about 2.2, about 2.3, about 2.4, about 2.5, about 2.6, about 2.7, about 2.8, about 2.9, about 3.0, about 3.1, about 3.2, about
20 3.3, about 3.4, about 3.5, about 3.6, about 3.7, about 3.8, about 3.9, about 4.0, about 4.1, about 4.2, about 4.3, about 4.4, about 4.5, about 4.6, about 4.7, about 4.8, about 4.9, about 5.0, about 5.1, about 5.2, about 5.3, about 5.4, about 5.5, about 5.6, about 5.7, about 5.8, about 5.9, about 6.0, about 6.1, about 6.2, about 6.3, about 6.4, about 6.5, about 6.6, about 6.7, about 6.8, about 6.9, about 7.0, about 7.1, about 7.2, about 7.3,
25 about 7.4, about 7.5, about 7.6, about 7.7, about 7.8, about 7.9, about 8.0, about 8.1, about 8.2, about 8.3, about 8.4, about 8.5, about 8.6, about 8.7, about 8.8, about 8.9, about 9.0, about 9.1, about 9.2, about 9.3, about 9.4, about 9.5, about 9.6, about 9.7, about 9.8, about 9.9, about 10.0, or more micrograms/kg/body.

30 Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including e.g., oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

A. Alimentary Delivery

The term "alimentary delivery" refers to the administration, directly or otherwise, to a portion of the alimentary canal of an animal. The term "alimentary canal" refers to the tubular passage in an animal that functions in the digestion and absorption of food

5 and the elimination of food residue, which runs from the mouth to the anus, and any and all of its portions or segments, *e.g.*, the oral cavity, the esophagus, the stomach, the small and large intestines and the colon, as well as compound portions thereof such as, *e.g.*, the gastro-intestinal tract. Thus, the term "alimentary delivery" encompasses several routes of administration including, but not limited to, oral, rectal, endoscopic and

10 sublingual/buccal administration. A common requirement for these modes of administration is absorption over some portion or all of the alimentary tract and a need for efficient mucosal penetration of the nucleic acid(s) so administered.

1. Oral Delivery

In certain applications, the pharmaceutical compositions disclosed herein may be

15 delivered via oral administration to an animal, patient, or subject. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft- shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active components may even be incorporated with excipients and used in the

20 form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U.S. Patents 5,641,515; 5,580,579 and 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as

25 dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other

30 materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup or elixir may contain the active component sucrose as a sweetening agent

methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and 5 formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active 10 compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of 15 dosages and treatment regimens may be desirable.

2. Rectal Administration

Therapeutics administered by the oral route can often be alternatively administered by the lower enteral route, *i.e.*, through the anal portal into the rectum or lower intestine. Rectal suppositories, retention enemas or rectal catheters can be used for 20 this purpose and may be preferred when patient compliance might otherwise be difficult to achieve (*e.g.*, in pediatric and geriatric applications, or when the patient is vomiting or unconscious). Rectal administration may result in more prompt and higher blood levels than the oral route, but the converse may be true as well (Harvey, 1990). Because about 50% of the therapeutic that is absorbed from the rectum will bypass the 25 liver, administration by this route significantly reduces the potential for first-pass metabolism (Benet *et al.*, 1996).

B. Parenteral Delivery

The term "parenteral delivery" refers to the administration of a therapeutic of the invention to an animal, patient or subject in a manner other than through the digestive 30 canal. Means of preparing and administering parenteral pharmaceutical compositions are known in the art (see, *e.g.*, Avis, 1990).

C. Intraluminal administration

Intraluminal administration, for the direct delivery of a therapeutic to an isolated portion of a tubular organ or tissue (e.g., such as an artery, vein, ureter or urethra), may be desired for the treatment of patients with diseases or conditions afflicting the lumen of 5 such organs or tissues. To effect this mode of administration, a catheter or cannula is surgically introduced by appropriate means. After isolation of a portion of the tubular organ or tissue for which treatment is sought, a composition comprising a therapeutic of the invention is infused through the cannula or catheter into the isolated segment. After 10 incubation for from about 1 to about 120 minutes, during which the therapeutic is taken up or in contact with the cells of the interior lumen of the vessel, the infusion cannula or catheter is removed and flow within the tubular organ or tissue is restored by removal of the ligatures which effected the isolation of a segment thereof (Morishita *et al.*, 1993). Therapeutic compositions of the invention may also be combined with a biocompatible 15 matrix, such as a hydrogel material, and applied directly to vascular tissue *in vivo*.

15 D. Intraventricular administration

Intraventricular administration, for the direct delivery of a therapeutic to the brain of a patient, may be desired for the treatment of patients with diseases or conditions afflicting the brain. One method to affect this mode of administration, a silicon catheter is surgically introduced into a ventricle of the brain of a human patient, and is connected 20 to a subcutaneous infusion pump (Medtronic Inc., Minneapolis, Minn.) that has been surgically implanted in the abdominal region (Zimm *et al.*, 1984; Shaw, 1993). The pump is used to inject the therapeutic and allows precise dosage adjustments and variation in dosage schedules with the aid of an external programming device. The reservoir capacity of the pump is 18-20 mL and infusion rates may range from 0.1 mL/h 25 to 1 mL/h. Depending on the frequency of administration, ranging from daily to monthly, and the dose of drug to be administered, ranging from 0.01 µg to 100 g per kg of body weight, the pump reservoir may be refilled at 3-10 week intervals. Refilling of the pump may be accomplished by percutaneous puncture of the self-sealing septum of the pump.

30 E. Intrathecal drug administration

Intrathecal drug administration, for the introduction of a therapeutic into the spinal column of a patient may be desired for the treatment of patients with diseases of

the central nervous system. To effect this route of administration, a silicon catheter may be surgically implanted into the L3-4 lumbar spinal interspace of a human patient, and is connected to a subcutaneous infusion pump which has been surgically implanted in the upper abdominal region (Luer and Hatton, 1993; Ettinger *et al.*, 1978; Yaida *et al.*, 1995). The pump is used to inject the therapeutic and allows precise dosage adjustments and variations in dose schedules with the aid of an external programming device. The reservoir capacity of the pump is 18-20 mL, and infusion rates may vary from 0.1 mL/h to 1 mL/h. Depending on the frequency of drug administration, ranging from daily to monthly, and dosage of drug to be administered, ranging from 0.01 µg to 100 g per kg of body weight, the pump reservoir may be refilled at 3-10 week intervals. Refilling of the pump is accomplished by a single percutaneous puncture to the self-sealing septum of the pump.

To effect delivery to areas other than the brain or spinal column via this method, the silicon catheter is configured to connect the subcutaneous infusion pump to, *e.g.*, the hepatic artery, for delivery to the liver (Kemeny *et al.*, 1993).

F. Vaginal Delivery

Vaginal delivery provides local treatment and avoids first pass metabolism, degradation by digestive enzymes, and potential systemic side-effects. Vaginal suppositories (Block, Chapter 87 In: Remington's Pharmaceutical Sciences, 18th Ed., Gennaro, ed., Mack Publishing Co., Easton, Pa., 1990, pages 1609-1614) or topical ointments can be used to effect this mode of delivery.

G. Liposome-, Nanocapsule-, and Microparticle-Mediated Delivery

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Lasic, 1998; which describes the use of liposomes and

nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U.S. Patent 5,741,516, specifically incorporated herein by reference in its entirety).
5 Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran et al., 1997; Margalit, 1995; U.S. Patent 5,567,434; 5,552,157; 5,565,213; 5,738,868 and 5, 795,587, each specifically incorporated herein by reference in its entirety).

10 Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μm . Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 \AA , containing an aqueous solution in the core.

15 The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places
20 where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood,
25 liver, spleen, bone marrow, and lymphoid organs.

30 Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland et al., 1987; Quintanar-Guerrero et al., 1998; Douglas et al., 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be are easily made, as described

(Couvreur et al., 1980; 1988; zur Muhlen et al., 1998; Zambaux et al. 1998; Pinto-Alphandry et al. , 1995 and U.S. Pat. No. 5,145,684, specifically incorporated herein by reference in its entirety).

EXAMPLES

5 The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the present invention in any fashion. One skilled in the art will appreciate readily that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those objects, ends and advantages inherent herein. The present examples, along with 10 the cells and methods described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses which are encompassed within the spirit of the invention as defined by the scope of the claims will occur to those skilled in the art.

A. Materials and Methods

15 **Isolation of the mouse soluble CRFR2 α cDNA.** The soluble CRFR2a splice variant was isolated in parallel with that of the mouse CRFR2 α ortholog. PCR primers were designed based on the homology between known mammalian CRFR2 genes. The following oligonucleotide primers, 5' CCCCGAAGCTGCCGACTGG 3' (SEQ ID NO:16) (sense) and 5' GGAAGGCTGTAAAGGATGGAGAAG 3' (SEQ ID NO:17) 20 (antisense) were used to screen cDNA prepared from mouse whole brain poly(A)+ RNA which was reverse transcribed using oligo dT or random primers. PCR was performed at 62°C for 35 cycles with 90 sec extension at 72°C. The amplified fragments were subcloned into pCRITOPO vector (Invitrogen, Carlsbad, CA), sequenced, and found to encode the full-length CRFR2 α novel splice variant lacking exon six, sCRFR2 α (Chen 25 *et al.*, 2005).

30 **Semi-quantitative RT-PCR and Southern analysis.** The following mouse peripheral and CNS tissues were dissected and directly subjected to total RNA isolation as previously described (Chen *et al.*, 2005): total brain, olfactory bulb, hypothalamus, cortex, cerebellum, hippocampus, midbrain, pons/medulla oblongata, spinal cord and pituitary. The cDNA products were used as templates for semi-quantitative and RT PCR analysis using specific primers for CRFR2 α , sCRFR2 α and the ribosomal protein S 16.

The locations of the oligonucleotide primers at exons three and seven result in the amplification of two products of 418 and 309 corresponding to CRFR2 α and sCRFR2 α , respectively. Oligonucleotide primers sequence and PCR conditions can be found in the supporting text.

5 **Extracellular receptor kinase 1/2 (ERK1/2) assay.** CATH.a cells were equilibrated with DMEM supplemented with 1% (w/v) bovine serum albumin (BSA) for 6 hr and then stimulated with 0.1 % DMEM/BSA (vehicle) or 10 nM Ucn I in the presence or absence of 0.4 or 4 nM sCRFR2 α diluted in 0.1% DMEM/BSA. Cells were harvested immediately and analyzed for phosphorylated ERK1/2-p42, 44, as previously
10 described (Chen *et al.*, 2005).

15 **Transient transfections and luciferase assay.** The HEK293T cells were transfected with a luciferase reporter containing a fragment of the EVX1 gene containing a potent CRE site. The cells were harvested and the luciferase reporter activity was assayed as previously described (Chen *et al.*, 2005). Twenty hours posttransfection, cells were treated for 4h with vehicle or with Ucn 1 (0.0001-100 nM) in the presence or absence of 0.1 nM sCRFR2 α .

20 **Radio-Immuno Assays (RIA).** Antisera was raised in rabbits immunized with a synthetic peptide fragment encoding the unique C-terminal tail (aa 113-143) of mouse sCRFR2 α conjugated to Keyhole Limpet Hemocyanin using a protocol previously described for inhibin subunits (Vaughan *et al.*, 1989). The analog Tyr¹¹³ sCRFR2 α (113-143) was radiolabelled with Na¹²⁵I using chloramine-T and purified by HPLC (Vaughan *et al.*, 1989) for use as tracer in the HA. The procedure for sCRFR2 α RIA was similar to that previously described in detail for inhibin subunits (Vaughan *et al.*, 2005). Briefly, anti-sCRFR2 α was used at 1/300,000 final dilution and synthetic sCRFR2 α (113-143)
25 was used as standard. Murine tissues were acid extracted and partially purified using octadecyl silica cartridges as described (Vaughan *et al.*, 1989). Lyophilized samples were tested at three to seven dose levels. Free tracer was separated from bound by the addition of sheep anti-rabbit γ -globulins and 10% (wt/vol) polyethylene glycol. The EC50 and minimum detectable dose for sCRFR2 α are ~5 pg and 100 pg per tube,
30 respectively.

Immunohistochemistry. Adult male C57B6J mice (Jackson Laboratories) and Sprague-Dawley albino rats (Harlan Sprague-Dawley) were anesthetized with chloral

hydrate (350 mg/kg, ip) and perfused with Zamboni's fixative (Bittencourt *et al.*, 1999), followed by 0-4 hr. post-fixation. Regularly spaced (1-in-4) series of 30 μ m thick frontal sections throughout the brain were prepared for nickel-enhanced avidin-biotin-immunoperoxidase localization of sCRFR2 α -ir using Vectastain Elite reagents (Vector Laboratories, Burlingame, CA). Primary sCRFR2 α antisera were adsorbed against the carrier, affinity purified and used at a dilution of 1:2000. Specificity of immunostaining was evaluated using primary antisera preincubated overnight at 4°C with 0-300 μ M synthetic immunogen. Labeling was also evaluated in mutant mice deficient in either or both CRFRs (Smith *et al.*, 1998; Bale *et al.* 2000). Detailed description of the 10 fluorescence immunocytochemical analysis of COSM6 cells transfected with sCRFR2 α can be found in the supporting text.

15 **Mammalian expression of sCRFR2 α :** A cDNA corresponding to amino acids 1-143, modified by PCR to include a FLAG epitope following amino acid 143, was subcloned into pSec-Tag2 HygroA (Invitrogen, Carlsbad, CA) and used for transfection of COSM6 cells as described (Perrin *et al.*, 2001). After 4 days, the media was collected and sCRFR2 α was enriched by purification using FLAG-agarose (Sigma, St. Louis, MO) immunoaffinity chromatography. The protein was detected by immunoblot analysis using either the anti-FLAG antibody or the antibody generated to the unique sCRFR2 α C-terminus.

20 **Bacterial expression of sCRFR2 α :** A cDNA corresponding to amino acids 20-143 was generated by PCR using mCRFR2 α as the template. The cDNA was subcloned into pET-32a(+) (Novagen, La Jolla, CA) and the protein purified by S-protein affinity chromatography as described (Perrin *et al.*, 2001). The protein was detected by immunoblot analysis using the antibody generated to the unique sCRFR2 α C-terminus.

25 **Radioreceptor assays.** The soluble protein, purified either from COS M6 cell media or *E.coli* was incubated in triplicate wells with [125 I-DTyr 0]-astressin and increasing concentration of unlabeled peptides as described (Perrin *et al.*, 2003).

B. Results

30 A cDNA transcript of smaller (~100 bp) size was observed during the isolation of the mouse CRFR2 α (Van Pett *et al.*, 2000). This smaller fragment was isolated and found to encode a variant of CRFR2 α bearing a deletion of exon six. Translation of the

variant transcript predicts a novel 143 amino acid protein, sCRFR2 α , comprising the majority of the first extracellular domain of CRFR2 α followed by a unique 38 amino acid C terminus (FIG. 1A). Screening of GenBank showed homology of the C-terminus to no other protein. The genomic arrangement of the sCRFR2 α is shown in FIG. 1B.

5 If the sCRFR2 α mRNA is merely a product of splicing errors, it should be much less abundant than the correctly spliced RNA. In order to examine this question, semi-quantitative RT-PCR followed by Southern hybridization analysis was used to compare the relative abundance of CRFR2 α and sCRFR2 α mRNA in several brain regions. Total RNA prepared from mouse tissues was reverse-transcribed to generate cDNAs that were
10 used as templates for semi-quantitative RT-PCR analysis, followed by Southern hybridization, using specific primers and probes for CRFR2 α and sCRFR2 α (FIG. 2). The oligonucleotide primer pair (located in exons three and seven) allowed the simultaneous amplification of both the soluble form and the full-length membrane bound receptor in a single reaction (FIG. 2A). The sCRFR2 α is highly expressed in the
15 olfactory, cortex, midbrain and the pituitary (FIGs. 2B and 2C). Lower levels of expression were found in the hippocampus, hypothalamus, pons, medulla and spinal cord (FIGs. 2B and 2C). As shown in FIG. 2, the abundance of sCRFR2 α mRNA is lower, but comparable, to that of CRFR2 α mRNA. The sequences of cDNA fragments from RT-PCR were found to encode a splice variant of the mouse CRFR2 α gene (FIG.
20 1A).

Computer analysis of the sequence predicted that the first 19 amino acids serve as a putative signal peptide. Because the sequence contains no obvious sites for membrane attachment, the protein is hypothesized to be secreted as a soluble form. To explore this hypothesis, the protein was expressed in COS M6 cells. Following
25 purification from the media, a protein band of ~30 kD was visualized by immunoblot analyses using either anti-FLAG antiserum or the anti-sCRFR2 α , an antiserum raised against a synthetic peptide fragment encoding the unique C-terminal tail of sCRFR2 α protein (aa 113-143) (FIG. 3A). The larger size of the protein compared to that predicted from the cDNA is probably a result of glycosylation.

30 In order to obtain a larger quantity of sCRFR2 α , a protein lacking the putative signal peptide was expressed as a fusion protein in *E. coli* (Perrin *et al.*, 2001). Following cleavage and purification, the protein was visualized (using the anti-

sCRFR2 α) by immunoblot analysis as a narrow band of size ~20kD. The anti-sCRFR2 α serum detects the sCRFR2 α proteins both in radioimmunoassay (FIG. 3B) as well as in immunocytochemistry (FIG. 3C).

5 Immunohistochemical studies using anti-sCRFR2 α serum revealed the distribution of sCRFR2 α -ir in rodent brain. The cellular distribution of immunolabeling for sCRFR2 α -ir was widespread and conformed more closely to the location of CRFR1 mRNA expression pattern than to that of CRFR2 (FIGs. 4A-4F). The results described are from studies in mice; a similar pattern of labeling was observed in rats. Major sites of cellular expression include mitral and tufted cells of the olfactory bulb, the medial 10 septavdiagonal band complex, piriform cortex, substantia nigra, red nucleus, basolateral amygdaloid, deep cerebellar and dorsal column nuclei, all of which are prominent sites of CRFR1 expression. Similar to CRFR1, sCRFR2 α -ir cell bodies are numerous throughout isocortex, although the laminar distributions are only partly overlapping. Thus, while both CRFR1- and sCRFR2 α -expressing cell bodies are numerous in layer 15 2/3, the dominant cortical seat of CRFR1 expression is in layer 4, while that of sCRFR2 α is in layer 5. Major sites of CRFR2 expression, including the lateral septal, midbrain raphe, ventromedial hypothalamic and medial amygdaloid nuclei were all lacking in sCRFR2 α -stained cell bodies, although interestingly the latter two sites were among the few invested with labeled varicosities that the inventors take to be representative of 20 sCRFR2 α -ir terminal fields. The paraventricular nucleus of the hypothalamus also contained a presumed sCRFR2 α -ir terminal field of moderate density.

Labeling throughout the brain was blocked by pre-incubation of the antiserum with low micromolar concentrations ($\geq 30 \mu\text{M}$) of the sCRFR2 α (113-143) peptide used as immunogen; competition with the corresponding peptide predicted from the CRFR1 sequence did not interfere with immunolabeling at concentrations as high as 3 mM. Further support for the specificity of labeling are observations that all immunolocalizations persisted in CRFR1- and/or CRFR2-deleted mice; note that the targeting construct used for generating each of the existing receptor-knockout lines would be expected to spare the sCRFR2 α coding region (Smith *et al.*, 1998; Timpl *et al.*, 25 1998; Bale *et al.*, 2000).

