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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 November 2001 (29.11.2001)

PCT

(10) International Publication Number
WO 01/90141 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/CA01/00768

(22) International Filing Date: 24 May 2001 (24.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/207,031 25 May 2000 (25.05.2000) US

(71) Applicant *(for all designated States except US)*: **MERCK FROSST CANADA & CO.** [CA/CA]; 16711 Trans-Canada Highway, Kirkland, Québec H9H 3L1 (CA).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **NG, Gordon, Yiu, Kon** [CA/CA]; 16711 Trans-Canada Highway, Kirkland, Québec H9H 3L1 (CA).

(74) Agents: **SWABEY OGILVY RENAULT** et al.; Suite 1600, 1981 McGill College Avenue, Montreal, Québec H3A 2Y3 (CA).

(81) Designated States *(national)*: CA, JP, US.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

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

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

(54) Title: NOVEL GB1C ISOFORM AND NUCLEOTIDES ENCODING SAME

MERCK hgb1C DNA AND PROTEIN SEQUENCE

```
1    atgttgctgctgctgctactggcgccactcttcctccgccccccgggcgcgggcggggcg
     tacaacgacgacgacgatgaccgcggtgagaaggaggcgggggggcccgcgcccgccccgc
     M  L  L  L  L  L  A  P  L  F  L  R  P  P  G  A  G  G  A

61   cagacccccaacgccacctcagaaggttgccagatcatacacccgccctgggaaggggggc
     gtctgggggttgcggtggagtcttccaacggtctagtatgtgggcgggaccccttcccccg
     Q  T  P  N  A  T  S  E  G  C  Q  I  I  H  P  P  W  E  G  G

121  atcaggtaccggggcctgactcgggaccaggtgaaggctatcaacttcctgccagtggac
     tagtccatggccccggactgagccctggtccacttccgatagttgaaggacggtcacctg
     I  R  Y  R  G  L  T  R  D  Q  V  K  A  I  N  F  L  P  V  D

181  tatgagattgagtatgtgtgccgggggggagcgcgaggtggtgggcccaaggtccgcaag
     atactctaactcatacacacgggccccctcgcgctccaccaccccgggttccaggcgttc
     Y  E  I  E  Y  V  C  R  G  E  R  E  V  V  G  P  K  V  R  K

241  tgcctggccaacggctcctggacagatatggacacacccagccgctgtgtgaatcgaacg
     acggaccggttgccgaggacctgtctatacctgtgtgggtcggcgacacacttagcttgc
     C  L  A  N  G  S  W  T  D  M  D  T  P  S  R  C  V  N  R  T

301  ccacactcagaacggcgcgcagtgtacatcggggcactgtttcccatgagcgggggctgg
     ggtgtgagtcttgccgcgcgtcacatgtagccccgtgacaaagggtactcgcccccgacc
     P  H  S  E  R  R  A  V  Y  I  G  A  L  F  P  M  S  G  G  W

361  ccaggggggccaggcctgccagcccgcggtggagatggcgctggaggacgtgaatagccgc
     ggtcccccggtccggacggtcgggcgccacctctaccgcgacctcctgcacttatcggcg
     P  G  G  Q  A  C  Q  P  A  V  E  M  A  L  E  D  V  N  S  R

421  agggacatcctgccggactatgagctcaagctcatccaccacgacagcaagtgtgatcca
     tccctgtaggacggcctgatactcgagttcgagtaggtggtgctgtcgttcacactaggt
     R  D  I  L  P  D  Y  E  L  K  L  I  H  H  D  S  K  C  D  P

481  ggccaagccaccaagtacctatatgagctgctctacaacgaccctatcaagatcatcctt
     ccggttcggtggttcatggatatactcgacgagatgttgctgggatagttctagtaggaa
     G  Q  A  T  K  Y  L  Y  E  L  L  Y  N  D  P  I  K  I  I  L

541  atgcctggctgcagctctgtctccacgctggtggctgaggctgctaggatgtggaacctc
     tacggaccgacgtcgagacagaggtgcgaccaccgactccgacgatcctacaccttggag
     M  P  G  C  S  S  V  S  T  L  V  A  E  A  A  R  M  W  N  L
```

(57) Abstract: A new human isoform of the gamma-aminobutyric acid receptor gb1c, termed Gb1c, is described as well as nucleic acids encoding same and various assays employing said receptor. The novel Gb1c cDNA encodes a protein of 899 amino acids differering from the previously noted gb1c sequence (Genbank Accession No. AJ012187) by a single amino acid substitution (a glutamine in place of a histidine at position 21). This single amino acid substitution affords the receptor higher affinity for $GABA_B$-selective antagonists. Accordingly, the disclosed Gb1c receptor can be used to differentially screen for $GABA_B$ receptor modulators such as agonists or antagonists for the treatment of various CNS disorders, drug addiction, neuropathic pain and gastroesophageal reflux and incontinence.

WO 01/90141 A2

TITLE OF THE INVENTION

NOVEL GB1C ISOFORM AND NUCLEOTIDES ENCODING SAME

FIELD OF THE INVENTION

This invention relates to a novel receptor, designated Gb1c, to nucleotides encoding same, and to assays employing said receptor.

BACKGROUND OF THE INVENTION

Gamma aminobutyric acid (GABA), an endogenous neurotransmitter in the central and peripheral nervous systems, is responsible for mediating synaptic transmission through two ligand-gated channels, $GABA_A$ and $GABA_C$ receptors, and a third receptor sub-type, $GABA_B$, which acts through G proteins to regulate potassium and calcium channels. Specifically, $GABA_B$ receptor activation increases K^+ or decreases Ca^{++} conductance and also inhibits or potentiates stimulated adenylyl cyclase activity. The expression of $GABA_B$ receptors is widely distributed in the mammalian brain (*e.g.*, frontal cortex, cerebellar molecular layer, interpeduncular nucleus) and has been observed in many peripheral organs as well.

Two $GABA_B$ receptors, $GABA_BR1a$ and $GABA_BR1b$, have been identified that bind $GABA_B$ receptor antagonists in transfected cells; Kaupmann *et al.*, 1997 *Nature* 386:239-246. More recently, a third $GABA_B$ receptor isoform, gb1c, was reported to GenBank (Accession No. AJ012187). These receptors are extremely valuable in identifying modulators of the $GABA_B$ receptor, such as $GABA_B$ agonists and antagonists.

Pharmacological and physiological evidence indicates that a large number of $GABA_B$ receptors remain to be cloned. Accordingly, this field continues to be of intense interest. Modified or variant receptors can differ significantly in their functionality. Even a single amino acid change in G protein-coupled receptors can lead to large differences in the receptor's ability to bind ligand (agonist/antagonist) and to couple functionally to signal transduction pathways. For example, a single amino acid change in the ligand binding pocket formed by the 7 transmembrane segments of the rhodopsin receptor (Lys296Glu) leads to constitutive receptor activity and the pathology termed retinitis pigmentosa; Keen *et al.*, 1991 *Genomics* 11:199-205. Mutations in GPCRs and disease have been extensively reviewed by Coughlin,

S.R., 1994 *Current Biology* 6:191-197, and Spiegel, A.M., 1995 *Ann. Rev. Physiol.* 58:143-170.

It would be desirable to identify other $GABA_B$ receptors which could be used in the identification of modulators of the $GABA_B$ receptor, particularly compounds that could perhaps combat some of the significant pharmacological activities that have been found to be associated with $GABA_B$ receptor activation such as various CNS disorders, drug addiction, neuropathic pain and gastroesophogeal reflux and incontinence.

SUMMARY OF THE INVENTION

This invention relates to a novel γ-aminobutyric acid type B ($GABA_B$) receptor polypeptide designated Gb1c. In particular the present invention relates to human Gb1c, illustrated in SEQ ID NO:2 and, more specifically, to the sequence of amino acids 18-899 of SEQ ID NO:2. The particular $GABA_B$ receptor sequence taught herein comprises a glutamine at position 21 in place of a histidine. Accordingly, those $GABA_B$ receptor polypeptides of the gb1c subgroup comprising this amino acid substitution are expressly included within the present invention.

A further aspect of this invention are isolated polynucleotides which encode the Gb1c receptor polypeptides of the present invention. These nucleic acids may be free from associated nucleic acids, or they may be isolated or purified. The nucleic acids may be any type of nucleic acid, with preferred forms being DNA (genomic and cDNA) although this invention specifically includes RNAs as well. Nucleic acid constructs may also contain regions which control transcription and translation such as one or more promoter regions, termination regions, and if desired enhancer regions. The nucleic acids may be inserted into any known vector including plasmids, and used to transfect suitable host cells using techniques generally available to one of ordinary skill in the art.