In order to determine the presence of sCRFR2 α -like ir in brain, a highly specific radioimmunoassay was developed using anti-sCRFR2 α -and [^{125}I -Tyr 113] sCRFR2 α

(113-143)] as the tracer. Tissue from mouse brain was acid-extracted, partially purified on C18 cartridges and assayed at multiple doses in the radioimmunoassay. The tissue extracts displaced [^{125}I -Tyr¹¹³] sCRFR2 α (113-143)] bound to anti-sCRFR2 α in a dose-dependent manner (FIG. 4G). Highest levels of expression were found in the olfactory bulb, hypothalamus, cortex and midbrain, all of which correlate with the presence of ir cells and fibers, determined by the immunohistochemical studies (FIG. 4). A putative soluble form of CRFRI (generated by deletion of exon 5) would comprise a different unique C-terminal sequence. A protein corresponding to that sequence did not displace [^{125}I -Tyr¹¹³] sCRFR2 α (113-143) in the radioimmunoassay. These results further confirm the existence of sCRFR2 α protein in rodent CNS.

The interactions of the sCRFR2 α with CRF family ligands were assessed by radioreceptor assay using competitive displacement of [^{125}I -D Tyr⁰]-astressin bound to sCRFR2 α . The soluble proteins, secreted by COS M6 cells or produced in bacteria, bind the agonists, Ucn 1 and CRF, as well as the antagonist, astressin, with nanomolar affinities, whereas, the affinities for Ucn 2 and Ucn 3 are much lower (Table 2).

Table 2. Inhibitory binding constants, Ki (nM) for CRF ligands binding to sCRFR2 α proteins.

Protein	CRF	rUcn1	mUcn2	mUcn3	Astressin
mam sCRFR2 α	23 (14-39)	6.6 (3.5-12)	113 (68-190)	>200	6.7 (3.6-12)
bact sCRFR2 α	14.8 (9.2-24)	5.8 (2.5-13.3)	116 (85-158)	>200	10 (7.9-12.5)

20. Binding of CRF family members to sCRFR2 α proteins purified from either COS M6 cell media (mam sCRFR2 α) or *E.coli* (bactsCRFR2 α). See Methods for details.

To delineate the possible hnctions of sCRFR2 α , the inventors studied its effects on signaling by CRF family ligands. Both the mammalian and bacterially expressed sCRFR2 α proteins inhibit, in a dose dependent manner, the cAMP response to Ucn 1 and CRF in HEK293T cells transfected with mouse CRFR2 α as measured by the CRE luciferase activity of the EVX1 gene (FIG. 5A). Because the urocortins activate MAPK signaling (Brar *et al.*, 2002), the inventors measured the ability of sCRFR2 α to inhibit the activation by Ucn 1 of ERK1/2-p42,44 in CATH.a cells, which endogenously express CRFRI and CRFR2 α . The sCRFR2 α inhibits the induction of phosphorylated ERK by Ucn 1 in CATH.a cells (FIG. 5B).

Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined by the appended claims. Moreover, the scope of the present application is not intended to be limited to the particular embodiments of the process, machine, manufacture, composition of matter, means, methods and steps described in the specification. As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed that perform substantially the same function or 5 achieve substantially the same result as the corresponding embodiments described herein may be utilized according to the present invention. Accordingly, the appended claims are intended to include within their scope such processes, machines, manufacture, 10 compositions of matter, means, methods, or steps.

Throughout this specification and the claims, unless the context requires 15 otherwise, the word "comprise" and its variations, such as "comprises" and "comprising," will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps

REFERENCES

The references listed below are incorporated herein by reference to the extent that they supplement, explain, provide a background for, or teach methodology, techniques, and/or compositions employed herein. The reference to any prior art in this specification is not, and should not be taken as an acknowledgement or any form of suggestion that such art forms part of the common general knowledge in Australia.

U.S. Patent 4,196,265

U.S. Patent 4,215,051

U.S. Patent 4,554,101

U.S. Patent 4,683,195

U.S. Patent 4,683,202

U.S. Patent 4,800,159

U.S. Patent 4,883,750

U.S. Patent 5, 795,587

U.S. Patent 5,063,245

U.S. Patent 5,145,684

U.S. Patent 5,225,538

U.S. Patent 5,552,157

U.S. Patent 5,565,213

U.S. Patent 5,567,434

U.S. Patent 5,580,579

U.S. Patent 5,641,515

U.S. Patent 5,738,868

U.S. Patent 5,741,516

U.S. Patent 5,792,451,

U.S. Patent 6,831,158

Allen and Choun, *FEBS Lett.*, 223:42-46, 1987.

Andresz *et al.*, *Makromol. Chem.*, 179: 301, 1978

Avis, Remington's Pharmaceutical Sciences, 18th Ed., Gennaro (Ed.), Mack Publishing Co., Pa., 84:1545-1569, 1990.

Baichwal and Sugden, *In: Gene Transfer*, Kucherlapati (Ed.), NY, Plenum Press, 117-148, 1986.

Bale and Vale, *Annu. Rev. Pharmacol. Toxicol.*, 44, 525-557, 2004.

Bale *et al.*, *Nat. Genet.*, 24,410-414, 2000.

Baud *et al.*, *Genomics*, 26(2):334-344, 1995.

Bellus, *J. Macromol. Sci. Pure Appl. Chem.*, A31(1): 1355-1376, 1994.

Benet *et al.*, In: *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, Hardman *et al.* (Eds.), McGraw-Hill, NY, Chap. 1, 9th Ed., 1996.

Berger *et al.*, *Annu. Rev. Immunol.*, 17:657-700, 1999.

Bittencourt *et al.*, *J. Comp. Neurol.*, 415, 285-312, 1999.

Bittner *et al.*, *Methods in Enzymol.*, 153:516-544, 1987.

Block, Remington's Pharmaceutical Sciences, 18th Ed., Gennaro (Ed.), Mack Publishing Co., Pa., 87:1609-1614, 1990.

Brar *et al.*, In: *Encyclopedia of Hormones & Related Cell Regulators*, Henry (Ed.), AN (Academic Press), 3:13-325, 2002.

Chalmers *et al.*, *Trends Pharmacol. Sci.*, 17(4):166-172. 1996.

Chandran *et al.*, *Indian J. Exp. Biol.*, 35(8):801-809., 1997.

Chang *et al.*, *Neuron.*, 11, 1187-1 195, 1993.

Chen and Okayama, *Mol. Cell Biol.*, 7(8):2745-2752, 1987.

Chen *et al.* *Mol. Endocrinol.*, 19:441-458, 2005.

Chen *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:8967-8971, 1993.

Colbere-Garapin *et al.*, *J. Mol. Biol.*, 150:1-14, 1981.

Coupar *et al.*, *Gene*, 68:1-10, 1988.

Couvreur *et al.*, *FEBS Lett.*, 84(2):323-326, 1977.

Couvreur *et al.*, *J. Pharm. Sci.*, 69(2):199-202, 1980.

Couvreur, *Crit. Rev. Ther. Drug Carrier Syst.*, 5(1):1-20, 1988.

Dautzenberg *et al.*, *J. Neuroendocrinol.*, 11(12):941-949, 1999.

Delgado *et al.*, *Crit Rev Ther Drug Carrier Syst.*, 9(3-4):249-304, 1992.

Douglas *et al.*, *Crit Rev Ther Drug Carrier Syst.*, 3(3):233-61, 1987.

Eason *et al.*, *Transplantation*, 61(2):224-228, 1996.

Ettinger *et al.*, *Cancer*, 41:1270, 1978.

Fechheimer, *et al.*, *Proc Natl. Acad. Sci. USA*, 84:8463-8467, 1987.

Fraley *et al.*, *Proc. Natl. Acad. Sci. USA*, 76:3348-3352, 1979.

Freifelder, In: *Physical Biochemistry Applications to Biochemistry and Molecular Biology*, 2nd Ed. Wm. Freeman and Co., NY, 1982.

Friedmann, *Science*, 244:1275-1281, 1989.

Gabizon and Papahadjopoulos, *Proc. Natl. Acad. Sci. USA*, 85(18):6949-6953, 1988.

GB Appln. 2 202 328

Ghosh and Bachhawat, In: *Liver Diseases, Targeted Diagnosis and Therapy Using Specific Receptors and Ligands*, Wu *et al.* (Eds.), Marcel Dekker, NY, 87-104, 1991.

Goding, In: *Monoclonal Antibodies: Principles and Practice*, 2d ed., Academic Press, Orlando, Fl, pp 60-61, 71-74, 1986.

Gopal, *Mol. Cell Biol.*, 5:1188-1190, 1985.

Graham and Van Der Eb, *Virology*, 52:456-467, 1973.

Grammatopoulos *et al.*, *Mol. Endocrinol.*, 13:2189-2202, 1999.

Graves *et al.*, *Biochem. Biophys. Res. Commun.*, 187: 1135-1 143, 1992.

Hamann *et al.*, *J. Immunol.*, 155(4):1942-1950, 1995.

Harlow and Lane, In: *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 346-348, 1988.

Harvey, Remington's Pharmaceutical Sciences, 18th Ed., Gennaro Ed.), Mack Publishing Co., Pa., 35:711, 1990.

Henry-Michelland *et al.*, *Int J Pharm*, 35: 121-7, 1987.

Hermonat and Muzychka, *Proc. Natl. Acad. Sci. USA*, 81:6466-6470, 1984.

Hess *et al.*, *J. Adv. Enzyme Reg.*, 7:149, 1968.

Hitzeman *et al.*, *J. Biol. Chem.*, 255:2073, 1980.

Holland *et al.*, *Biochemistry*, 17:4900, 1978.

Horwich *et al.* *J. Virol.*, 64:642-650, 1990.

Hsu and Hsueh, *Nat. Med.*, 7605-611, 2001.

Hwang *et al.*, *Crit. Rev. Ther. Drug Carrier Syst.*, 15(3):243-284, 1998.

Innis, *et al.*, In: *PCR Protocols. A guide to Methods and Application*, Academic Press, Inc. San Diego, 1990.

Jacobson *et al.*, *J. Acquir. Immune. Defic. Syndr.*, 21(1):S34-41, 1999.

Ji *et al.*, *Biochem. Biophys. Res. Commun.*, 247:414-419, 1998.

Jones, *Genetics*, 85: 12, 1977.

Kehne and Lombaert, *Curr. Drug Targets CNS Neurol. Disord.*, 1(5):467-493, 2002.

Kemeny *et al.*, *Cancer*, 71:1964, 1993.

Kenakin, *Life Sci.*, 43(14):1095-1101, 1988.

Khan *et al.*, *Biochem. Biophys. Res. Commun.*, 190:888-894, 1993.

Kingsman *et al.*, *Gene*, 7:141, 1979.

Kishimoto *et al.*, *Proc Natl. Acad. Sci. USA*, 92:1108-1112, 1995.

Klein *et al.*, *Nature*, 327:70-73, 1987.

Koob and Heinrichs, *Brain Res.*, 848:141-152, 1999.

Kostich *et al.*, *Mol. Endocrinol.*, 12:1077-1085, 1998.

Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:1173, 1989.

Kyte and Doolittle, *J. Mol. Biol.*, 157(1):105-132, 1982.

Lasic, *Trends Biotechnol.*, 16(7):307-321, 1998.

Lewis *et al.*, *Proc. Natl. Acad. Sci. USA*, 98:570-7575, 2001.

Lin *et al.*, *Science*, 282(5390):943-946, 1998.

Liu *et al.*, *J. Biol. Chem.*, 269:29220-29226, 1994.

Lovenberg *et al.*, *Endocrinology*, 136:4139-4142, 1995.

Lovenberg *et al.*, *Proc. Natl. Acad. Sci. USA*, 92:836-840, 1995.

Lowy *et al.*, *Cell*, 22:817-823, 1980.

Luer and Hatton, In: *The Annals of Pharmacotherapy*, 27:912, 1993.

Malherbe *et al.*, *Brain Res. Mol. Brain Res.*, 67:201-210, 1999.

Mannstadt *et al.*, *Am. J. Physiol.*, 277(5 Pt 2):F665-675, 1999.

Marchese *et al.*, *Trends Pharmacol Sci.*, 20(9):370-375., 1999.

Margalit, *Crit. Rev. Ther. Drug Carrier Syst.*, 12(2-3):233-261, 1995.

Mathiowitz *et al.*, *Nature*, 386(6623):410-414, 1997.

Meij *et al.*, *Mol. Cell Biochem.*, 157(1-2):31-38, 1996.

Morishita *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:8474, 1993.

Muglia *et al.*, *Nature*, 373:427-432, 1995.

Mulligan *et al.*, *Proc. Natl. Acad. Sci. USA*, 78:2072, 1981.

Murphy *et al.*, *J. Virol.*, 74(17):7745-7754, 2000.

Nicolas and Rubinstein, In: *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez and Denhardt, eds., Stoneham: Butterworth, pp. 494-513, 1988.

Nicolau and Sene, *Biochim. Biophys. Acta*, 721:185-190, 1982.

O'Hare *et al.*, *Proc. Natl. Acad. Sci. USA*, 78: 1527, 1981.

Owens *et al.*, *J. Pharmacol. Exp. Ther.*, 258(1):349-356, 1991.

Palczewski *et al.*, *Science*, 289(5480):739-745, 2000.

PCT Appln. PCT Application WO 88/10315

PCT Appln. PCT/US87/00880

PCT Appln. PCT/US89/01025

PCT Appln. WO 00/23114

PCT Appln. WO 92/16655

Perrin and Vale, *Ann. N. Y. Acad. Sci.*, 885:312-328, 1999.

Perrin *et al.*, *J. Biol. Chem.*, 276:31528-31534, 2001.

Perrin *et al.*, *J. Biol. Chem.*, 278:15595-15600, 2003.

Perrin *et al.*, *Proc. Natl. Acad. Sci. USA*, 92:2969-2973, 1995.

Pinto-alphandary *et al.*, *J. Drug Target*, 3(2):167-169, 1995.

Pisarchik and Slominski, *Eur. J. Biochem.*, 271:2821-2830, 2004.

Potter *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:7161-7165, 1984.

Quintanar-Guerrero *et al.*, *Pharm. Res.*, 15(7):1056-1062, 1998.

Rekasi *et al.*, *Proc. Natl. Acad. Sci. USA*, 97:0561-10566, 2000.

Remington's Pharmaceutical Sciences, 15th ed., pages 1035-1038 and 1570-1580, Mack Publishing Company, Easton, PA, 1980.

Reyes *et al.*, *Proc. Natl. Acad. Sci. USA*, 98:2843-2848, 2001.

Ridgeway, *In: Vectors: A survey of molecular cloning vectors and their uses*, Stoneham: Butterworth, pp. 467-492, 1988.

Rippe, *et al.*, *Mol. Cell Biol.*, 10:689-695, 1990.

Rivier and Vale, *Nature*, 305:325-327, 1983.

Sambrook *et al.*, *In: Molecular cloning: a laboratory manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Schwarz *et al.*, *J. Biol. Chem.*, 275:32174-32181, 2000.

Seck *et al.*, *J. Biol. Chem.*, 278:23085-23093, 2003.

Shaw, *Cancer*, 72(11):3416, 1993.

Smith *et al.*, *Neuron*, 20:1093-1102, 1998.

Stacey *et al.*, *Trends Biochem. Sci.*, 25(6):284-289, 2000.

Stenzel *et al.*, *Mol. Endocrinol.*, 9:637-645, 1995.

Stinchcomb *et al.*, *Nature*, 282(5734):39-43, 1979.

Szybalska *et al.*, *Proc. Natl. Acad. Sci. USA*, 48:2026, 1962.

Takakura, *Nippon Rinsho*, 56(3):691-695, 1998.

Temin, *In: Gene Transfer*, Kucherlapati (Ed.), NY, Plenum Press, 149-188, 1986.

Timpl *et al.*, *Nat. Genet.*, 19:162-166, 1998.

Tschemper *et al.*, *Gene*, 10: 157, 1980.

Tur-Kaspa *et al.*, *Mol. Cell Biol.*, 6:716-718, 1986.

Vale *et al.*, *Science*, 213:1394-1397, 1981.

Valerio *et al.*, *Neuroreport*, 12:2711-2715, 2001.

van Dullemen *et al.*, *Gastroenterology*, 109, 129-35, 1995.

VanPett *et al.*, *J. Comp. Neurol.*, 428:191-212, 2000.

Vaughan *et al.*, *Methods Enzymol.*, 168:588-617, 1989.

Vaughan *et al.*, *Nature*, 378:287-292, 1995.

Vita *et al.*, *FEBS Lett.*, 335:1-5, 1993.

Walker *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:392-396, 1992.

Wigler *et al.*, *Cell*, 11(1):223-232, 1977.

Wigler *et al.*, *Proc. Natl. Acad. Sci. USA*, 77(6):3567-3570, 1980.

Wilson *et al.*, *Br. J. Pharmacol.*, 125(7):1387-1392, 1998.

Wilson *et al.*, In: *G-protein-coupled receptors*, CRC press, Boca Raton, 97-116, 1999.

Wu and Wu, *Adv. Drug Delivery Rev.*, 12:159-167, 1993.

Wu and Wu, *Biochemistry*, 27: 887-892, 1988.

Wu and Wu, *J. Biol. Chem.*, 262:4429-4432, 1987.

Yaida *et al.*, *Regul. Pept.*, 59:193, 1995.

Yan and Wold, *Biochemistry*, 23(16):3759-3765.1984.

Yang and Russell, *Proc. Natl. Acad. Sci. USA*, 87:4144-4148, 1990.

You *et al.*, *Biol. Reprod.*, 62:108-116, 2000.

Zambaux *et al.*, *J. Control Release*, 50(1-3):31-40, 1998.

Zhu *et al.*, *Brain Res. Mol. Brain Res.*, 73:3-103, 1999.

Zimm *et al.*, *Cancer Research*, 44:1698, 1984.

zur Muhlen *et al.*, *Eur. J. Pharm. Biopharm.*, 45(2):149-155, 1998.

The claims defining the invention are as follows.

1. An isolated, soluble corticotropin releasing factor receptor type 2 alpha (sCRFR2 alpha) ligand binding domain having at least 75% similarity to SEQ ID NO:4, the domain capable of binding a CRFR2 alpha ligand.
2. The sCRFR2 alpha ligand binding domain of claim 1, wherein the sCRFR2 alpha ligand binding domain has at least 80% similarity to SEQ ID NO:4.
3. The sCRFR2 alpha ligand binding domain of claim 1, wherein the sCRFR2 alpha ligand binding domain has at least 85% similarity to SEQ ID NO:4.
4. The sCRFR2 alpha ligand binding domain of claim 1, wherein the sCRFR2 alpha ligand binding domain has at least 90% similarity to SEQ ID NO:4.
5. The sCRFR2 alpha ligand binding domain of claim 1, wherein the sCRFR2 alpha ligand binding domain has at least 95% similarity to SEQ ID NO:4.
6. The sCRFR2 alpha ligand binding domain of claim 1, wherein the sCRFR2 alpha ligand binding domain has at least 98% similarity to SEQ ID NO:4.
7. The sCRFR2 alpha ligand binding domain of claim 1, wherein the sCRFR2 alpha ligand binding domain comprises the amino acid sequence of SEQ ID NO:4.
8. The sCRFR2 alpha ligand binding domain of any one of claims 1 through 7, wherein the isolated sCRFR2 alpha ligand binding domain further comprises an affinity tag, a label, a radionuclide, an enzyme, a fluorescent marker, a chemiluminescent marker, an immunoglobulin domain or a combination thereof, in particular and immunoglobulin Fc domain.
9. The sCRFR2 alpha ligand binding domain of any one of claims 1 through 7, further comprising a leader sequence.
10. The sCRFR2 alpha ligand binding domain of any one of claims 1 through 7, wherein the sCRFR2 alpha ligand binding domain is conjugated to a polymer, in particular polyethylene glycol (PEG).