Another aspect of this invention are vectors comprising the above polynucleotides, and host cells comprising the vectors. Still another aspect of this invention is a method of making Gb1c comprising introducing a vector comprising polynucleotides encoding Gb1c into a host cell under culturing conditions.

Yet another aspect of this invention are assays for γ-aminobutyric acid type B ($GABA_B$) receptor polypeptide modulators, e.g., agonists and antagonists (compounds which stimulate or which inhibit the function of $GABA_B$ receptors).

One type of assay in particular comprises determining whether binding occurs between a candidate compound and a receptor polypeptide comprising the extracellular region of the Gb1c polypeptide. This can be carried out by means of a label directly or indirectly associated with the candidate compound. Another type of assay constituting part of the instant invention comprises determining if a candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells or cell membranes bearing the polypeptide.

Assays which make up further aspects of this invention include binding assays such as competition for ^{125}I-GABA$_B$ receptor agonist binding, coupling assays including agonist-mediated inhibition of forskolin-stimulated adenylate cyclase in cells expressing the Gb1c receptor polypeptide, measurement of agonist-stimulated calcium release in cells expressing Gb1c receptors and promiscuous G proteins such as G15/G16/Gqi5/Gqo5 such as aequorin assays and calcium-based dye assays, stimulation of inward rectifying potassium channels (GIRK channels, measured by voltage changes) in cells expressing Gb1c receptors, and measurement of pH changes upon agonist stimulation of cells expressing Gb1c receptors as measured with a microphysiometer. In certain cellular environments, stimulation of Gb1c receptors may also lead to inhibition of calcium channels.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 shows the nucleic acid (SEQ ID NO:1) and deduced amino acid (SEQ ID NO:2) sequences of human Gb1c. The polypeptide sequence comprises a signal peptide (M^1-A^{17}), a single Sushi Repeat (T^{26}-R^{98}), two putative glycosylation sites (N^{24}, N^{84}), and exhibits weak amino acid identity (26% identity over 26 amino acid overlap) to selectins. The putative transmembrane domains, and their corresponding amino acid numbers, are as follows: TM1:Ser526-Phe550; TM2:Ser563-Gly589; TM3:Leu605-Trp625; TM4:Trp646-Ile665; TM5:Thr706-Ala726; TM6:His740-Thr761; and TM7:Gln767-Phe791.

FIGURE 2 shows the amino acid sequence of human gb1c (Genbank Accession No. AJ012187; SEQ ID NO:8) aligned over the Gb1c receptor polypeptide sequence of the instant invention.

FIGURE 3 shows the complete cDNA sequence of HG20 (SEQ ID NO:3).

FIGURE 4 shows the complete amino acid sequence of HG20 (SEQ ID NO:4).

FIGURE 5A shows the major schematic topology/structural domains of gb1a, gb1b and gb1c receptors, namely, the signal peptide, Sushi repeats, coiled-coil and PDZ domains. Figure 5B illustrates the amino acid alignment of the extracellular N-termini of human gb1a, gb1b, and Gb1c isoforms. The proposed signal peptide cleavage site of gb1 is marked with scissors, and arrows delimit the Sushi domains (SU). An arrow (Ψ) marks the start of the LIV-BP-like domain. Gaps (...) were introduced to maximize alignment. The novel Gb1c isoform differs from gb1a by an in-frame 62 amino acid deletion, and elimination of one protein-protein interacting Sushi Repeat (also known as short consensus repeat) leaving a single Sushi Repeat interacting module. The N-terminus common to the gb1 isoforms exhibits weak amino acid identity to the bacterial periplasmic binding protein LIV-BP and is predicted to form two lobes that bind ligand according to a Venus flytrap model for receptor activation. The entire N-terminal domain of gb1, in the absence of the other portions of the receptor, is sufficient to specify agonist and antagonist binding.

FIGURE 6 shows a summary of BMAX (estimated maximal receptor density) and Kd for gb1c and Gb1c from ^{3}H-CGP71872 saturation binding.