11. An isolated nucleic acid encoding a sCRFR2 alpha ligand binding domain of any one of claims 1-7.
12. The nucleic acid of claim 11, wherein the nucleic acid is an expression cassette, in particular wherein the expression cassette is comprised in an expression vector such as a linear nucleic acid, a plasmid expression vector, or a viral expression vector, more in particular wherein the expression vector is operably coupled to a delivery vector, such as a liposome, a polypeptide, a polycation, a lipid, a bacterium, or a virus.
13. Use of the isolated, soluble corticotropin releasing factor receptor type 2 alpha (sCRFR2 alpha) ligand binding domain of any one of claims 1 through 7 in the therapeutic treatment of type II diabetes, insulin sensitivity, anxiety-related disorders; a mood disorders; bipolar disorders; post-traumatic stress disorder; inflammatory disorders; chemical dependencies and addictions; gastrointestinal disorders; or skin disorders.
14. The isolated, soluble corticotropin releasing factor receptor type 2 alpha (sCRFR2 alpha) ligand binding domain of claim 13, being contained in a medicament formulated for administration by ingestion, injection, endoscopy, or perfusion, in particular wherein injection is intravenous injection, intramuscular injection, subcutaneous injection, intradermal injection, intracranial injection or intraperitoneal injection.
15. A medicament comprising the isolated, soluble corticotropin releasing factor receptor type 2 alpha (sCRFR2 alpha) ligand binding domain of any one of claims 1 through 7.

Fig. 1

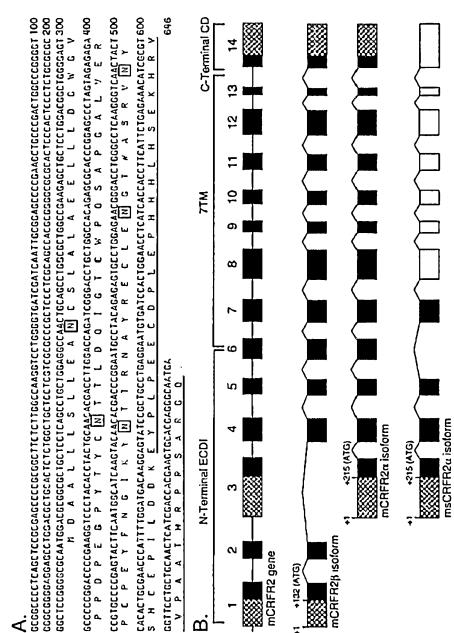


Fig. 2

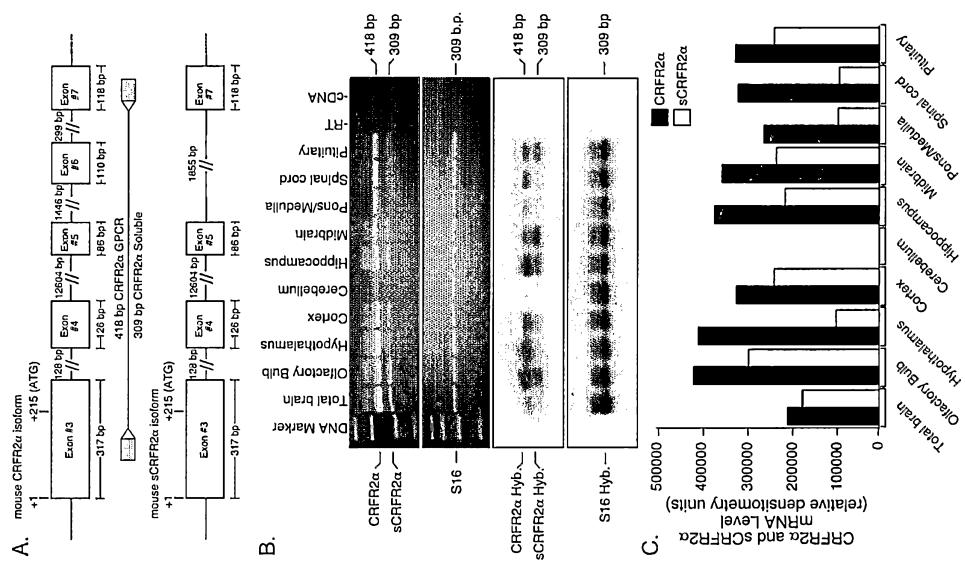
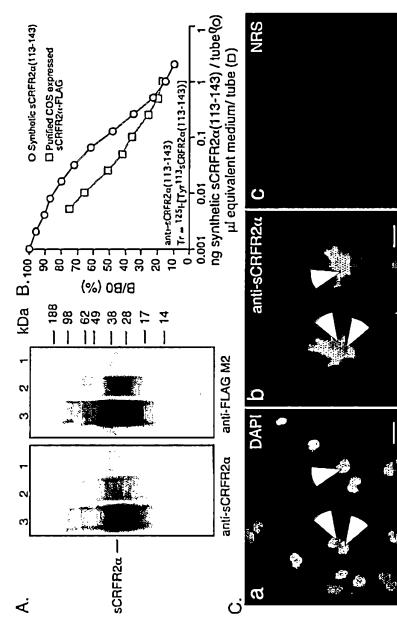


Fig. 3



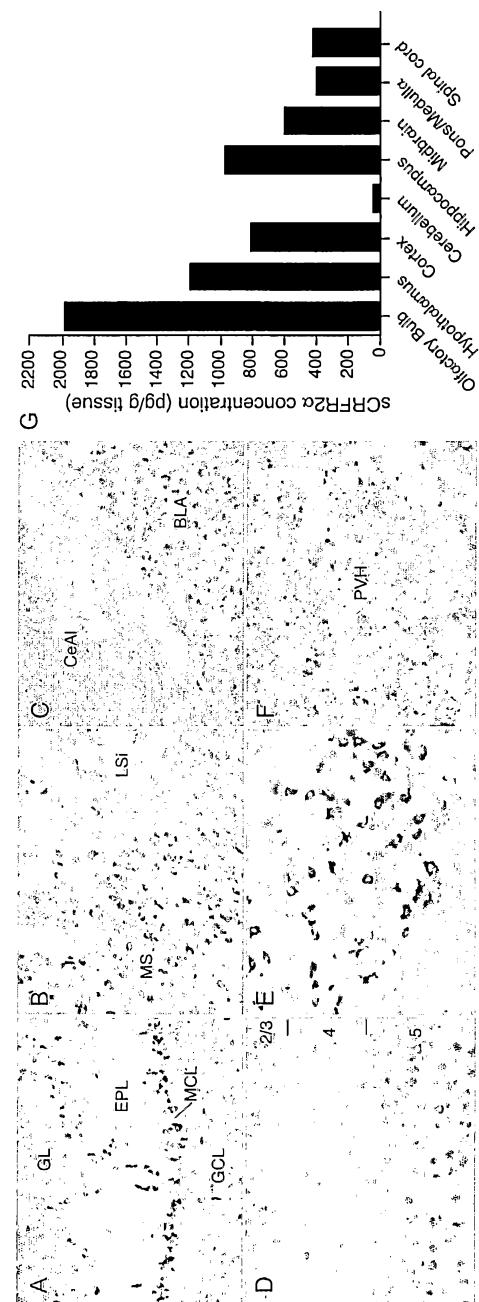


Fig. 4

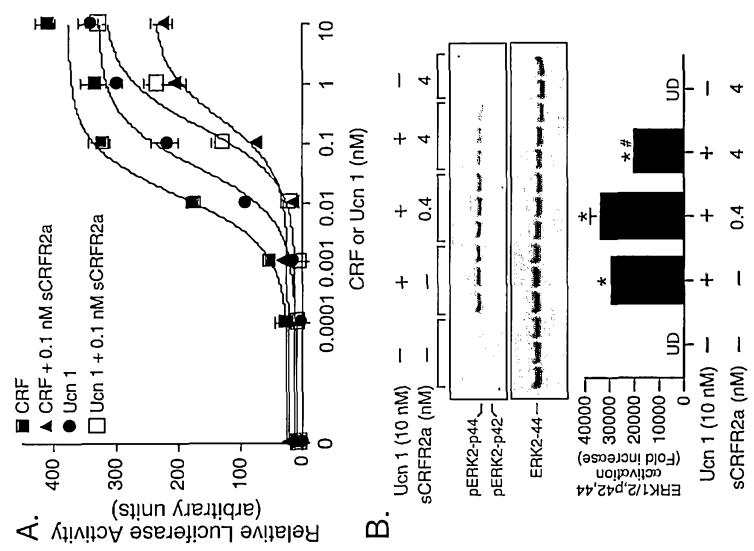


Fig. 5

SEQUENCE LISTING

<110> Research Development Foundation
Chen, Alon
Vale, Wyllie
Perrin, Marilyn

<120> Compositions and Methods Related to Soluble G-Protein Coupled Receptors (sGPCRs)

<130> CLFR:247WO

<140> UNKNOWN

<141> 2006-02-01

<150> US 60/650,866
<151> 2005-02-08

<160> 17

<170> PatentIn version 3.3

<210> 1
<211> 2107
<212> DNA
<213> homo sapien

<220>
<221> CDS
<222> (1)..(1233)

<400> 1
atg gac gcg gca ctg ctc cac agc ctg ctg gag gcc aac tgc agc ctg
48
Met Asp Ala Ala Leu Leu His Ser Leu Leu Glu Ala Asn Cys Ser Leu
1 5 10 15
gcg ctg gct gaa gag ctg ctc ttg gac ggc tgg ggg cca ccc ctg gac
96
Ala Leu Ala Glu Glu Leu Leu Asp Gly Trp Gly Pro Pro Leu Asp
20 25 30
ccc gag ggt ccc tac tcc tac tgc aac acg acc ttg gac cag atc gga
144
Pro Glu Gly Pro Tyr Ser Tyr Cys Asn Thr Thr Leu Asp Gln Ile Gly
35 40 45
acg tgc tgg ccc cgc agc gct gcc gga gcc ctc gtg gag agg ccg tgc
192
Thr Cys Trp Pro Arg Ser Ala Ala Gly Ala Leu Val Glu Arg Pro Cys
50 55 60
ccc gag tac ttc aac ggc gtc aag tac aac acg acc cgg aat gcc tat
240
Pro Glu Tyr Phe Asn Gly Val Lys Tyr Asn Thr Thr Arg Asn Ala Tyr
65 70 75 80
cga gaa tgc ttg gag aat ggg acg tgg gcc tca aag atc aac tac tca
288
Arg Glu Cys Leu Glu Asn Gly Thr Trp Ala Ser Lys Ile Asn Tyr Ser
85 90 95
cag tgt gag ccc att ttg gat gac aag cag agg aag tat gac ctg cac
336
Gln Cys Glu Pro Ile Leu Asp Asp Lys Gln Arg Lys Tyr Asp Leu His
100 105 110

tac cgc atc gcc ctt gtc gtc aac tac ctg ggc cac tgc gta tct gtc
384
Tyr Arg Ile Ala Leu Val Val Asn Tyr Leu Gly His Cys Val Ser Val
115 120 125
gca gcc ctg gtg gcc gcc ttc ctg ctt ttc ctg gcc ctg cgg agc att
432
Ala Ala Leu Val Ala Ala Phe Leu Leu Phe Leu Ala Leu Arg Ser Ile
130 135 140
cgc tgt ctg cgg aat gtg att cac tgg aac ctc atc acc acc ttt atc
480
Arg Cys Leu Arg Asn Val Ile His Trp Asn Leu Ile Thr Thr Phe Ile
145 150 155 160
ctg cga aat gtc atg tgg ttc ctg ctg cag ctc gtt gac cat gaa gtc
528
Leu Arg Asn Val Met Trp Phe Leu Leu Gln Leu Val Asp His Glu Val
165 170 175
cac gag agc aat gag gtc tgg tgc cgc tgc atc acc acc atc ttc aac
576
His Glu Ser Asn Glu Val Trp Cys Arg Cys Ile Thr Thr Ile Phe Asn
180 185 190
tac ttc gtg gtg acc aac ttc ttc tgg atg ttt gtc gaa ggc tgc tac
624
Tyr Phe Val Val Thr Asn Phe Phe Trp Met Phe Val Glu Gly Cys Tyr
195 200 205
ctg cac acg gcc att gtc atg acc tac tcc act gag cgc ctg cgc aag
672
Leu His Thr Ala Ile Val Met Thr Tyr Ser Thr Glu Arg Leu Arg Lys
210 215 220
tgc ctc ttc ctc ttc atc gga tgg tgc atc ccc ttc ccc atc atc gtc
720
Cys Leu Phe Leu Phe Ile Gly Trp Cys Ile Pro Phe Pro Ile Ile Val
225 230 235 240
gcc tgg gcc atc ggc aag ctc tac tat gag aat gaa cag tgc tgg ttt
768
Ala Trp Ala Ile Gly Lys Leu Tyr Tyr Glu Asn Glu Gln Cys Trp Phe
245 250 255
ggc aag gag cct ggc gac ctg gtg gac tac atc tac caa ggc ccc atc
816
Gly Lys Glu Pro Gly Asp Leu Val Asp Tyr Ile Tyr Gln Gly Pro Ile
260 265 270
att ctc gtg ctc ctg atc aat ttc gta ttt ctg ttc aac atc gtc agg
864
Ile Leu Val Leu Leu Ile Asn Phe Val Phe Leu Phe Asn Ile Val Arg
275 280 285
atc cta atg aca aag tta cgc gcg tcc acc aca tcc gag aca atc cag
912
Ile Leu Met Thr Lys Leu Arg Ala Ser Thr Thr Ser Glu Thr Ile Gln
290 295 300
tac agg aag gca gtg aag gcc acc ctg gtg ctc ctg ccc ctc ctg ggc
960
Tyr Arg Lys Ala Val Lys Ala Thr Leu Val Leu Leu Pro Leu Leu Gly
305 310 315 320
atc acc tac atg ctc ttc ttc gtc aat ccc ggg gag gac gac ctg tca
1008
Ile Thr Tyr Met Leu Phe Phe Val Asn Pro Gly Glu Asp Asp Leu Ser
325 330 335
cag atc atg ttc atc tat ttc aac tcc ttc ctg cag tcg ttc cag ggt
1056
Gln Ile Met Phe Ile Tyr Phe Asn Ser Phe Leu Gln Ser Phe Gln Gly
340 345 350

ttc ttc gtg tct gtc ttc tac tgc ttc ttc aat gga gag gtg cgc tca
 1104
 Phe Phe Val Ser Val Phe Tyr Cys Phe Phe Asn Gly Glu Val Arg Ser
 355 360 365
 gcc gtg agg aag agg tgg cac cgc tgg cag gac cat cac tcc ctt cga
 1152
 Ala Val Arg Lys Arg Trp His Arg Trp Gln Asp His His Ser Leu Arg
 370 375 380
 gtc ccc atg gcc cggttccatc cct aca tca ccc aca cgg atc
 1200
 Val Pro Met Ala Arg Ala Met Ser Ile Pro Thr Ser Pro Thr Arg Ile
 385 390 395 400
 agc ttc cac agc atc aag cag acg gcc gct gtg tgaccctcg gtcgcccacc
 1253
 Ser Phe His Ser Ile Lys Gln Thr Ala Ala Val
 405 410
 tgcacagctc ccctgtcctc ctccaccccttc ttccctctggg ttctctgtgc tggcaggct
 1313
 ctcgtggggc aggagatggg aggggagaga ccagctctcc agcctggcag gaaagagggg
 1373
 gtgcggcagc caagggggac tgcaagggac agggatgagt gggggccacc aggtcagcg
 1433
 caagaggaag cagagggaat tcacaggacc ccctgagaag agccagtcag atgtctgcag
 1493
 gcatttgccc atcccagcct ctctggccag ggccttactg ggcccagagc agagaaggac
 1553
 ctgtccaaaca cacacagcta ttatagtag cagacacagg gtccttcgc cctactcatg
 1613
 gagccagcag ccaggcaatg gtgtggccct gcactggccc ttggactcca cactcagtgg
 1673
 tgccctgcag ttgggtgggt tacgccagca aaggatcagt ttggctgcct tatcccaggg
 1733
 ctgtcaccta gagaggctca ctgttacccc accctgttcc tgtgtccct ccccagccat
 1793
 cctcccgccct tggggctcc atgaaggatg caggcttcca ggcctggcct cctctttgg
 1853
 gagaccctt ctctgcctag tccacagatt aggaatcaa ggaagacgcc atcaggaaag
 1913
 ccacatcctt agtcaaccag ttgcattgtc cggggcaaaa tgaggagcag agcatggag
 1973
 gagggaggcg tggatggga atagcagaac caccatgtct tcagtgattt aaactcatac
 2033
 cccattgccc ttggccctcc agtctccct tcagaaacat ctctgcttc tgtgaaataa
 2093
 accatgcctc ttgg
 2107

 <210> 2
 <211> 411
 <212> PRT
 <213> homo sapien

 <400> 2

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

50	55	60													
Pro	Glu	Tyr	Phe	Asn	Gly	Val	Lys	Tyr	Asn	Thr	Thr	Arg	Asn	Ala	Tyr
65					70			75							80
Arg	Glu	Cys	Leu	Glu	Asn	Gly	Thr	Trp	Ala	Ser	Lys	Ile	Asn	Tyr	Ser
					85				100	105	90				95
Gln	Cys	Glu	Pro	Ile	Leu	Asp	Asp	Lys	Gln	Arg	Lys	Tyr	Asp	Leu	His
					100				105	110					110
Tyr	Arg	Ile	Ala	Leu	Val	Val	Asn	Tyr	Leu	Gly	His	Cys	Val	Ser	Val
					115			120			125				
Ala	Ala	Leu	Val	Ala	Ala	Phe	Leu	Leu	Phe	Leu	Ala	Leu	Arg	Ser	Ile
					130			135			140				
Arg	Cys	Leu	Arg	Asn	Val	Ile	His	Trp	Asn	Leu	Ile	Thr	Thr	Phe	Ile
					145			150			155				160
Leu	Arg	Asn	Val	Met	Trp	Phe	Leu	Leu	Gln	Leu	Val	Asp	His	Glu	Val
					165			170			175				
His	Glu	Ser	Asn	Glu	Val	Trp	Cys	Arg	Cys	Ile	Thr	Thr	Ile	Phe	Asn
					180			185			190				
Tyr	Phe	Val	Val	Thr	Asn	Phe	Phe	Trp	Met	Phe	Val	Glu	Gly	Cys	Tyr
					195			200			205				
Leu	His	Thr	Ala	Ile	Val	Met	Thr	Tyr	Ser	Thr	Glu	Arg	Leu	Arg	Lys
					210			215			220				
Cys	Leu	Phe	Leu	Phe	Ile	Gly	Trp	Cys	Ile	Pro	Phe	Pro	Ile	Ile	Val
					225			230			235				240
Ala	Trp	Ala	Ile	Gly	Lys	Leu	Tyr	Tyr	Glu	Asn	Glu	Gln	Cys	Trp	Phe
					245			250			255				
Gly	Lys	Glu	Pro	Gly	Asp	Leu	Val	Asp	Tyr	Ile	Tyr	Gln	Gly	Pro	Ile
					260			265			270				
Ile	Leu	Val	Leu	Leu	Ile	Asn	Phe	Val	Phe	Leu	Phe	Asn	Ile	Val	Arg
					275			280			285				
Ile	Leu	Met	Thr	Lys	Leu	Arg	Ala	Ser	Thr	Thr	Ser	Glu	Thr	Ile	Gln
					290			295			300				
Tyr	Arg	Lys	Ala	Val	Lys	Ala	Thr	Leu	Val	Leu	Leu	Pro	Leu	Leu	Gly
					305			310			315				320
Ile	Thr	Tyr	Met	Leu	Phe	Phe	Val	Asn	Pro	Gly	Glu	Asp	Asp	Leu	Ser
					325			330			335				
Gln	Ile	Met	Phe	Ile	Tyr	Phe	Asn	Ser	Phe	Leu	Gln	Ser	Phe	Gln	Gly
					340			345			350				
Phe	Phe	Val	Ser	Val	Phe	Tyr	Cys	Phe	Phe	Asn	Gly	Glu	Val	Arg	Ser
					355			360			365				
Ala	Val	Arg	Lys	Arg	Trp	His	Arg	Trp	Gln	Asp	His	His	Ser	Leu	Arg
					370			375			380				
Val	Pro	Met	Ala	Arg	Ala	Met	Ser	Ile	Pro	Thr	Ser	Pro	Thr	Arg	Ile
					385			390			395				400
Ser	Phe	His	Ser	Ile	Lys	Gln	Thr	Ala	Ala	Val					
					405			410							