FIGURE 7 shows how GABA mediated a dose-dependent inhibition of forskolin-stimulated cAMP (GABA EC_{50} value of 21±8 nM). Basal cAMP levels were approximately 1 pmol cAMP/2 X 10^5 cells and forskolin-induced cAMP levels typically greater than 20-fold. The functional and stably expressed Gb1c-gb2 heterodimer subtype exhibited nanomolar potency for GABA similar to the reported affinity of GABA at neuronal $GABA_B$ receptors.

DETAILED DESCRIPTION OF THE INVENTION

As used throughout the specification and claims, the following definitions apply:

"Functional $GABA_B$ receptor" refers to the receptor formed following co-expression of Gb1c and a subunit known variously as $GABA_BR2$ (White *et al.*, 1998 *Nature* 396:679-682; and Jones *et al.*, 1998 *Nature* 396:674-679); GBR2 (Kuner *et al.*, 1999 *Science* 283:74-77); or gb2 (Ng *et al.*, 1999 *J. Biol. Chem.* 274;7607-7610), also known as HG20 (International Patent Application PCT/US99/02361, filed February 3, 1999). A functional $GABA_B$ receptor, further, displays at least one

functional response when exposed to $GABA_B$ receptor agonists such as GABA or (-) baclofen. Examples of functional responses are: pigment aggregation in *Xenopus* melanophores, modulation of cAMP levels, coupling to inwardly rectifying potassium channels, mediation of late inhibitory postsynaptic potentials in neurons, increases in potassium conductance, and decreases in calcium conductance.

"HG20 or a functional equivalent thereof" refers to HG20 (or gb2) and those polypeptides capable of interacting with a Gb1c polypeptide to form a heterodimer active as a functional $GABA_B$ receptor, e.g., $GABA_BR2$, and GBR2.

In accordance with the instant invention, Applicants have cloned a novel cDNA from adult human cerebellum mRNA which encodes an isoform of the γ-aminobutyric acid receptor "gb1c" previously published by GlaxoWellcome in Genbank, Accession No. AJ012187. The novel receptor, designated "Gb1c", similarly encodes a protein of 899 amino acids, but differs by a single amino acid substitution (a glutamine in place of a histidine at position 21). This difference is found in the extracellular N-terminus of the gb1c receptor which is entirely responsible for ligand binding. As a result, the novel isoform disclosed herein exhibits higher affinity for antagonist and is more similar to native neuronal $GABA_B$ receptors which exhibit, e.g., approximately 1 nM Kd for CGP71872.

The human polypeptide sequence of the Gb1c polypeptide of the instant invention is provided in Figure 1. As noted above, the particular $GABA_B$ receptor sequence taught herein comprises a glutamine at position 21 in place of a histidine. Accordingly, those $GABA_B$ receptor polypeptides of the gb1c subgroup comprising this amino acid substitution are expressly included within the present invention. Polypeptides of the gb1c subgroup are readily distinguished structurally from those of the 1a and 1b subgroups. Gb1c subgroup polypeptides differ from gb1a subgroup polypeptides by an in-frame 62 amino acid deletion, and elimination of one protein-protein interacting Sushi Repeat. Gb1b subgroup polypeptides lack both N-terminus Sushi Repeats.

Gb1c proteins contain various functional domains, including one or more domains which anchor the receptor in the cell membrane, and at least one ligand binding domain. As with many receptor proteins, it is possible to modify many of the amino acids, particularly those which are not found in the ligand binding domain, and still retain at least a percentage of the biological activity of the original receptor. Thus this invention specifically includes modified functionally equivalent Gb1cs which

have truncated, or mutated N-terminal portions. This invention also specifically includes modified functionally equivalent Gb1cs which contain modifications and/or deletions in other domains, which are not accompanied by a loss of functional activity.

Additionally, it is possible to modify other functional domains such as those that interact with second messenger effector systems, by altering binding specificity and/or selectivity. Such functionally equivalent mutant receptors are also within the scope of this invention.

Particular embodiments of the instant invention are those truncated forms of Gb1c which comprise the extracellular portion of the receptor (amino acids 18-525; 1-17 encodes the putative signal domain), but lack the intracellular signaling portion of the receptor, and to nucleic acids encoding these truncated forms. Such truncated receptors are useful in various binding assays. Receptor chimeras (i.e., fusion proteins) which contain modifications and/or deletions not accompanied by a loss of functional activity are also included within the instant invention.