<210> 3
<211> 1997
<212> DNA
<213> homo sapien

<220>
<221> CDS
<222> (1)..(405)

<400> 3
atg gac gcg gca ctg ctc cac agc ctg ctg gag gcc aac tgc agc ctg
48
Met Asp Ala Ala Leu Leu His Ser Leu Leu Glu Ala Asn Cys Ser Leu
1 5 10 15

gcg ctg gct gaa gag ctg ctc ttg gac ggc tgg ggg cca ccc ctg gac
 96
 Ala Leu Ala Glu Glu Leu Leu Leu Asp Gly Trp Gly Pro Pro Leu Asp
 20 25 30
 ccc gag ggt ccc tac tcc tac tgc aac acg acc ttg gac cag atc gga
 144
 Pro Glu Gly Pro Tyr Ser Tyr Cys Asn Thr Thr Leu Asp Gln Ile Gly
 35 40 45
 acg tgc tgg ccc cgc agc gct gcc gga gcc ctc gtg gag agg ccg tgc
 192
 Thr Cys Trp Pro Arg Ser Ala Ala Gly Ala Leu Val Glu Arg Pro Cys
 50 55 60
 ccc gag tac ttc aac ggc gtc aag tac aac acg acc ccg aat gcc tat
 240
 Pro Glu Tyr Phe Asn Gly Val Lys Tyr Asn Thr Thr Arg Asn Ala Tyr
 65 70 75 80
 cga gaa tgc ttg gag aat ggg acg tgg gcc tca aag atc aac tac tca
 288
 Arg Glu Cys Leu Glu Asn Gly Thr Trp Ala Ser Lys Ile Asn Tyr Ser
 85 90 95
 cag tgt gag ccc att ttg gat gac aag gag cat tcg ctg tct gcg gaa
 336
 Gln Cys Glu Pro Ile Leu Asp Asp Lys Glu His Ser Leu Ser Ala Glu
 100 105 110
 tgt gat tca ctg gaa cct cat cac cac ctt tat cct gcg aaa tgt cat
 384
 Cys Asp Ser Leu Glu Pro His His His Leu Tyr Pro Ala Lys Cys His
 115 120 125
 gtg gtt cct gct gca gct cgt tgaccatgaa gtgcacgaga gcaatgaggt
 435
 Val Val Pro Ala Ala Ala Arg
 130 135
 ctgggtccgc tgcatcacca ccatcttcaa ctacttcgtg gtgaccaact tcttcggat
 495
 gtttgtggaa ggctgctacc tgcacacggc cattgtcatg acctactcca ctgagcgcct
 555
 ggcgaagtgc ctcttcctct tcacatggatg gtgcattcccc ttccccatca tcgtccctg
 615
 ggcacatcgcc aagctctact atgagaatga acagtgcgtgg tttggcaagg agcctggcga
 675
 cctgggtggac tacatctacc aaggccccat catttcgtg ctccctgatca atttcgtatt
 735
 tctgttcaac atcgtcagga tcctaattgac aaagtacgc ggcgtccacca catccgagac
 795
 aatccagttac aggaaggcag tgaaggccac cctggtgctc ctgcccctcc tgggcatcac
 855
 ctacatgctc ttcttcgtca atccccgggg ggacgacctg tcacagatca tggatcatcta
 915
 tttcaactcc ttccctgcagt cgttccagg tttcttcgtg tctgtcttct actgcttctt
 975
 caatggagag gtgcgtcag ccgtgaggaa gaggtggcac cgctggcagg accatcactc
 1035
 ccttcgagtc cccatggccc gggccatgtc catccctaca tcacccacac ggatcagctt
 1095
 ccacagcata aagcagacgg cgcgtgtgtg acccctcggt cgcccacctg cacagctccc
 1155
 ctgtccctcct ccaccttctt cctctgggtt ctctgtgctg ggcaggctct cgtggggcag
 1215
 gagatgggag gggagagacc agctctccag cctggcagga aagaggggggt gcggcagcca
 1275

agggggactg caaggacag ggatgagtgg gggccaccag gctcagcgca agaggaagca
 1335
 gagggaattc acaggacccc ctgagaagag ccagtcagat gtctgcaggc atttgc当地
 1395
 cccagcctct ctggccaggg cttactggg cccagagcag agaaggacct gtccaaacaca
 1455
 cacagctatt tata>tagca gacacagggc tcccctgccc tactcatgga gccagcagcc
 1515
 aggcaatggt gtggccctgc actggccctt ggactccaca ctcagtggtg ccctgc当地
 1575
 gggtggtta cgccagcaaa ggatcagttt ggctgc当地ta tcccagggt gtcacctaga
 1635
 gaggctcaact tgtacccac cctgttcctg tgtccctcc ccagccatcc tcccgc当地
 1695
 ggggctccat gaaggatgca ggcttccagg cctggctcc tctcttggga gacccttct
 1755
 ctgcctagtc cacagattag gcaatcaagg aagacgcat cagggaaagcc acatccttag
 1815
 tcaaccagg gcacgtgcg gggcaaaatg aggaggcagag gcacggagga gggaggcgtg
 1875
 ggatggaaat agcagaacca ccatgtcttc agtgattgaa actcatacccatgc当地
 1935
 tgccctccag tctcccttc agaaacatct ctgctcttg taaaataaac catgc当地
 1995
 gg
 1997

<210> 4
 <211> 135
 <212> PRT
 <213> homo sapien

<400> 4

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

<210> 5
 <211> 1320
 <212> DNA
 <213> homo sapien

<220>
 <221> CDS
 <222> (1)..(1314)

<400> 5
 atg agg ggt ccc tca ggg ccc cca ggc ctc ctc tac gtc cca cac ctc
 48
 Met Arg Gly Pro Ser Gly Pro Pro Gly Leu Leu Tyr Val Pro His Leu
 1 5 10 15
 ctc ctc tgc ctg ctc tgc ctc ctc cca ccg ccg ctc caa tac gca gcc
 96
 Leu Leu Cys Leu Leu Cys Leu Leu Pro Pro Pro Leu Gln Tyr Ala Ala
 20 25 30
 ggg cag agc cag atg ccc aaa gac cag ccc ctg tgg gca ctt ctg gag
 144
 Gly Gln Ser Gln Met Pro Lys Asp Gln Pro Leu Trp Ala Leu Leu Glu
 35 40 45
 cag tac tgc cac acc atc atg acc ctc acc aac ctc tca ggt ccc tac
 192
 Gln Tyr Cys His Thr Ile Met Thr Leu Thr Asn Leu Ser Gly Pro Tyr
 50 55 60
 tcc tac tgc aac acc ttc gac cag atc gga acg tgc tgg ccc cgc
 240
 Ser Tyr Cys Asn Thr Thr Leu Asp Gln Ile Gly Thr Cys Trp Pro Arg
 65 70 75 80
 agc gct gcc gga gcc ctc gtg gag agg ccg tgc ccc gag tac ttc aac
 288
 Ser Ala Ala Gly Ala Leu Val Glu Arg Pro Cys Pro Glu Tyr Phe Asn
 85 90 95
 ggc gtc aag tac aac acg acc cgg aat gcc tat cga gaa tgc ttg gag
 336
 Gly Val Lys Tyr Asn Thr Thr Arg Asn Ala Tyr Arg Glu Cys Leu Glu
 100 105 110
 aat ggg acg tgg gcc tca aag atc aac tac tca cag tgt gag ccc att
 384
 Asn Gly Thr Trp Ala Ser Lys Ile Asn Tyr Ser Gln Cys Glu Pro Ile
 115 120 125
 ttg gat gac aag cag agg aag tat gac ctg cac tac cgc atc gcc ctt
 432
 Leu Asp Asp Lys Gln Arg Lys Tyr Asp Leu His Tyr Arg Ile Ala Leu
 130 135 140
 gtc gtc aac tac ctg ggc cac tgc gta tct gtg gca gcc ctg gtg gcc
 480
 Val Val Asn Tyr Leu Gly His Cys Val Ser Val Ala Ala Leu Val Ala
 145 150 155 160
 gcc ttc ctg ctt ttc ctg gcc ctg cgg agc att cgc tgt ctg cgg aat
 528
 Ala Phe Leu Leu Phe Leu Ala Leu Arg Ser Ile Arg Cys Leu Arg Asn
 165 170 175
 gtg att cac tgg aac ctc atc acc acc ttt atc ctg cga aat gtc atg
 576
 Val Ile His Trp Asn Leu Ile Thr Thr Phe Ile Leu Arg Asn Val Met
 180 185 190
 tgg ttc ctg ctg cag ctc gtt gac cat gaa gtg cac gag agc aat gag
 624
 Trp Phe Leu Leu Gln Leu Val Asp His Glu Val His Glu Ser Asn Glu
 195 200 205
 gtc tgg tgc cgc tgc atc acc acc atc ttc aac tac ttc gtg gtg acc
 672
 Val Trp Cys Arg Cys Ile Thr Thr Ile Phe Asn Tyr Phe Val Val Thr
 210 215 220
 aac ttc ttc tgg atg ttt gtg gaa ggc tgc tac ctg cac acg gcc att
 720
 Asn Phe Phe Trp Met Phe Val Glu Gly Cys Tyr Leu His Thr Ala Ile

225 230 235 240
 gtc atg acc tac tcc act gag cgc ctg cgc aag tgc ctc ttc ctc ttc
 768
 Val Met Thr Tyr Ser Thr Glu Arg Leu Arg Lys Cys Leu Phe Leu Phe
 245 250 255 255
 atc gga tgg tgc atc ccc ttc ccc atc atc gtc gcc tgg gcc atc ggc
 816
 Ile Gly Trp Cys Ile Pro Phe Pro Ile Ile Val Ala Trp Ala Ile Gly
 260 265 270 270
 aag ctc tac tat gag aat gaa cag tgc tgg ttt ggc aag gag cct ggc
 864
 Lys Leu Tyr Tyr Glu Asn Glu Gln Cys Trp Phe Gly Lys Glu Pro Gly
 275 280 285 285
 gac ctg gtg gac tac atc tac caa ggc ccc atc att ctc gtg ctc ctg
 912
 Asp Leu Val Asp Tyr Ile Tyr Gln Gly Pro Ile Ile Leu Val Leu Leu
 290 295 300 300
 atc aat ttc gta ttt ctg ttc aac atc gtc agg atc cta atg aca aag
 960
 Ile Asn Phe Val Phe Leu Phe Asn Ile Val Arg Ile Leu Met Thr Lys
 305 310 315 320
 tta cgc gcg tcc acc aca tcc gag aca atc cag tac agg aag gca gtg
 1008
 Leu Arg Ala Ser Thr Thr Ser Glu Thr Ile Gln Tyr Arg Lys Ala Val
 325 330 335 335
 aag gcc acc ctg gtg ctc ctg ccc ctc ctg ggc atc acc tac atg ctc
 1056
 Lys Ala Thr Leu Val Leu Leu Pro Leu Leu Gly Ile Thr Tyr Met Leu
 340 345 350 350
 ttc ttc gtc aat ccc ggg gag gac gac ctg tca cag atc atg ttc atc
 1104
 Phe Phe Val Asn Pro Gly Glu Asp Asp Leu Ser Gln Ile Met Phe Ile
 355 360 365 365
 tat ttc aac tcc ttc ctg cag tcc ttc cag ggt ttc ttc gtg tct gtc
 1152
 Tyr Phe Asn Ser Phe Leu Gln Ser Phe Gln Gly Phe Phe Val Ser Val
 370 375 380 380
 ttc tac tgc ttc ttc aat gga gag gtg cgc tca gcc gtg agg aag agg
 1200
 Phe Tyr Cys Phe Phe Asn Gly Glu Val Arg Ser Ala Val Arg Lys Arg
 385 390 395 400
 tgg cac cgc tgg cag gac cat cac tcc ctt cga gtc ccc atg gcc cgg
 1248
 Trp His Arg Trp Gln Asp His His Ser Leu Arg Val Pro Met Ala Arg
 405 410 415 415
 gcc atg tcc atc cct aca tca ccc aca cgg atc agc ttc cac agc atc
 1296
 Ala Met Ser Ile Pro Thr Ser Pro Thr Arg Ile Ser Phe His Ser Ile
 420 425 430 430
 aag cag acg gcc gct gtg tgaccc
 1320
 Lys Gln Thr Ala Ala Val
 435
 <210> 6
 <211> 438
 <212> PRT
 <213> homo sapien
 <400> 6

Met Arg Gly Pro Ser Gly Pro Pro Gly Leu Leu Tyr Val Pro His Leu
 1 5 10 15
 Leu Leu Cys Leu Leu Cys Leu Leu Pro Pro Pro Leu Gln Tyr Ala Ala
 20 25 30
 Gly Gln Ser Gln Met Pro Lys Asp Gln Pro Leu Trp Ala Leu Leu Glu
 35 40 45
 Gln Tyr Cys His Thr Ile Met Thr Leu Thr Asn Leu Ser Gly Pro Tyr
 50 55 60
 Ser Tyr Cys Asn Thr Thr Leu Asp Gln Ile Gly Thr Cys Trp Pro Arg
 65 70 75 80
 Ser Ala Ala Gly Ala Leu Val Glu Arg Pro Cys Pro Glu Tyr Phe Asn
 85 90 95
 Gly Val Lys Tyr Asn Thr Thr Arg Asn Ala Tyr Arg Glu Cys Leu Glu
 100 105 110
 Asn Gly Thr Trp Ala Ser Lys Ile Asn Tyr Ser Gln Cys Glu Pro Ile
 115 120 125
 Leu Asp Asp Lys Gln Arg Lys Tyr Asp Leu His Tyr Arg Ile Ala Leu
 130 135 140
 Val Val Asn Tyr Leu Gly His Cys Val Ser Val Ala Ala Leu Val Ala
 145 150 155 160
 Ala Phe Leu Leu Phe Leu Ala Leu Arg Ser Ile Arg Cys Leu Arg Asn
 165 170 175
 Val Ile His Trp Asn Leu Ile Thr Thr Phe Ile Leu Arg Asn Val Met
 180 185 190
 Trp Phe Leu Leu Gln Leu Val Asp His Glu Val His Glu Ser Asn Glu
 195 200 205
 Val Trp Cys Arg Cys Ile Thr Thr Ile Phe Asn Tyr Phe Val Val Thr
 210 215 220
 Asn Phe Phe Trp Met Phe Val Glu Gly Cys Tyr Leu His Thr Ala Ile
 225 230 235 240
 Val Met Thr Tyr Ser Thr Glu Arg Leu Arg Lys Cys Leu Phe Leu Phe
 245 250 255
 Ile Gly Trp Cys Ile Pro Phe Pro Ile Ile Val Ala Trp Ala Ile Gly
 260 265 270
 Lys Leu Tyr Tyr Glu Asn Glu Gln Cys Trp Phe Gly Lys Glu Pro Gly
 275 280 285
 Asp Leu Val Asp Tyr Ile Tyr Gln Gly Pro Ile Ile Leu Val Leu Leu
 290 295 300
 Ile Asn Phe Val Phe Leu Phe Asn Ile Val Arg Ile Leu Met Thr Lys
 305 310 315 320
 Leu Arg Ala Ser Thr Thr Ser Glu Thr Ile Gln Tyr Arg Lys Ala Val
 325 330 335
 Lys Ala Thr Leu Val Leu Leu Pro Leu Leu Gly Ile Thr Tyr Met Leu
 340 345 350
 Phe Phe Val Asn Pro Gly Glu Asp Asp Leu Ser Gln Ile Met Phe Ile
 355 360 365
 Tyr Phe Asn Ser Phe Leu Gln Ser Phe Gln Gly Phe Phe Val Ser Val
 370 375 380
 Phe Tyr Cys Phe Phe Asn Gly Glu Val Arg Ser Ala Val Arg Lys Arg
 385 390 395 400
 Trp His Arg Trp Gln Asp His His Ser Leu Arg Val Pro Met Ala Arg
 405 410 415
 Ala Met Ser Ile Pro Thr Ser Pro Thr Arg Ile Ser Phe His Ser Ile
 420 425 430
 Lys Gln Thr Ala Ala Val
 435

<210> 7
 <211> 1249
 <212> DNA
 <213> homo sapien

<220>
 <221> CDS
 <222> (1)..(525)

 <400> 7
 atg agg ggt ccc tca ggg ccc cca ggc ctc ctc tac gtc cca cac ctc
 48
 Met Arg Gly Pro Ser Gly Pro Pro Gly Leu Leu Tyr Val Pro His Leu
 1 5 10 15
 ctc ctc tgc ctg ctc ctc ctc cca ccg ccg ctc caa tac gca gcc
 96
 Leu Leu Cys Leu Leu Cys Leu Leu Pro Pro Pro Leu Gln Tyr Ala Ala
 20 25 30
 ggg cag agc cag atg ccc aaa gac cag ccc ctg tgg gca ctt ctg gag
 144
 Gly Gln Ser Gln Met Pro Lys Asp Gln Pro Leu Trp Ala Leu Leu Glu
 35 40 45
 cag tac tgc cac acc atc atg acc ctc acc aac ctc tca ggt ccc tac
 192
 Gln Tyr Cys His Thr Ile Met Thr Leu Thr Asn Leu Ser Gly Pro Tyr
 50 55 60
 tcc tac tgc aac acg acc ttg gac cag atc gga acg tgc tgg ccc cgc
 240
 Ser Tyr Cys Asn Thr Thr Leu Asp Gln Ile Gly Thr Cys Trp Pro Arg
 65 70 75 80
 agc gct gcc gga gcc ctc gtg gag agg ccg tgc ccc gag tac ttc aac
 288
 Ser Ala Ala Gly Ala Leu Val Glu Arg Pro Cys Pro Glu Tyr Phe Asn
 85 90 95
 ggc gtc aag tac aac acg acc cgg aat gcc tat cga gaa tgc ttg gag
 336
 Gly Val Lys Tyr Asn Thr Thr Arg Asn Ala Tyr Arg Glu Cys Leu Glu
 100 105 110
 aat ggg acg tgg gcc tca aag atc aac tac tca cag tgt gag ccc att
 384
 Asn Gly Thr Trp Ala Ser Lys Ile Asn Tyr Ser Gln Cys Glu Pro Ile
 115 120 125
 ttg gat gac aag cag agg aag tat gac ctg cac tac cgc atc gcc ctt
 432
 Leu Asp Asp Lys Gln Arg Lys Tyr Asp Leu His Tyr Arg Ile Ala Leu
 130 135 140
 gtc gag cat tcg ctg tct gcg gaa tgt gat tca ctg gaa cct cat cac
 480
 Val Glu His Ser Leu Ser Ala Glu Cys Asp Ser Leu Glu Pro His His
 145 150 155 160
 cac ctt tat cct gcg aaa tgt cat gtg gtt cct gct gca gct cgt
 525
 His Leu Tyr Pro Ala Lys Cys His Val Val Pro Ala Ala Ala Arg
 165 170 175
 tgaccatgaa gtgcacgaga gcaatgaggt ctggtgccgc tgcacatcacca ccatcttcaa
 585
 ctacttcgtg gtgaccaact tcttctggat gtttggaa ggctgttacc tgcacacggc
 645
 cattgtcatg acctactcca ctgagcgcct ggcgaatgtc ctcttcctct tcatacgatg
 705
 gtgcacatcccc ttccccatca tcgtcgccctt ggccatcgcc aagctctact atgagaatga
 765
 acagtgttgg tttggcaagg agcctggcga cctgggtggac tacatctacc aaggccccat
 825

cattctcggt ctcctgatca atttcgtatt tctgttcaac atcgtcagga tcctaattgac
 885
 aaagttacgc gcgtccacca catccgagac aatccagtac aggaaggcag tgaaggccac
 945
 cctggtgctc ctgcccctcc tgggcacatcac ctacatgctc ttcttcgtca atcccgaaa
 1005
 ggacgacctg tcacagatca tgttcatcta tttcaactcc ttccctgcagt cgttccaggg
 1065
 tttcttcgtg tctgtttct actgcttctt caatggagag gtgcgctcag ccgtgaggaa
 1125
 gaggtggcac cgctggcagg accatcaatc cttcgagtc cccatggccc gggccatgtc
 1185
 catccctaca tcacccacac ggatcagctt ccacagcatc aagcagacgg ccgctgtgtg
 1245
 accc
 1249

<210> 8
 <211> 175
 <212> PRT
 <213> homo sapien

<400> 8

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

<210> 9
 <211> 1199
 <212> DNA
 <213> homo sapien

<220>
 <221> CDS
 <222> (1)..(1197)

<400> 9
 atg gga aga gag cct tgg cct gaa gac agg gac ctg ggc ttt cct cag
 48
 Met Gly Arg Glu Pro Trp Pro Glu Asp Arg Asp Leu Gly Phe Pro Gln
 1 5 10 15

ctc ttc tgc caa ggt ccc tac tcc tac tgc aac acg acc ttg gac cag
 96
 Leu Phe Cys Gln Gly Pro Tyr Ser Tyr Cys Asn Thr Thr Leu Asp Gln
 20 25 30
 atc gga acg tgc tgg ccc cgc agc gct gcc gga gcc ctc gtg gag agg
 144
 Ile Gly Thr Cys Trp Pro Arg Ser Ala Ala Gly Ala Leu Val Glu Arg
 35 40 45
 ccg tgc ccc gag tac ttc aac ggc gtc aag tac aac acg acc cgg aat
 192
 Pro Cys Pro Glu Tyr Phe Asn Gly Val Lys Tyr Asn Thr Thr Arg Asn
 50 55 60
 gcc tat cga gaa tgc ttg gag aat ggg acg tgg gcc tca aag atc aac
 240
 Ala Tyr Arg Glu Cys Leu Glu Asn Gly Thr Trp Ala Ser Lys Ile Asn
 65 70 75 80
 tac tca cag tgt gag ccc att ttg gat gac aag cag agg aag tat gac
 288
 Tyr Ser Gln Cys Glu Pro Ile Leu Asp Asp Lys Gln Arg Lys Tyr Asp
 85 90 95
 ctg cac tac cgc atc gcc ctt gtc gtc aac tac ctg ggc cac tgc gta
 336
 Leu His Tyr Arg Ile Ala Leu Val Val Asn Tyr Leu Gly His Cys Val
 100 105 110
 tct gtg gca gcc ctg gtg gcc gcc ttc ctg ctt ttc ctg gcc ctg cgg
 384
 Ser Val Ala Ala Leu Val Ala Ala Phe Leu Leu Phe Leu Ala Leu Arg
 115 120 125
 agc att cgc tgt ctg cgg aat gtg att cac tgg aac ctc atc acc acc
 432
 Ser Ile Arg Cys Leu Arg Asn Val Ile His Trp Asn Leu Ile Thr Thr
 130 135 140
 ttt atc ctg cga aat gtc atg tgg ttc ctg ctg cag ctc gtt gac cat
 480
 Phe Ile Leu Arg Asn Val Met Trp Phe Leu Leu Gln Leu Val Asp His
 145 150 155 160
 gaa gtg cac gag agc aat gag gtc tgg tgc cgc tgc atc acc acc atc
 528
 Glu Val His Glu Ser Asn Glu Val Trp Cys Arg Cys Ile Thr Thr Ile
 165 170 175
 ttc aac tac ttc gtg gtg acc aac ttc ttc tgg atg ttt gtg gaa ggc
 576
 Phe Asn Tyr Phe Val Val Thr Asn Phe Phe Trp Met Phe Val Glu Gly
 180 185 190
 tgc tac ctg cac acg gcc att gtc atg acc tac tcc act gag cgc ctg
 624
 Cys Tyr Leu His Thr Ala Ile Val Met Thr Tyr Ser Thr Glu Arg Leu
 195 200 205
 cgc aag tgc ctc ttc ctc atc gga tgg tgc atc ccc ttc ccc atc
 672
 Arg Lys Cys Leu Phe Leu Phe Ile Gly Trp Cys Ile Pro Phe Pro Ile
 210 215 220
 atc gtc gcc tgg gcc atc ggc aag ctc tac tat gag aat gaa cag tgc
 720
 Ile Val Ala Trp Ala Ile Gly Lys Leu Tyr Tyr Glu Asn Glu Gln Cys
 225 230 235 240
 tgg ttt ggc aag gag cct ggc gac ctg gtg gac tac tac caa ggc
 768
 Trp Phe Gly Lys Glu Pro Gly Asp Leu Val Asp Tyr Ile Tyr Gln Gly
 245 250 255

```

ccc atc att ctc gtg ctc ctg atc aat ttc gta ttt ctg ttc aac atc
816
Pro Ile Ile Leu Val Leu Leu Ile Asn Phe Val Phe Leu Phe Asn Ile
260 265 270
gtc agg atc cta atg aca aag tta cgc gcg tcc acc aca tcc gag aca
864
Val Arg Ile Leu Met Thr Lys Leu Arg Ala Ser Thr Thr Ser Glu Thr
275 280 285
atc cag tac agg aag gca gtg aag gcc acc ctg gtg ctc ctg ccc ctc
912
Ile Gln Tyr Arg Lys Ala Val Lys Ala Thr Leu Val Leu Leu Pro Leu
290 295 300
ctg ggc atc acc tac atg ctc ttc ttc gtc aat ccc ggg gag gac gac
960
Leu Gly Ile Thr Tyr Met Leu Phe Phe Val Asn Pro Gly Glu Asp Asp
305 310 315 320
ctg tca cag atc atg ttc atc tat ttc aac tcc ttc ctg cag tcg ttc
1008
Leu Ser Gln Ile Met Phe Ile Tyr Phe Asn Ser Phe Leu Gln Ser Phe
325 330 335
cag ggt ttc ttc gtg tct gtc tac tgc ttc ttc aat gga gag gtg
1056
Gln Gly Phe Phe Val Ser Val Phe Tyr Cys Phe Phe Asn Gly Glu Val
340 345 350
cgc tca gcc gtg agg aag agg tgg cac cgc tgg cag gac cat cac tcc
1104
Arg Ser Ala Val Arg Lys Arg Trp His Arg Trp Gln Asp His His Ser
355 360 365
ctt cga gtc ccc atg gcc cgg gcc atg tcc atc cct aca tca ccc aca
1152
Leu Arg Val Pro Met Ala Arg Ala Met Ser Ile Pro Thr Ser Pro Thr
370 375 380
cggtt atc agc ttc cac agc atc aag cag acg gcc gct gtg tga ccc ct
1199
Arg Ile Ser Phe His Ser Ile Lys Gln Thr Ala Ala Val Pro
385 390 395

```

<210> 10
<211> 397
<212> PRT
<213> homo sapien

<400> 10

 Met Gly Arg Glu Pro Trp Pro Glu Asp Arg Asp Leu Gly Phe Pro Gln

 1 5 10 15

 Leu Phe Cys Gln Gly Pro Tyr Ser Tyr Cys Asn Thr Thr Leu Asp Gln

 20 25 30

 Ile Gly Thr Cys Trp Pro Arg Ser Ala Ala Gly Ala Leu Val Glu Arg

 35 40 45

 Pro Cys Pro Glu Tyr Phe Asn Gly Val Lys Tyr Asn Thr Thr Arg Asn

 50 55 60

 Ala Tyr Arg Glu Cys Leu Glu Asn Gly Thr Trp Ala Ser Lys Ile Asn

 65 70 75 80

 Tyr Ser Gln Cys Glu Pro Ile Leu Asp Asp Lys Gln Arg Lys Tyr Asp

 85 90 95

 Leu His Tyr Arg Ile Ala Leu Val Val Asn Tyr Leu Gly His Cys Val

 100 105 110

 Ser Val Ala Ala Leu Val Ala Ala Phe Leu Leu Phe Leu Ala Leu Arg

 115 120 125

 Ser Ile Arg Cys Leu Arg Asn Val Ile His Trp Asn Leu Ile Thr Thr

130	135	140
Phe Ile Leu Arg Asn Val Met Trp Phe Leu Leu Gln Leu Val Asp His		
145	150	155
Glu Val His Glu Ser Asn Glu Val Trp Cys Arg Cys Ile Thr Thr Ile		160
165	170	175
Phe Asn Tyr Phe Val Val Thr Asn Phe Phe Trp Met Phe Val Glu Gly		
180	185	190
Cys Tyr Leu His Thr Ala Ile Val Met Thr Tyr Ser Thr Glu Arg Leu		
195	200	205
Arg Lys Cys Leu Phe Leu Phe Ile Gly Trp Cys Ile Pro Phe Pro Ile		
210	215	220
Ile Val Ala Trp Ala Ile Gly Lys Leu Tyr Tyr Glu Asn Glu Gln Cys		
225	230	235
Trp Phe Gly Lys Glu Pro Gly Asp Leu Val Asp Tyr Ile Tyr Gln Gly		240
245	250	255
Pro Ile Ile Leu Val Leu Leu Ile Asn Phe Val Phe Leu Phe Asn Ile		
260	265	270
Val Arg Ile Leu Met Thr Lys Leu Arg Ala Ser Thr Thr Ser Glu Thr		
275	280	285
Ile Gln Tyr Arg Lys Ala Val Lys Ala Thr Leu Val Leu Leu Pro Leu		
290	295	300
Leu Gly Ile Thr Tyr Met Leu Phe Phe Val Asn Pro Gly Glu Asp Asp		
305	310	315
Leu Ser Gln Ile Met Phe Ile Tyr Phe Asn Ser Phe Leu Gln Ser Phe		
325	330	335
Gln Gly Phe Phe Val Ser Val Phe Tyr Cys Phe Phe Asn Gly Glu Val		
340	345	350
Arg Ser Ala Val Arg Lys Arg Trp His Arg Trp Gln Asp His His Ser		
355	360	365
Leu Arg Val Pro Met Ala Arg Ala Met Ser Ile Pro Thr Ser Pro Thr		
370	375	380
Arg Ile Ser Phe His Ser Ile Lys Gln Thr Ala Ala Val		
385	390	395

<210> 11
<211> 1128
<212> DNA
<213> homo sapien

<220>
<221> CDS
<222> (1)..(402)

<400> 11
atg gga aga gag cct tgg cct gaa gac agg gac ctg ggc ttt cct cag
48
Met Gly Arg Glu Pro Trp Pro Glu Asp Arg Asp Leu Gly Phe Pro Gln
1 5 10 15
ctc ttc tgc caa ggt ccc tac tcc tac tgc aac acg acc ttg gac cag
96
Leu Phe Cys Gln Gly Pro Tyr Ser Tyr Cys Asn Thr Thr Leu Asp Gln
20 25 30
atc gga acg tgc tgg ccc cgc agc gct gcc gga gcc ctc gtg gag agg
144
Ile Gly Thr Cys Trp Pro Arg Ser Ala Ala Gly Ala Leu Val Glu Arg
35 40 45
ccg tgc ccc gag tac ttc aac ggc gtc aag tac aac acg acc cgg aat
192
Pro Cys Pro Glu Tyr Phe Asn Gly Val Lys Tyr Asn Thr Thr Arg Asn
50 55 60

gcc tat cga gaa tgc ttg gag aat ggg acg tgg gcc tca aag atc aac
 240
 Ala Tyr Arg Glu Cys Leu Glu Asn Gly Thr Trp Ala Ser Lys Ile Asn
 65 70 75 80
 tac tca cag tgt gag ccc att ttg gat gac aag cag agg aag tat gac
 288
 Tyr Ser Gln Cys Glu Pro Ile Leu Asp Asp Lys Gln Arg Lys Tyr Asp
 85 90 95
 ctg cac tac cgc atc gcc ctt gtc gag cat tcg ctg tct gcg gaa tgt
 336
 Leu His Tyr Arg Ile Ala Leu Val Glu His Ser Leu Ser Ala Glu Cys
 100 105 110
 gat tca ctg gaa cct cat cac cac ctt tat cct gcg aaa tgt cat gtg
 384
 Asp Ser Leu Glu Pro His His His Leu Tyr Pro Ala Lys Cys His Val
 115 120 125
 gtt cct gct gca gct cgt tgaccatgaa gtgcacgaga gcaatgaggt
 432
 Val Pro Ala Ala Ala Arg
 130
 ctgggtgccgc tgcatcacca ccatttcaa ctacttcgtg gtgaccaact tcttcggat
 492
 gtttggaa ggctgtacc tgcacacggc cattgtcatg acctactcca ctgagcgcct
 552
 ggcgaagtgc ctcttcctct tcacatggatg gtgcattcccc ttccccatca tcgtgcctg
 612
 gcccattggc aagctctact atgagaatga acagtgtgg tttggcaagg agcctggcga
 672
 cctgggtggac tacatctacc aaggccccat catttcgtg ctccatgtca atttcgtatt
 732
 tctgttcaac atcgtcagga tcctaattgac aaagttacgc ggcgtccacca catccgagac
 792
 aatccagtac aggaaggcag tgaaggccac cctggtgctc ctgcccctcc tgggcatcac
 852
 ctacatgctc ttcttcgtca atccggggg ggacgacctg tcacagatca tggatcatcta
 912
 tttcaactcc ttccctgcagt cgttccaggg tttcttcgtg tctgtcttct actgcttctt
 972
 caatggagag gtgcgctcag ccgtgaggaa gaggtggcac cgctggcagg accatcactc
 1032
 ccttcgagtc cccatggccc gggccatgtc catccataca tcacccacac ggtcagctt
 1092
 ccacagcatc aagcagacgg ccgctgtgtg accccct
 1128

<210> 12
 <211> 134
 <212> PRT
 <213> homo sapien

<400> 12

Met Gly Arg Glu Pro Trp Pro Glu Asp Arg Asp Leu Gly Phe Pro Gln
 1 5 10 15
 Leu Phe Cys Gln Gly Pro Tyr Ser Tyr Cys Asn Thr Thr Leu Asp Gln
 20 25 30
 Ile Gly Thr Cys Trp Pro Arg Ser Ala Ala Gly Ala Leu Val Glu Arg
 35 40 45
 Pro Cys Pro Glu Tyr Phe Asn Gly Val Lys Tyr Asn Thr Thr Arg Asn
 50 55 60
 Ala Tyr Arg Glu Cys Leu Glu Asn Gly Thr Trp Ala Ser Lys Ile Asn

65 70 75 80
 Tyr Ser Gln Cys Glu Pro Ile Leu Asp Asp Lys Gln Arg Lys Tyr Asp
 85 90 95
 Leu His Tyr Arg Ile Ala Leu Val Glu His Ser Leu Ser Ala Glu Cys
 100 105 110
 Asp Ser Leu Glu Pro His His His Leu Tyr Pro Ala Lys Cys His Val
 115 120 125
 Val Pro Ala Ala Ala Arg
 130

<210> 13
 <211> 32990
 <212> DNA
 <213> homo sapien

<220>
 <221> gene
 <222> (1)..(32990)

<400> 13
 tcacacagcg gccgtctgct tcatgtgtg gaagctgatc cgtgtgggtg atgttagggat
 60
 ggcacatggcc cgggcccattgg ggactcgaag ggagtgtatgg tcctgccacgc ggtgccacct
 120
 ctccctcactg gctgagcgca cctgtgggaa aggcaaggc tcagctggct cccaggacc
 180
 aaccctgggc ttctgggacc atcccctcct ctgtttctg ctctcatgg tcgactgcca
 240
 ccctcatgac aaggaactgt ctgcttccaa aacaggtcct tccctgatgc tggtgcctct
 300
 ctccaggggc ctcttcctta tcctttcctt ggagaagctt gactccacac ctcccttact
 360
 ccacactgtc ctcccagatc atccagtttc ctctggacac actgccttcc ctccctgac
 420
 tcactctgcc acctcaaaat tagcatcttgg accagaga ggcgcagctt gagaacttggc
 480
 cccctcagcc ttcttgggccc cctgctcctt gcagggcggtt ggtgggtgtgc gcccagctca
 540
 cacacctggt gcgcctcctc ctccctcctt aagaccgtcc cctctgcacc cctccacctt
 600
 ctctaccccttcc tgcttagactc ccctgccttga ctaaccatgg catttactgc ctgaaaggga
 660
 aagctctctg tcttggtcat ttctaaagtt ggctctaccc ggtgccattt aagcttccatt
 720
 tgtttggtaa ttgaattttt ttggaaattt gatgggtgtct ctcacatcct agcttgcaca
 780
 tggttcgaac agaataaaatg ctccataaaac atcccttggt tgtttgcctt actggccccc
 840
 tctcttcccttc tcatccagcc cacaggaccc gtctctgctt atggactttt gcaagtggaaat
 900
 gtgctatcat ccacagagac acaggcctgc atagacaccc cggccagtgcc ctgtgcagag
 960
 acaggcaagt gtattcctct ggagtggtcg gaagctccca ggccttaggg gctgtgtctg
 1020
 ctctttgctc catgtcctca tccacccctt cccaggtacg agtgcgccttc ttccctgcgt
 1080
 cagctgcact ggggctttgg gcctcatgtat tggcttgac atgccaagct tcacggccgc
 1140
 ctgactcccc agagcctgctt ctgggtgtgtt gggctttctt gtggcaggtt ggcggggaaat
 1200

gtgctggct ctcttctgct ccctgagtcc atacagaccc gattgcttgg cacacatctc
 1260
 tttctgaaag cttccttctg tcaccaccc tcgtctggct tctcttcc atgaggccct
 1320
 gcggcctggg cttccttgct caccagctc tgagtgcaca tgtgggtgc tggcatgtta
 1380
 tttcgccggct gcatagttga ctttctaaac tggctccagc ccctgtgaga aggattcctg
 1440
 ttcccctaag gccgggtggt atctcaaggc ttcctctccc tggcactcca gcaaggccgg
 1500
 gcagcaccaa tggagctgcc ctggagtgccc gtctgaggac ctggatatcc cagggccaccc
 1560
 cgagggccca gttcataacc tctccattga agaagcagta gaagacagac acgaagaaac
 1620

 cctggaaagg agggaaagga gggagtggtc agtgacccat ctgcggagct gttctgcctg
 1680
 gtgggggtggc actggggaca agatggtgcc ggggggacaa tttgacccag gaaccctgg
 1740
 agtcagagct gggtggggc ctagcctcag ggtcagata ttccacggtc cactctgagt
 1800
 gagcagtggt cacaggcctg gagccaggca gactccggc tcctcaatat tgtgtgaccc
 1860
 gtcccagggc atttgggtgg gctgcctgca tcctcagaac aggtgctggg acagtaggtg
 1920
 cgggctggta ggtgaagac ccaggctggg atggggataa ctggccaggc tcaggctgag
 1980
 catgtacggg cttctaactc tgtggatgt actgagccat gggtgtgcca gtcccacggg
 2040
 acccccttag tgagggcatg gctgccagag cagcctcagg aaagctcact gtggggccccc
 2100
 catctggta caggccccac ctggaaacgac tgcaggaagg agttgaaata gatgaacatg
 2160
 atctgtgaca ggtcgccctc cccgggattt acgaagaaga gcatgttaggt gatgccagg
 2220
 aggggcagga gcaccagggt ggccttcaact gccttcctgg gggcgagagg tggacacagg
 2280
 tctgagccca tgcggcaggc agggcctcac ccagtggcgc gccccggagaca gtggtggtgt
 2340
 ttcttcggg agcttgagcg agctccctgg gtgaggcctg gggcagaggg ctctgccagg
 2400
 aacccgggg acggcccgat gacccctcc tgcggatgt ggtggacgcg cgtaactttc tcattaggat
 2460
 cacatacctg tactggattt tctcgatgt ggtggacgcg cgtaactttc tcattaggat
 2520
 cctgacgatg ttgaacagaa atacgaaatt gatctggagg gagggcgggc atggaaaga
 2580
 gggaaaagga aggagcacgt gtttgagatg agccgagagg cagccccctt cccgcagac
 2640
 ccctggaaac cgatgtccca cgcacacacc tgcggatgt ggtggacgcg cgtaactttc tcattaggat
 2700
 catcccaagg cctgtgctcc tccctcgcca ggtggggag acagccccat tttccaggc
 2760
 ctgtgtctcc agcccaggac tgaggaggaa aggtggcgc ctctgtgggt cgtgaccgag
 2820
 agcacggggc atgcctctg aggctgagaa agcccccaac cctctccca gtatagggac
 2880
 cctatgggag ccccttccc ctcactgcca cggggccctt acttggatgt gtcggggc
 2940
 tctgagtc ttcactgttc actggctcta ggtcgccctg aggtccttct gactgaggac
 3000

agcaactgcc a tgggtggggcg gacaaggcca gatggaggga ttaagaactc agttgccgaa
 3060
 tgagtgatta atctgccagc atcccacccc tgtgcccagg aactccagag ctttcgtagg
 3120
 caccctagac aggaaagatg gatgggctgt ggggctggct ctgaggcccc aggaagggcc
 3180
 tcatttatct tcattctgag ttttccctca gcaccagagc ctccccaccc agcacctggc
 3240
 cctggaactg gcctgtgtgg ctccccagct gctccctgct gggtgtgggt ccccatggc
 3300
 ctgcgagtgt ctgttcatct gtattggctg tgactatgac tgtgttgtcc cctggagct
 3360
 gcagtgtggg tctgtgggtc tattcttaa gcacttgtgc aagtgtgttt gcctcacaga
 3420
 gcgtgtgcat gtgtgagcat ggttctgggg cccgggcaca tctgcgttcc tttcacctg
 3480
 cctccgagca taggcaggca ggcaggcaag ggagtgtgtg tcggcctgac acagtggggc
 3540
 tgggtttctga caaagtgcgtc gtgggcccaga gggagacaca tggtagccca ggggtgtctg
 3600
 ctggctcggt tgcacacacc cactcacaca cacacacaca cacacacaca cctctgacac
 3660
 tctgtcaaga aaatccattt cttcttaagc ttgggtggg cccttctaacc cctcccaagt
 3720
 tctgaatcct ggctgggaga ggagggacac aggatacaga cagatatacct gagactgtca
 3780
 acctgttagcc tctggatgct ccccaacaagc ttccgggttag ccctaggggc agggagagct
 3840
 cacacccac ctcagccact gtctactgct ctcttaccac atagtggtat ggcctgggg
 3900
 tcccagggag ggctgggat gaaagcctt cagcgggctt gccatgaccc gatacccca
 3960
 gctttctcctt agggctcacc taatttccag ctccctgggtc tctggcatat tccctggat
 4020
 catggacac aagtccaccc aacatataag tgaatctaa ttttctata caagtaatac
 4080
 aaagtatttg acctttctg attcttccgg acttccctga gagtagaaac tggtagaaatc
 4140
 aaaatataattt ctcatcttccat ccacattttt ttgaatttcaa tattatctt ccaacatcca
 4200
 tcaacccact aatcaattca attatccacc cacttttattt atctatccat ctatccactt
 4260
 acctaccaac cagtttccatc atttacccac ctacttttattt ccatccgtcc atccatccat
 4320
 ccatccagcc attcatccat tcatccttcc ttccttccat ccatccatcc atctacccaa
 4380
 caaccaatttcc attcaatttattt ccacctactg ttttatctt ccatccacgt acccatccat
 4440
 ctacccatcc atctacctac ccaataacca atccatttccat ttatccaccc tttttttat
 4500
 ctggcccatcc actggcccat ccgtctatcc attcatctac cttagcaacca aggccttgcc
 4560
 tccatcctta cacttggcag acattttagtt atgtggcccc actcatctca gcttgggtct
 4620
 atgtaacctg aaaatcccttcc tttcccttgc gctgtttcc cagcctcccc ctacacacca
 4680
 ccacaggtac cttccctctt ctccagggag cttctgggtc accctgcccc cactgagctc
 4740
 tccggcttcc catttttggc ctctacccctt ctgtaaaaaa cacactccag ttcttttagta
 4800

agaactgtct attgtaccct gccctcggttc agcctgtttc tggttttatt taattgttca
4860
ttcattcaat gaatcaatgg accatgtgtc aggctctgag ccaggcatca gcgatggcga
4920
agtggacagg cagtcacagc ccctgcttgc agagggcttg tggcttagct gggggcctgg
4980
aggggttggg ggtgggggag atgacatctt tcctaaggtg tcaaaaagac ctggagaaat
5040
ggcagatggg ctaagaactg aagatgaggc ccccctccca aatgaactga agcaccaagt
5100
cctagtctca tggctcaaattt tttgactgtt ccaatttgc a gttgtgtggc cttggtaag
5160
ttgcttaacc tcatgagctc ctgcttcctt gtctgcccc gaggggccta accacagagc
5220
ttgccttaga agatcagcat gggctgaaat agtcacagag accactaaga cctgtgtctg
5280
gccttagtgc tgccaatcct ttccccatgac accaacaatg acaacaccag caagaagtca
5340
ggcttcctgc aaattcaacc acctgattca cttctcagaa ccagaaataaa gtgagtgct
5400
cttagcagca gaaatgcagc tgtgagctgc tggggagtgt agggatggca ggaacaccaa
5460
acacgtgtcc aacctcgagg acaacaggca gatggtgggg acaaaggcagg agacctgcct
5520
ttgagcctct gctctgctag ctattagtgg tgaccttggg caagccactt catctctc
5580
agcttcttcc tctgtaaaat ggagagcata accctacttc ttttaggact tgagcatgca
5640
agggttggaa ataatgtaga tagagcttct ggaaggggctc agcttggaaat gtggaagccc
5700
ctgttgggc aaagagctgg gcccaaagga aagaaggaaa gggctcgta gcacagacgc
5760
aggggtggag ccagagaatg ggtctggaag aggaagcatg gagctgggtc ttgacactga
5820
gctcttgcgt tcctccctc cctttctt ggcatttcct caccaccctt ccaacatctt
5880
tacagttggg gcttactctt tgccaccatt ggtccccat ccaggccagc cttccttggc
5940
tcagttccag gctcatcaaa agccttattt ctcaggatg gcccttggca gctgggtgg
6000
agggaaagga tctcatgggg tttccataca cactgagggg tgagtgactc actgcagctg
6060
ggagtccaaag gacaatccta ctagctgctg acaaaacctg gcttctgcct ctgaagcgaa
6120
ccaagccctg acggtgaaat ttccagcagt gggtaaaaaa ctttccttagt cttcatcagt
6180
tcacatataat caaggacact ccaaagatgg aggtcaaattt aacccgaag tcacagaaaa
6240
aggagagtaa gaaaataaaa gagataaata ataattgtct ccaaggtaga caaaagaaaag
6300
ccccaccagg cacctctgt gtgcaggagg agatggact ctcgctgagt ctcccagaag
6360
gaactgtggc ccttaagga agctggtaca tacaggagcc tggaaattact gagcctgaaa
6420
taagtcttac gtaaatgcca atttcttac gggctttttt gtatgtctt ctttctttat
6480
ctctctctct ttttttaaa ggtatgatataaatgggtt cttttgcag aatgcagc
6540
ctacttgaac tctctcttaa ttaaagctca ttgtggatta aagatgtgtg gcaattatac
6600

cttagatgca aatacatttgc acaaaggca gggagctgga caaggaacaa ttcctcatta
 6660
 tgctgcattcc cggtattcac cgcttcaggc tgggggtggt gggctggacc tggagagggt
 6720
 cactggggc tgagggAAC gtaggtctgg aatgaacaga agcccgacca ctcttctt
 6780
 gtcaccagct gcctgctcac ctggctttct cagcattgt a cagtgtcaga ccgggaggc
 6840
 acaggaagtc taccatcag atccttctc ctcccagaga aagagccctg aggtcagcag
 6900
 agtggtgcag ggctgcgc a gggcatcgag catgaggaag ggaaggccgg gtgggtctgg
 6960
 ggcacagggt ggcattggaca gggta cgggg gtggcatata gagtgtgtgc gcaggcgtgc
 7020
 ccatgccaca tcttccag cgtctctgccc tgggtggcc ctcttctgtc ctcttggcac
 7080
 ccagccccat cccagccacc gctgagggt taccaggagc acgagaatga tggggcctt
 7140
 gtagatgttag tccaccagg t cgcaggctc cttgcacaa ac cagcacctgt gaagatgggg
 7200
 tggctgttagg gggcctcctg agctggaaact gggggacccc cacagacttgg gcccagggt
 7260
 cctccagctt tacccttcctt gagacccaaat ctccaggact gccccttccc agaaagcctt
 7320
 ggtggaaata ccagctccac tgaccaccct taccaccag gctgtctttt gctctacatg
 7380
 gtgggtctga tctccca cccactgagg acaggccatg ggtctcttcc tccctctt
 7440
 ttacctgccc cattgcac tgc ttaggcat aggactggc atgcaggagg cacaggaata
 7500
 aaggagagc tgggctcccc cttgctggcc cagggcaat tgccttaggc cacctgttcc
 7560
 catattgctt gaaacccgct gttccgtgtg ccaggcactc actcgccac acaaagcagc
 7620
 agatgtggaa gtggagcctt cttacccca gcttggatt ttgcgtctttt gagccaaatgc
 7680
 ttgtgcaccc agtgaggggaa gctgctgcct ggctggcctc actggctgccc agaggcataa
 7740
 ttcatccatcc qggatgtg gtcaggggaa atgttttagg gtctgggtcg gccagcaact
 7800
 actggaccac agggaggggc catcaactttt gggaaactca ttccctctgaa atcttggagg
 7860
 tcatccatcc cagacccgtc tca gatagga ttctccctgt ctccagggtt ttctggcctc
 7920
 tgcttggagc cttcccttga caggaagctc accaccatacc aaggcagccc tcttcttgg
 7980
 cagttttac tcttacaagc tcttccataat gttgagctcc tggaaactcta ctccatttt
 8040
 gccccccactc tgagaacaga tcccccctgc tcttccagg ctctgagaca gacccttgg
 8100
 agatgtcat acaggctaat gttccctggt ccccgctgct ctcagctggt gacacctgac
 8160
 ttctggctg ctccctggct gacccatggc cagagttgct gtcatttggg ctgtggctca
 8220
 gcatattaca gcttctggca tagggagaac ctcaggatgg gacactgcac ccaggcagag
 8280
 actgagactt cagaaaaaaa aacaagctct gttaaatgct catggactct ttaaatgctc
 8340
 atggacagct ctggatgtca tgtgcggcct tggcccttc cacatacccc gggccagagc
 8400

tgccacatcc aagtctgagg aggccttaca aagaccaaag gggcatgcg ttgggtcggg
8460
gggcacttca aggaaccaca attcctaact taattcagca aagttccttg agccctacca
8520
tgtgagtacc tctgagtgcg ggactatTTT tttcaattta aggacaacag gaacgtggat
8580
ctgtcttaggt gtggctacaa ttccctgctcc acggcttggc cagagccaa ggctgacctg
8640
tccttaacacc cccattccct tccccatgac ccccatggc tggcccatcc acttactgtt
8700
cattctcata gtagagcttgcgatggccc aggccgacat gatggggaaag gggatgcctg
8760
aaagaaggaa agacttgggc tgcagggac agatggacag ggactttctt gtaggacata
8820
ccgtgggta ccacaacagg gctaggatca tgTTTTact cttccaaacag caggccaccc
8880
acaaccccaag gggtgccctg tccctcaaca cctggccat gcccctgccc ctctctccaa
8940
gcagggctgc gtgatTTTgt gatagagaag tagagccagc cagTTTctga gccagagaca
9000
aaggccttgc gacaggtcct tcctggcagg gggagaagag ctatTTgagg aatatcctt
9060
ggagagatcc ttgcTTgtt cctcaccaggc atggagggaa gtagctgcat ccacccggact
9120
gctgctgggg tggagggcag gcccaggctc tggcTTgg acccaagacg aaggaaatg
9180
gcctggtaga aagtccccc cccaaactagg ccctgctgccc cctgggaccc tcacaccatc
9240
cgatgaagag gaagaggcac ttgcgcaggc gctcagtgga gtaggtcatg acaatggccg
9300
tgtgcaggta gcagcTTcc acaaacatcc agaagaagtt ggtcaccacg aagttagtga
9360
agatgggtgtt gatgcagcgg caccagacct gtgtgcaggg cagagggct gtcaggaggc
9420
agcttggggc ccaggttagga cataccatc cccaggcagg gcaacaagac acaggcgtcc
9480
ccaaagggggg ttgcgtggaca tgccatcaaa taccagcgaa cctcactctg aaaagctca
9540
tccttctcca gtgcTTttc acacttcaat attaagttaa ggtgccattc tccactgggg
9600
ccaaacgtgtt ttTTTTact tctctcaac tctttctaat ttTcattct agtaaagaga
9660
gcaagagtct ggctctgagc ttctgttagg cagaggctgg aattcaacca tcttgggttg
9720
ttttcattttt agtttttttgc agtcatttc aatccatagc aagtcagacc tgcttccctt
9780
tggggatggg atatgaaatt tcatttagaa aaaaatgaaa aaataaaaagt gagtaaagt
9840
agtcaagggt gtatgaagtgc gggctgcggc cagggagggg attctccaaa gactctgggt
9900
ttgggaactt ctggacttgg cacaattatt actagctctg gagggagact tgcaaagtac
9960
acggcccccg gcaagtcaact gcacccctct gaacctcaaa aagtaaccct gccttccagg
10020
gtgggttgaag gagatagagg atggcaaaga caggcatgag ggttagctgtt ttgtggctgt
10080
ggttgtgcct gctgtggttc cgtgcTTcc cagcagaggga gaatgtgtcc ctgtccctct
10140
gagcaaggcc acccttcccc aggcaccaag gctaccccttcc caaaggagggc agggagggga
10200

agaccctgcc ccttagagc cacgcagtgg gccatggcag ggccagggtc tagacattgg
 10260
 gcttccaggc cagagctcct ctaacagac cccacctggt catcttcccc acaggctcat
 10320
 ccccaggca ccctgagagc caaggctggg acatggggaa ttagtggagcc agctcaaggt
 10380
 ccggggagct gtgcgtcagg ggctctgctt ctgcacacag cccatcctct ctgtctggct
 10440
 ctgacagccc cagttctcag ataaccctcc tgtgctgagc tttggctgc ttctggctt
 10500
 ccctgcacag tccctgtggc tggctgtctc ctttttctg agaaagtccct gctgagggtgg
 10560
 gaagctacca agcccctcc cccaacccta cttttcatcc agggttgatg atgttctatt
 10620
 agcacaagcc cacgttggag ctagaaggca ccctcaactc gagtgaacct gtttattct
 10680
 gaggcagtgt tctatgtggg accatgttaa ggatcacacc aggctgggtgt ctgctcaggc
 10740
 acaggccacc caaaggaaat gtactgagaa gtctctgtcg gtgtgccaca gggctctgt
 10800
 atggcccaag actagtctac agttttacaa tagcttggac acagtacaca gcaatggaca
 10860
 gaaatccaga gtggacagtt agcatgtggg atagcccattt atgttagaggt atcatcactg
 10920
 catgtgaccc tggcgagtca cttAACCTCT gtgagtctca gtttccatgt ctatgtatg
 10980
 gggaaaatga tccctgctgg tctcattagg attaagttag agaaagctca acagaggtta
 11040
 gttctagctt cttttctca aagggtctt tgagggcacc tgaatccaca agatgaggag
 11100
 tggacttagga taaatgtgtc tagagtca ctttgtaagc tcccagcctg gcagcttc
 11160
 gctcctccca gcccagctct gttggacaa tggctagggt ggaggtgagc tcaggctgg
 11220
 ttttgcacct gagccacagc ccagatgaca gcattctggc catgggtcag ccaaggagca
 11280
 gccagaaagt cagttgtcac cagctgagag tagcagggac tggtaacat cagtctctgt
 11340
 gtgcaaatct ccaaggggtc tgtctcagag cctggagaga gcagagggga tctgttctca
 11400
 gagtgaaaaa aagaatgggg tggagttgca gaagttcaac attaggaaga gttttaaga
 11460
 gtaaaagctg ctcaagaaga agagggctgc tttagcaggt agtgagcttc ctgacaaagg
 11520
 aggggtccaa gcagagacca gacagacaga tgggtcccc cggagcccaag agccccccag
 11580
 gtatagcccc gagtctcccc gagcaatgac ctcattgctc tcgtgcactt catggtaac
 11640
 gagctgcagc aggaaccaca tgacatttcg caggataaag gtgggtatga gttccagtg
 11700
 aatcacattc cgccagacagc gaatgctcct gtgggagggtg caggtcaggg gtcagccagg
 11760
 ttcaggggtc aactggact gggcccccc tgagggcagg tagagactca gcctggatg
 11820
 agggcagggc tgcacttagga gccacttccc acccatggtg gccacagttt ggcctctgag
 11880
 tccagctccc actctgcacc ccacatgcct gctggtgatt catgccctgg caccccaccc
 11940
 aaaccccaact ttctccacgg gccctttat ctgctggcc ccagaatgga ggtgagaatg
 12000

tctgggagag gtgaagggggg tgctgttaggg ggagggatga ggagaaagca aggcggaagg
 12060
 gcagactcac cgcaaggccca ggaaaagcag gaaggcggcc accaggcgtg ccacagatac
 12120
 gcagtgccccc aggtagttga cgacaaggc gatgcggtag tgcaggtcat acttcctctg
 12180
 ctggacagac agacatgggc agggcagatg gaggcatggg cacgtggggg tggggctggg
 12240
 tattccagcc gtggccacct ctgtgtcctg accttggggg cagaagtgtc ccaggtgtca
 12300
 tgcgcgtgcc tggctcttag gttcgttcc tggctggccct gggtggctct tggctttata
 12360
 aatggctgtg gtcaggcctt ccaacatgca tttatttta ttttttaga gatggggtat
 12420
 cactctgttg cccaagctgg ggtccagtgg ctattcacag gtgtgatcac agtgcaccgc
 12480
 agccttgatc tgcaggcctg aagcaatcct cccacccatcag cctctgaagt agctggact
 12540
 acagacaagt gccactgtgc ctggcaccaa catgcattt tggggcaca ttttgagagg
 12600
 tatgggtaca gattctttt taaaattcta ggctctagaa tgcttctcct gagtttagtc
 12660
 tcagccccgga tcccagctgg ctgagtgact gagggtgagt cgattttat ttctggaccc
 12720
 cagttttagt gtctatacaa tggggccaac ctgccatcct acctaagaag caatggccta
 12780
 tgggaggccct ggggtggggt gcatttagat tagtctgccc tggagaccac acgaggggg
 12840
 gcactgtcta tagagaattt agaaattcta ttcaaactga ataaaagtca gttgacttt
 12900
 aattatcacc aagtgttggc aatttggaaac aaagtttagt atgaaataact cttcttgcc
 12960
 agggagagcc acccctccac cttaccaac ccctgtgtt attctccact gctggagggg
 13020
 cagagagagt ggaggtgagg accaagggtt ggaggccccctt ctgcccattt agtggcctcc
 13080
 ctgcagaacc cctgtggctc acattttgcgaaatcactt ctcaggccgg gtgaatgttt
 13140
 ttccctctca atttggactt catctggggc aaagtcccag cccactgag gactatgctg
 13200
 ggtaatgagg agggatgga gctgagatgt atccttcag gtcaggaatg aggcgtagcc
 13260
 tcaaggagag gagttcggtg ctcagcagc actgaattga attccagagc gagagctgga
 13320
 gcagggctcc agagaggcag agcaggaccc agcttctcat ggggacagcc ctggggggc
 13380
 tgcattttgtt tccatgtgga gtcacagaat cagaatgcca gatggggaaa ctgaggccca
 13440
 gagggagagg aaggtgtgca gtcacacaaac ccctaagatg ttaggagcat tgattacacg
 13500
 gttattccct ttttgtgtca catgccaccc tggcttgc tcttagcagc ccaagctctc
 13560
 ttgaaggtag gactccatgt ccccttctc tactccccta cagtgtcact aagcacaggg
 13620
 caggacactg gggcagggg caggaggtac agaaggagt gactgggtga caaaaggact
 13680
 ggtctgcccc ctttggatc ttctctgctc aacctgagtc caaatacctg tgtgaggcct
 13740
 ggggtcacag caggtgaggg ccactcacct tgtcatccaa aatgggtca cactgtgagt
 13800

agttgatctt tgaggcccac gtcccattct ccaagcattc tcgataggca ttccctacaa
 13860
 aaaatgccaa ctgccaagag tcaggtcact cccctcctca agaaccctcc ctggctccct
 13920
 ggtgcccaca ggataaggtg tacgcacctc agcttccta gggcagcaaa tacaatgtgc
 13980
 atctgaaagt ttacatgtaa gtcactttagt tgaaaatagc acaaaggct catcttacaa
 14040
 aatagctaaa ggcatgctaa aaccatttt agacacaatc tttgactagc attttgtaca
 14100
 ctttcatttt gttcatttgc ttcaaaactg aatccaactg tgaggagtgg gttgtgcctg
 14160
 aagattcact gcttcagca tgccacataa tttacatcct tgtctcaatt gttttatatt
 14220
 tatgaggtcc atacttcacc aatatcagca cttccatttt tataatagct accattata
 14280
 tagcatatgt ttcccatgta ccacttcctc ttcttaactgc tatggatata catgtcaatt
 14340
 caatctccac aaccatccca tgaaggaggg attaatatga acccatttca aagaggagga
 14400
 aactgaggca taaagagatt aagctactgg ccaaggatac aaagatgaca agaggattca
 14460
 aactcagaag ttgtggcttc gggcttctg ttcttaacct ttaggcata tcaagtagtg
 14520
 gcaaacagga atgagtgaat gctggggact cagggctggg ccaccgcct cgggctgctg
 14580
 ctgcctggga ggctaccacg agatgtttct aaatgtcaa gaaccgcct ctgtatattaa
 14640
 gttggacat gatthaagtg tgaaaagaca tttcaagcag caatgtcctg tgcacaaaag
 14700
 tgcaggtttt gtcaaggggg aaggtaggaa agttaaaaaa tgctcacatt tgccttctc
 14760
 tctatctcag ctccaaactgt caaggtctgg ctccctaagcc accccctcca tacaccagct
 14820
 ccaatcagga ggaaggccctc ccacctgaga aggcccgag ctctcgaagc ctgcctctgc
 14880
 ttggccttca tcactagtgt gtttctgacc taccgggtgg gttacaaatg ctgaacgttc
 14940
 cattgtttgg ggggtatttg taccagggtt cagctctctg actaatggc agttgtctgt
 15000
 gaattttct ttcttagcata tgtagatgt acaatgtaaa gctaattaaa ataaatagct
 15060
 tgcagagcac agagttgcag agctggaggg ggaaccttag gttgttttg gaagcagtgt
 15120
 gttctgaaat aatitagttt cttaaaaacc cacttcgtt gagcccgatg agttgaaagc
 15180
 aatggaaattt ggaaggagca cttgccgaag agcaaaatca atggggaaaga ttctattagc
 15240
 ttaattgttt tttagtttg tgctggagc tcatccattc ttcaaaccctt gggacgtgac
 15300
 tggccttattt cttccctcctg ggccaaggcc catccctggc agggccctgc catccccctg
 15360
 ccaagtgagt cagggaatgc cctggtctga tgctgattct gacttcagg aagaggaagc
 15420
 ctgctccccca cccctagcca tggcgcccaa ctccccaggt gggatctaat ttgataccta
 15480
 gcactatctt tccttaccaa cattcgtgct catgaaagag aaaggttacc tcaactcgtg
 15540
 agggtcagtg atagtgatgt cactgaacta aaaaagcaaa agtatgtaaag gagggtaagt
 15600

tcttttgtga aatgaacagt ccctccctga tgggggtta cggtgcctct gaacagtcta
15660
tgtgaggtga ggcagaccag gatcctgcct gtgcttcaaa gggcagaaaa atttatctc
15720
tatatgttca tagatattat ccagcttcc tgaagctcag tgctaggccc cttccttcag
15780
gaagacactcc ttgattgctt ctactctata gctctctc tcctgagcac ctccagtcct
15840
gaccgcctga gccccacact ccagcccttg ccccatgacc agcctggagc tgtctaggtg
15900
agtgagtctg gtcactctga tcaatgtggg ggctccctga ggacagggcc ttggaatact
15960
tgtttcccgaa atatgatatac tcatggtggc actgatcatg gggtgggctt gcaggtggga
16020
gggggtcagg atagaaatgc tgcaaatcag agaccttcc tctccctcac accaatgccc
16080
atggggtccc aagttccatg gattctgtct ctccttttc ttttccaga gtcaactccc
16140
ctcctgcaca tccctccatc tgctctctag tcttcttgcc tgggctgcta caacagcctc
16200
ctctcctctt gcctccctgc ttctttctt gccaccttga ccatgttgat ttctgcttaa
16260
agccatcagt ggctctttat tggctcaag aataaagtcc aatttcttag catgatagtc
16320
aaggcccttg acaccaggt cccagcctaa ctgtcctgac ccatctccag cattctacat
16380
ggccactggc attggcacac atagggttgg cacataccct cttgtgtgg atccaccact
16440
taaaggcacct ctccttctt acccactgct cacggatgaa ctcctaccca tcttaaccc
16500
cacactcaaa tgccgcctcc tgcattggagc catcccgac acccttaggt ctgaagcagt
16560
cctttctttt tttccctcctc tatcccttcc cttctcttgc aacgtagttt ttccacttta
16620
ttgcccttga accacaataa cacatacatt ttatgttata atccagtaca cacacacaca
16680
cacaacaca gaatccagct gtactatccc ctttccagc ctattctaat gtcttctatt
16740
catttcacac acactgatca tgacccacta aattgactcc acagtcact cttgtgtcac
16800
aatccatagt ttgaaaaata caaattctgt ttttagtgcat ttgccataat tcactatgaa
16860
cttcatctct ttggacctaa tccttctttt ctttgcact ggacttgagc tcctgggga
16920
tagacaagta agtagaagcc atattggagc caccatatct ccctcaggac agagccattg
16980
agggaaatgtc ggctgaacag aattgactca gacctgctga cccctggaa agcaggtgg
17040
atgcagaagc gggagggac ttccctctct ggcagccag ccctgcctgc agatgagctt
17100
ctggttacag acactggta tcaaaggact ggaggataga tggccccac tcttcagggg
17160
agctgtctgc tggccaaat gagggactg ggcctcagg cacagcctcg gacaggaggg
17220
agtaagacag aaagaatctc atccaccccg tggaaacgt agacggatgg gcacacactc
17280
tgaggggctga tggcaaggct agaattgtgg ggctggatgc agagaggtgg gtgccattaa
17340
ccatgcaac acaggtgccc ctgctgagga tcagtgtga tggatgtggag tggggcttc
17400

caggggcatt agggtctgaa tcatgcaggc agggtcttct cttgcaggac tctgtgtct
17460
gggaggcctg gttccatgc ccagagtggc ccaggcctgc caaaaacccc cattacaatg
17520
ggctctcccc gttcagatg ctgacacttt gcaaagagtt cctggttcag gcagttacct
17580
ggggtggtcc tgaagcctgg ctgggacatg agaccctata ttctagctgt agtttgc
17640
ctagatagct aggtgaccat gactaagccc cttcccttctt ggtctcagtt ttgccttta
17700
gaaatattac agtaaattga ttctctatga ttttattggc atctcaggtg gatcaattat
17760
atgttatgtg agactccccca gtctactaaa tgccagttag caccctcaag ctattgtgac
17820
aactccaagc atcctgacat caccatcaaa tgggtggcac ttctatctct gataccttct
17880
ggaaagacat ggaccatagg agacagggac catgagggac tacttttggt ggtagtttg
17940
ggggaggag tacaagggac aggggttaggc atggggaaaaa gagttaatgg gactgtggta
18000
aggcacaaca atggggtcat tggctatttgc caaagaaaag gacaggaagt gagggagagg
18060
ctggatatgg tggctcatgc ctgtatccc aacatgggagggc gagaccaagg caggtgaatt
18120
gtttgagccc aagagtttga aactagcctg agcaacatga taaaacccca tctctcaaaa
18180
aaaaaaaaaaa aaaaaaaaaat agccaggcgt ggtgggtgc gcctgttagtc tcagctactc
18240
agaaggctga agtgagagaa ttgcttgagc ctgggaggca gaggttggc tgagctgaga
18300
tttgtccact gcactccagc ctgggtgaca gagtgagacc ctgtgtatgat gatgtatgatg
18360
atgtatgatga tgatgtatgat gataatgatg atgtatgatg aagaagaagg aggatgaaga
18420
ggaggaggag gaggagaaga agaagaggtt gtaacttagtag tagtagttggt ggtgagggag
18480
agaaggtgat gacattgagg tggggagggc cagaagcaat atttacagaa ggaatgaagt
18540
: :
cactacatgg gatcgaaact ttcagcacccg cggacagcac aactccattc ttcatccctg
18600
cctatctcta aggtggggc cctcgagag ctcctaagaa ccagccttcc cacccgatata
18660
tccgaccccttgc gcatgggccc cagccatttgc aggtcaagct aattcatgcc cctttcaag
18720
ccccagctcc agccccagct cctccaggtt gttccatga ctgttccagc tcccaaggg
18780
tccagctgct tgtgactgccc tggctcatgc taggctggc ttgtatcttc actgaccct
18840
ataaaagggtgt tgtaactttcc taaccagccc aggttctctg gagcagaagt ttaccttatt
18900
ttgtacaagg ctgagagccct ctaaaaggcat gtcctgggtt gctgtcatct ctggccttt
18960
cttaagcact gcacagagct gagcacacag gatctcaggta tggaccaag ctgtatgt
19020
ttccttaggtt acttgggaac tcattgaagg taggaggctc agagtggcc aggaggacag
19080
tctgccacct tgtatgccc tgcttcacca agatgcatttgg caaatacaaa acagtcacac
19140
atacatgcag gaggaaagcaa ggtctcacac aaagacacca tgggttaggtt gaccctg
19200

ccaaaactgc agacattgca gtccaacccc acatggggga ggggtgtcag tctcaaacag
 19260
 caaacctgtg ggcagcattc tccaggttcc tggcatgaaa gctttgactg ttccaaggaa
 19320
 agcaagttgg aaactgagtt atcaggcattc tccttggaa ttaggaagga agacaacttc
 19380
 tgcatttggc ctggggac cagaagagag aaactgacac atctgggtc actcaaccat
 19440
 gaagagggcag aagatctctc tcccttggaa cacgcttcct cacgggagac cctggagttg
 19500
 gtgtacctag gagagacaga ctctccctgt gaccctgtgt ttcagcagat gcagccagtt
 19560
 atgccacctc ttcagtgtca ttaccacttg gtgtccagat cctcagaaga gaataaccta
 19620
 ggggccagat atcccaccct cagttccctt tactgcctgt agttggcca ctccagatcc
 19680
 agccatcttc ttgtctggcc tggggcttag ttgagaatac tttgaccatg actttgagtc
 19740
 acctttctta cctactttt ttttttttga gatgggtct cactctgtca cccaggttgg
 19800
 ggggcagtgg tgtgatctta gctcactgca gccttgaact cctatactca agtgatcctc
 19860
 ctgcctcagc ctcctgagta atggactac aggcatgtgc caccatgcca tgcattttg
 19920
 ttatTTTTTTT gtttgggggg tagaaacggg gtctcactat gtgggggggg ctggctcaa
 19980
 actcctggcc tcaagtgatc ctcctgcctc agcctccaa aatgttggga ttgcaggcat
 20040
 gagccaccat gcctggccat ctttcttgct ttctgtggaa aaaacctctc aatcaatgtc
 20100
 tcctccca cttggcctcca cttcctggag tggattgctt ttcatccccaa gttaagagac
 20160
 ttttggggaa ggtggacaga gctgggggtgg tcagacgtac agtcctagta ctaacagcat
 20220
 gttttccctg agcttccaa acgagggtgca ttccttgag tcagcatcct gcattgctt
 20280
 gaggccactg caggtaagg aggactgacc aactttgagg cccacctggc atatttctgt
 20340
 ctgtatccga tcacagcaaa ttgggtgtca gagggaaaca ggcagctgg gatgttcc
 20400
 caagctcaga cccagactga cgtagaagtg ctagattgag tttgggagtc taatattgcc
 20460
 aaccccagcc actgggtggg aaaggcatcc ccaggccct ttgctgggtgc tgcattggat
 20520
 tcaggccagg gaacttggat tccatttctt agtgttatat ggacacagct ctctgagcct
 20580
 cgactgcctg ttctggacct tcagagaccc tcaactccct attccacag gactttcac
 20640
 tcatgggtgcc ctgccccctt taagcaggac ctgttaagcac acttacagga gacagtgggg
 20700
 aaactgagtc atagaggcac tgagcagttt aatggagacc atactgggtc tatgaaaaag
 20760
 ggaagggggaa gtggaaacca gctcactgtg taagctctgc cctcccccaga gagggcagtt
 20820
 gggaggaggc ccctacccca tgctcccgcc tctactgtac agggctccct cttccagcc
 20880
 atccttagcaa cactcagtgg gctcgctgcc atgacgcccag cctgtggagg aaagtgggg
 20940
 ggggaccaac acaggacccc tggcagaa gctgccttgg aactgagaaa catcaactaga
 21000

actcatcaag ccctccaccc acctggtgca gatgaactga ggtctgaaga ggggagacca
21060
cctgccccaaa gggagaaaag cagtcagtag gatggccggg attagatctg gctctcagtt
21120
cctagttcct atgaagtaat gcagggagaa gacagctggc tggcaggatg ccagcagcat
21180
ccctccaggg gggcaagggg ctgcctttct ctacaggctt ttaggtacca gaccttctca
21240
atctagatag acagaatcct ccctcccagg acatccccag aagccacaga gttctgggg
21300
ctctcagaga tagcaggaga ccaccacccc agaatgagga tagccattct tggtgtgagc
21360
aggatttccc ctacccaagg acatgatggc ccctccttcc aggccccagg ccaccccaa
21420
ctccccctccc cttgctgaca atgccttagc tgtctacagg gagccccaaag cagcatcatc
21480
tccccctgtgt gccatggccc cacgaggtca gcatgttctc tgtcccttc acacagataa
21540
aaaaactggg acttggacaa ggagggcctg ccagtcctc agtgagtcat ggcaaacc
21600
ggacttagat ccagccctgc taaatctgag cccaggttcc tcccactctc cttgcccc
21660
gctgctctcc tggcaggtgc tgtgtgtgaa agggaccgccc tgcctgactc tgaagcac
21720
ggtgagggtg ggcagtcaga ggggccccaaa tgcctgtacc tggggcccaag ccaagaagcc
21780
ctgtggggag ctccctgagg atcactgaga tggggctct cttcagccc gtctcagg
21840
ctccaggctc tgctgtggca ctggtggtaa ggagtgcaca gggaggatg ctgggac
21900
tgacttaagg agcaggtggg aggagaggaa agggccaagg cccaggtccc cagccagccc
21960
ttgattgaga ttttagatggc acatttgaa aagcagtatc cttccagagt attctgg
22020
tgtgccatag ggctacggac aagcagccgc tgcctctaaa gccagcagaa tcgaggccc
22080
tgccctggc caacatttga gcctccatg atctggctta ttctccctc cctcccttc
22140
cactccgt cccaccctcc ctccatctg acacatcatt agctcaattt ttcaagg
22200
tgttcaacaa cacttcctcc atgaagtctt ccaaatttac tttcccttac tcttgg
22260
actcccttcc atgtgctctg actgtcttca ctgtgttatt ttacttggga gcctat
22320
aaaagttctc tttgagagtc tgcctgtctg tcactcttcc tagaacagga gcccctgg
22380
ggcaggctca ggtcttatgc atcttgaaa agcttgcctc taggctcctc aattcttct
22440
gggggaaagg gtaaaatact cagaacccca ataagggtg agcctgagca agacgat
22500
gtggctggag gattcctggg gagagcagga gacaggaaag atcaagatgc atgcag
22560
gggttagaagc tagagcagaa gccaggagtt cccagagcca gcagaggcct atcagg
22620
agacttgctg tagaactctg agcagctgt tttcccttc ctggccagtc atttc
22680
cttaagtggg gagggaaagg ctggactcgg aacatagagc gctctgcagc cggcag
22740
tggggtgtct ggatggccac aagggcatcc acttctgctc tatttctt cttc
22800

tttcttagct aaacctctgt gatggccatg cctgtccac cttccctctc tcccagcagg
22860
gaagttgttc tcacacatgg agtaacttgt ggccttggaa gaatggaata gagtcagggg
22920
ggatcagggtc tcgctggagt ctgagaatgc agacctgagt ttccggattt acagcttcta
22980
cttcttcaac ccagagggca gggcttatct gggctctcc tgaggcttgc acccctgcac
23040
tgccgctgtc cttacaataat gtggcatccc aactgctcca agacctttaa agtttacccc
23100
cactccctcc agaaagcctc ccaggaatgt cccagtgtcc accaagcccc tcttcccgat
23160
ctctgactgt tgatttgcac aagcctcctt gataagcagc ctggggcttc ctgagggcag
23220
gtctcagcct ttcttatcac ctctgactct tagggctgaa gaagggattc ctgcataaagc
23280
aggcagacccc aatgggagag acccctctgg ctggagacca ctcagcttat gtgtttccat
23340
tgtaacacaaa tcagtgccta agcatgtctc tagaatgggg ttctggaaag tggggacccc
23400
taacctcccc atgtggctag gttagatggg atgccccttc ttcccctgtc ctggcagatg
23460
cctcagtaca gatgacccca gccattccca gttagacatt gggagcttg aggtcaagga
23520
ggctgaggct cagccaagct gcacttagtg gttccacaga agaaaaatgg accatggcca
23580
ggagggaggg ggcaagaccc tccttgca gcaaaaggac ctagggaca ggcaagatga
23640
ggtcaggag gagggcagat agaaggaggg atgggagtgg ggagacagtg gttttggaga
23700
tggagtgcggg ggtatgaagaa ggcagttaggg agagatggag aaagagagag agaggtagaa
23760
aatgagagaa tctgagaaag acagaaacat acacccagaa acagaaccac acatagagaa
23820
aatcagaaac agacagggag acaatgagag agacacagag acagagatgt acacacagag
23880
atgggcttag aggagtccgt gtggaatggg gagaggtgg agaaaaatgg aatataagtg
23940
ccccacttct ggccaaacca cttccatgtc aatccacttc cttcggcct acagacaggg
24000
agacaggccc acaaaaggaa tgagacttgc cccaaataaa ttgtatatgg acattnagga
24060
ttgtttctag ccaccagga tttgaacctg ggttcggaga atcctgggtt aagaccgagg
24120
ctacctcccc gcctagagct aaaatgccag atccttactt cccaggatcc cttgttagcca
24180
gagcttggc atgggatttg gggctccaca tattccccca cccatcagat gcacctaccc
24240
gagggattha gtgaactgga ggcccccaga tggagacagg gagaaagcct ctccagagat
24300
aactgcagga agcttaagac tccaggttgg ccaggtgcag tggctcacgc ctataactccc
24360
agcactttgg gagccagaga caggaggatc gcttgagtct gggagttcaa gaccagctct
24420
ggcaaaatgg caagatccca actctacaaa aaatttaaaa attagccagg tgtggtggca
24480
tgtgcctgca gtctcagcta ctcgggaggc tgaggcagga ggattgcttgc ggcacaggag
24540
tttgagattt cagtgtatctt tgatcatgcc actagtggtg gccaggcgtg gtgggtttgg
24600

cagggcctt cagggagagg gaggtcccct cacagggcca gctctgccct gtgttcttgg
24660
cttggggccc caaatctggt tctctagccc actcagtgat tccataagct ccccaataatc
24720
tttttaaaaa atttcttctc tgcttaacct agcagagttg ctttttggaaa ggcagcagaa
24780
cctggggtttg aatccttggt ctgttacaag tgacttcatt gctccacacc tcagttccc
24840
catgtgtaaa atgaggataa tgccatgtct ctgtcactcg atggtgcaag gattaaatga
24900
gttaaaccac agtacaaaca tgtggaagct cagccactga agcgccagca caggttgtgt
24960
agagaacacc caaggagact cgtgtgctta acttggctct gccactgact aacatgtgtg
25020
gccatgcgct agtcccttcc cttcccttgg ccctgctgca tctggaaata actctggta
25080
agatggctca aggctctgac ccagcctccc aaactcacac actgttagcta tttgctaacc
25140
ccacatcctg aggacttcta aacgccttat ctccactctt ggtgctctct tgagctttc
25200
ccccacccaa ccagctgctt cctgaacatc tccactcagc tgcctccca gctcctcaaa
25260
gtcaacacat ccccaactgt gtcagctgt ttccctgtgt cccaacctt accttgcac
25320
ctctcaccat atcttgtcaa tttcaaatgc ccaagggtgc tggcgtcttc aacctcattt
25380
gctcttacag cctgggtct acctgggtct gtgcagaaat ccccttactt ctccttccc
25440
tccagaggac agtgtgctcc caaaatttag cttctctgct cacaagctca cccaggctcc
25500
ctctccagat ggtgaagccc tgttagaatgc cctgcacac ctgacccctt tgccctccc
25560
cagcctcatt gtttcttctt cctactccct gttctccata caacacactc ctcccaaggc
25620
acaaagccca gcctgtctcc ttgccttgc tcagtcctt cagccttagag ggcctctcct
25680
gtgccttatta cactcacctg ctaaaatcca acctgtcccc agaagtccag ctcagaggc
25740
taagtcctcc ctgatcctcc aggccggagg tatgagtctt cttccaaac ctctaagctc
25800
tcctcgcaaa ccgaatgcct cttgtggagg gagtactgccc ccatggtaa tagcgaggc
25860
tgtggattca cctgcctgca ctggtgacata ggagctgatt aggactttca ataagttact
25920
tcatgtgtct gagactcagt gttcttgcct gcaatatggg cataaaagca gtatgtatct
25980
cagagggagt gtggcgagt gggattatgg atgcctgaga tatggataaca aagctctc
26040
agtggtagct ggcacccgaa aaatgtcaa cacttagctt tgtggcagat tctctgtgct
26100
cagctgagtt gaaaaatcgc agagactaat atctaaactg ctatccccac cggggcgatt
26160
cctgctctct aaggaatgag gcttcaatgc gggtttggct atagcataac aaaattgggg
26220
caggaagtgg agcctgaacg ctccctgttct tccccttaggc ttctcggttggaa ggtttctccc
26280
tctgtcattc tttttaaagg aaggatccca agggaaaaggc aaacagagaa gcaggaagca
26340
gcactgatgt gaagaaaggg gaggagggaa atcaatcatg catttccaca gccatttagca
26400

gcctctgctc ttcccaccta ggctgacctc acaccccagg tgctggtag tggtaagtgc
 26460
 tccccacccct ccacaagctc ctctcattct ccagcagttg cagtgcagggc agcactttgg
 26520
 tctaaatgaa gaaaattgtt cattagatac caggggctaa ttggccctac ttttaccagc
 26580
 ctgggtggtg ccccaacctt ttatcagctt ctcattagac ttaattgata tgaatggcct
 26640
 gatatgtgct gtctgcagca ctttggaa gtcttccttg tcaccccgcc cccacacccc
 26700
 aatcccatag accttgcgt tccccctaa ttagggcatt catcattctg ttcttttgt
 26760
 ctgttgaat tgtgagttag agcctaata cagtgcctgt gcagtggctg gcacatagtg
 26820
 gttattccac aaattgactg agtattactg tactcagctt ggctatgaag aagagacca
 26880
 ggccctggga agggggggcc cccaggctgg tcggggagaa caaggcatag cccaggctgg
 26940
 ggtaagaacg aaagcagagc tccctgaggt tcacaggctg tggctgaga tgctccatca
 27000
 ccccaactta ccctggatgg cccttaatct ccggggccag tctgaatcag gcatgctgat
 27060
 ctccattgcc ccacccaca cagctcacgg aatcctaaaa atacagttt ataaggctcag
 27120
 gactaggaag gcacaggaa atcattcctt ctgcttgcac acctccttagc ttggggatct
 27180
 tacttccttc tagggagatc tgtttaatc atggatggaa agcatgtgct ccctgggccc
 27240
 accagagccc tcagctggtg cagccaggag agtgcggctt ggcattcaggt gcaggtacat
 27300
 ctctcttgc ttcttaaccat ctctcatggc aaagccctca gcctcagctt ctgaaagctt
 27360
 gtgtgtgtgt gtgtgtgtgt gtgtgtatgt gtgtgtggagg tggagctgg
 27420
 ctaccaatca gacacagaca ctctgaggat gtttcagga tgagaggctt cagacaaaaca
 27480
 atccagatag ggtctcagcg tgcttgacaa ctggggtccg gctggagata tcaggggggt
 27540
 tctaggtgg gcagtgattt ctgcccccac agcctccgc agaccacaca cctggctccc
 27600
 atcccatgtt ccacagggca cagacaggca cctaccacca cgcacacaccc cagaacccac
 27660
 actgaacgac cccacactgga ggctgtgtaa ctccagcaac ctctactgcc ctttcccc
 27720
 gagcctcaag ccagtcctatg atttgccagg aggaaggcag gcacaggttt ttctccctga
 27780
 gagatgggtgt tggggtgctg attcagacac cactcattat gggggcgcag gaataatgct
 27840
 gcccacagcc cctgactcag cccagcaggt cagtcctggg acctcggatt agggaaagatg
 27900
 caacaccccg caccctctgg cttagcagaa cctacgcctg catgtcctcc ccacgaaaca
 27960
 tgcgctcgcc ccagccctcc tgcctcctac cctcctcctc ccctctggta gtagggctg
 28020
 gtcactgcac tcactctggg ggatttggga ctttagcatat ccgcctcctt gggataggtc
 28080
 cccctcctgg gcaggggaag cattgtctaa acaatgaagc cctaatgggg aacttcaggc
 28140
 aacgggaacc tacaatgagg gaaggaggtt tacactccaa gaggaggaag tatttccttc
 28200

tcccctttt gggctgtcac cgcatggagc acgggaattt tgggcaaattc actggcttac
 28260
 ttagattaga cgctgtattt ggggtagaag cagaaggcgc gcccccaaca ctgttaaggg
 28320
 tccttaacgc ccgcggtccg cggtaccgcg gccgtcagca gctttgtacc gctgggtccg
 28380
 gaatccttt tactccctaa accgccttct caggaggccg gtgttagagca ggcagcgagg
 28440
 gccggggca ctcacgggtc gtgttgtact tgacgccgtt gaagtactcg gggcacggcc
 28500
 tctccacgag ggctccggca ggcgtcggg gccagcacgt tccgatctgg tccaaggtcg
 28560
 tgttgcagta ggagtaggga cctggcggcg ggagagagcg cagtagggct cagaggggcc
 28620
 cgcagggacg cggggctctc ggagcgcggg gtcagggcg cacccagcgc gcgagagaag
 28680
 gagcccgccg acgcctccgac cgctcgccctc ccgcctaccc tcgggttcca ggggtggccc
 28740
 ccagccgtcc aagagcagct cttcagccag cgccaggctg cagttggctt ccagcaggct
 28800
 gtggagcagt gccgcgttcca tcgcgtcccg cagccgcgtg cggagagggga gtgggagtgc
 28860
 ggcggccggcg tgactgcgag ggagtggacg cgagagttag cggccgagag ggcgcggggt
 28920
 cctggccccc gccagcccg ccccgatctc ccggcagcc tttggcgc acctccggc
 28980
 gcccagagct gtcaagtggg gacccccc gagaggagcc gccgagtgcg cggagctgcg
 29040
 ggtacagccg ctccggcg gccaatggct gcgcgggggg gcggggccgg gcggctccctc
 29100
 tcggagggggc tcaagtcttcca gccccgggc cttccaccct gcccggatgtt cggcttctca
 29160
 gttcgcagct ctcttccact cgcggcgtcc agaggaggc ggtggctgg agagcgtgg
 29220
 cctgggtga cggaaatgctc tgcgggggca tcgcaggccc ccggactgcg gggggcagca
 29280
 aagcgcgcac acctgcccgc ctgcccggg agaacaccgc agctctgcta atcgagggac
 29340
 agccgcaacc caaagttcg cagttcccg cctgagcttc ccctccctga gcgaggacac
 29400
 tggagggagg cagagacgga gagccatgg gcagatctgc atccctccaaa agcctccata
 29460
 gcctctggga aggaaatgtt tcgaggaggc ctggagggaa gcagggagac acccatgagt
 29520
 ctaatgagat caagtggctt gagtgggttg aagtccctgg atcagaccgt ggaagctgga
 29580
 atggagagag atttccggagg ggcagccaga cctgggtgagg aggaacttggg gtgggtggct
 29640
 ggtttaagt aggagcccg aggccaggat aacatcaggt ctcctgctt ggtactgg
 29700
 tggccattta atcagatggc gtaagtctt gggagaaatt ggctgaaatt atctggagg
 29760
 ctttatacac ccagttgtta gtgccttggg caccagaaag aataaaggaa gaagtggagaa
 29820
 tgggtctctc ttcccaagg cttggcagg gcttcacctg accctggttt ccacaccgt
 29880
 agcggggcag gggcgggag aatcaggccc cacagcaggt gtggaggagc tggaaacctc
 29940
 cccagagaac ttgcccaggc ccccccaccc accctggccc agaatgcctc cttgggtgatc
 30000

tgctgtggct gcaccctcag agtgcagg gcgggcctgc tggctgc agctgagtgg
 30060
 ggcattgggc tcggaaagca gggcctgggg taaggcatgc ttccactcag gctgcatttt
 30120
 ggtccagcct ctgactctgc tctccctgcc atggtacccc aggcaggatg cttgcctct
 30180
 cttctggagc tttctccatc agtgaatct tccccgaggc ccctctcagc tctgagactc
 30240
 tctggttctg agttatgaga cggagagtct gggaaagaca tacatgtgtg ttgtgtgtt
 30300
 ttgtttcta tcttagagg gcaaggagag ctggAACtta gtctcagaaa ccagtccgt
 30360
 tcccctgcca tcctccatc aaacctaagc tgcttaggaaa ggctgaaacc accatgaact
 30420
 agcaccatgc cctggcaga gagaggcaaa gcagcaggca ggctgctttt gtgtgtccg
 30480
 ctctcaccag cctccatatt aatggtgctg tcactgccc ggcagagcga gtgagaccac
 30540
 actaagacca gggctgagcc ctggaatctc tccagggct gctgactggc aggaacaaga
 30600
 tgctgagcag ccagccaggt ctaccctctg ctcctgaat ggacagttag agcccaggct
 30660
 cagctctggg cagctgcaga gtagaggtt ctttgttcat cattatcc agtgccagga
 30720
 gcagggggagg acctaggcag agtcctaccc tccgctccta gactgagccc tgtaacccag
 30780
 gttcccacac aggccagagg gtggctggcc aggtcagccc atgtattcac agggcaatag
 30840
 tggccctcat atacaaggga tagccctca cacagcacgg tgcagcacag aggaagggag
 30900
 aaatcctaag ccaagctaa gttattatcc ttgttcatta ctgcagccac cacctagg
 30960
 gtgcctaaag cacctataga tgcagctatg tcaagaggtg gtgtgtccc aatgaccaga
 31020
 gcccaggac ttctcatctc atgcagattt ctgcagaaca gaggggtggc ctgggtgaac
 31080
 tggattgctc ttaactggga gagctcacat accaaagatt cttctggaa gtgaccattt
 31140
 cagtggcaga gtcaaaggct gttctgcattc ctgaatgagc agttgggtc tgagcacata
 31200
 cccacagacc cacagacccg aggtcccctg gatgtggggc catttctca tggatctt
 31260
 tattataggc acagtgtca ttcgagatgt gacagaggaa aaactagaaa aggttagcagt
 31320
 ttggaaacaa attgattta cagctcaatt tagtgtctgc aaacaggcaa atgaggaaga
 31380
 aattggaaag agcccaaaa ttccctcaat ttactaaat ccaagtacaa acaaacaag
 31440
 acagggggcat tttgctaac taaagaagca gagtagatt agaggcttg ggatagtgt
 31500
 gggttccctcag cgctgcaggc cagccccatc ccagctggag gccaggaatt agggataag
 31560
 tatttggaaac agttgtgtg tcccaagct gttggggaca gttggcaaa tagttcagg
 31620
 gaagtgtgac aggtactttg gaagacccct gtgtaccatg agcagcagag taaggcaggt
 31680
 gcctctgggt ccctgcattcc catccatgg cccacccctg atgtcagctg cagcagcagt
 31740
 ggacagttcc aggccagctc agacacactt gcaaagccta atacatggca tcaaacaggc
 31800

atgttaggcac aacagaatca cacacagagc cacagttatg catcttcacc atgcacacac
 31860
 attctctctc taattcatcc cctaccactc aaatgtcagc tggaaaacag gatttaaagg
 31920
 gacaggatgc atcttgctt ctctaaggga ctgtccttgg ttagactaca tagagaggga
 31980
 gtctattcag gcacagctgg aatagtcgtg gtgttcctg ggaattacga cggggttgct
 32040
 gagggcacag aggttccaga ggggcctgaa ggtgctgcac ttttgcctg tagcaactgga
 32100
 gacaggggtc ccaggccagg cacccttcc tcaccctatc ctactccact gcaggacaga
 32160
 ggaattggcc tggtgtcttc atgtcaaacc atgggactaa gctgtggggc cagaaaggac
 32220
 tctcagtgac atccaatccc acattcttgc tttacagatg gacaatagga ttcccagaca
 32280
 agcaaaagag gtttccctcc cttaacttctc ccaggccatc ctcatctctg cagcagccca
 32340
 aatgggctcc ctgttccttc atggcagata aaacgatgtt ttataattac gatcctgtca
 32400
 ctcttcctca taggagtgcc gccaacccat tgccctttag atggagacca cctcttggt
 32460
 attgcaaaga aggtccttcg taatcatcta tccctcatcc tcacccttct gctccatttg
 32520
 aaaggccctc cagaacaaca tgcatttcct gggaaaccatc tttggatttg caaagaaggt
 32580
 tcttcatgat catcgcccc catcctcacc ctcctgcttc atttggaaagg ccctccagaa
 32640
 caatgtgtca ctaccagaaa ccattcgcta tctctgttcc ttgtttgtat ttctgcactg
 32700
 gaaatggcct tacagctcca cactcatctc tcagacatgc agtccttgg tgacccatc
 32760
 tgctgggtgc tggcttatgt cacatccagg ttgctttta ttcatgtct catctaccag
 32820
 cccatgcctc cccaaaggca gggctggc tggttacta acccatcccc agcacccatt
 32880
 ggaatgctct agacttcagc ccagttagtc actaagaaag gacagataacc ttggcagaag
 32940
 agctgaggaa agcccaggc cctgtcttca ggccaaggct ctcttccat
 32990

<210> 14
 <211> 644
 <212> DNA
 <213> mus musculus

<220>
 <221> CDS
 <222> (213)..(521)

<400> 14
 gcgccccctc agctccgcga gccccgcggc ttcttgc aaggtcctgg ggtgatcgat
 60
 caattgcgga gccccgaagc tgcccgactg gccggggtgg gcggggagga gcctggacgc
 120
 tgcactctct ggctgctcct cgtcgcgcgg gctccctcgc agccacgcgg ggcgcgcact
 180
 cccactccct ctgcgcgcgg ctccggggcg ca atg gac gcg gcg ctg ctc ctc
 233

Met Asp Ala Ala Leu Leu Leu
 1 5

agc ctg ctg gag gcc caa ctg cag cct ggc gct ggc cga aga gct gct
 281
 Ser Leu Leu Glu Ala Gln Leu Gln Pro Gly Ala Gly Arg Arg Ala Ala
 10 15 20
 cct gga cgg ctg ggg agt gcc ccc gga ccc cga agg tcc cta cac cta
 329
 Pro Gly Arg Leu Gly Ser Ala Pro Gly Pro Arg Arg Ser Leu His Leu
 25 30 35
 ctg caa cac gac ctt gga cca gat cgg gac ctg ctg gcc aca gag cgc
 377
 Leu Gln His Asp Leu Gly Pro Asp Arg Asp Leu Leu Ala Thr Glu Arg
 40 45 50 55
 acc cgg agc cct agt aga gag acc gtg ccc cga gta ctt caa tgg cat
 425
 Thr Arg Ser Pro Ser Arg Glu Thr Val Pro Arg Val Leu Gln Trp His
 60 65 70
 caa gta caa cac gac ccg gaa tgc cta cag aga gtg cct gga gaa cgg
 473
 Gln Val Gln His Asp Pro Glu Cys Leu Gln Arg Val Pro Gly Glu Arg
 75 80 85
 gac ctg ggc ctc aag ggt caa cta ctc aca ctg cga acc cat ttt gga
 521
 Asp Leu Gly Leu Lys Gly Gln Leu Leu Thr Leu Arg Thr His Phe Gly
 90 95 100
 tgacaaggag tatccgctgc ctgaggaatg tgatccactg gaactcatca ccacccat
 581
 tctgagaaac atcgcgtgg tccctgctgca actcatcgac cacgaagtgc acgagggcaa
 641
 tga
 644

 <210> 15
 <211> 103
 <212> PRT
 <213> mus musculus

 <400> 15

 Met Asp Ala Ala Leu Leu Ser Leu Leu Glu Ala Gln Leu Gln Pro
 1 5 10 15
 Gly Ala Gly Arg Arg Ala Ala Pro Gly Arg Leu Gly Ser Ala Pro Gly
 20 25 30
 Pro Arg Arg Ser Leu His Leu Leu Gln His Asp Leu Gly Pro Asp Arg
 35 40 45
 Asp Leu Leu Ala Thr Glu Arg Thr Arg Ser Pro Ser Arg Glu Thr Val
 50 55 60
 Pro Arg Val Leu Gln Trp His Gln Val Gln His Asp Pro Glu Cys Leu
 65 70 75 80
 Gln Arg Val Pro Gly Glu Arg Asp Leu Gly Leu Lys Gly Gln Leu Leu
 85 90 95
 Thr Leu Arg Thr His Phe Gly
 100

 <210> 16
 <211> 20
 <212> DNA
 <213> synthetic

 <400> 16
 ccccgaaagct gccccactgg
 20

<210> 17
<211> 24
<212> DNA
<213> synthetic

<400> 17
ggaaggctgt aaaggatgga gaag
24