Isolated polynucleotides encoding the above polypeptides also constitute an integral part of the instant invention.

A further aspect of the instant invention are vectors comprising the above polynucleotides. This includes a variety of expression vectors useful in effecting the expression of recombinant HG20 or Gb1c. Commercially available expression vectors which are suitable in this capacity include, but are not limited to, pMC1neo (Stratagene), pSG5 (Stratagene), pcDNAI and pcDNAIamp, pcDNA3, pcDNA3.1, pCR3.1 (Invitrogen), EBO-pSV2-neo (ATCC 37593), pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), and the PT7TS oocyte expression vector (or similar expression vectors containing the globin 5' UTR and the globin 3' UTR). The choice of vector will depend upon the cell type used, the level of expression desired, and the like.

Cells comprising the above vectors constitute a further aspect of the instant invention. Particularly in the assays of the instant invention, the present invention employs cells co-expressing HG20 and Gb1c, allowing for the formation of functional $GABA_B$ receptors and HG20/Gb1c heterodimers. Such cells are generally produced by transfecting cells that do not normally express functional $GABA_B$ receptors with either a single expression vector or independent vectors encoding

HG20 and Gb1c. The cells are then cultured under conditions such that heterodimers of HG20 and Gb1c are formed where the heterodimers constitute functional $GABA_B$ receptors. In this way, recombinant host cells expressing $GABA_B$ receptors are produced. This does not exclude the use of cells endogenously expressing a functional $GABA_B$ receptor, however.

Host cells may be prokaryotic or eukaryotic, including but not limited to, bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including, but not limited to, cell lines of human, bovine, porcine, monkey and rodent origin, and insect cells including but not limited to *Drosophila* and silkworm derived cell lines. Cells and cell lines which are suitable for recombinant expression of HG20 and Gb1c and which are commercially available, include but are not limited to, L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), HEK293 (ATCC CRL 1573), Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26), MRC-5 (ATCC CCL 171), *Xenopus* melanophores, and *Xenopus* oocytes. In the case of *Xenopus* oocytes, co-expression of HG20 and Gb1c is often effected by microinjecting RNA encoding HG20 and RNA encoding Gb1c into the oocytes rather than by transfecting the oocytes with expression vectors encoding HG20 and Gb1c. Microinjection of RNA into *Xenopus* oocytes in order to express proteins encoded by the RNA is well known in the art.

Another aspect of the instant invention are the assays afforded by disclosure of the novel $GABA_B$ receptor, Gb1c. These allow for the identification of modulators of the $GABA_B$ receptor, e.g., agonists and antagonists (i.e., compounds which stimulate or which inhibit the function of gamma-aminobutyric acid type B ($GABA_B$) receptors) potentially useful in the treatment and/or prevention of the various afflictions attributed to $GABA_B$ receptor function or lack thereof, including various CNS disorders, drug addiction, neuropathic pain and gastroesophogeal reflux and incontinence. Other types of modulators are inverse agonists and allosteric modulators of agonists.

One type of assay involves determining whether binding occurs between a candidate compound and a receptor polypeptide comprising the extracellular region (amino acids 18-525 of SEQ ID NO:2) of the Gb1c $GABA_B$ receptor polypeptide. This expressly includes binding to soluble proteins lacking

transmembrane and intracellular regions, cells or membrane preparations bearing a receptor polypeptide comprising the extracellular region, fusion proteins comprising the extracellular Gb1c region and fragments of the Gb1c polypeptide of SEQ ID NO:2. The use of membranes rather than cells is well known in the art. Moreover, the above assay as well as all the assays disclosed relating to binding of ligand can be run with Gb1c receptor polypeptide monomers. Determination and measurement of binding can be done by means of a label directly or indirectly associated with the candidate compound.

This assay as well as the other assays disclosed in the instant invention can further be carried out in the presence of a competitor, for instance, a $GABA_B$ receptor agonist.

Accordingly, the present invention includes a method for the identification of agonists or antagonists of the $GABA_B$ receptor which comprises exposing cells expressing HG20 or a functional equivalent thereof and a receptor polypeptide comprising the extracellular region of the Gb1c $GABA_B$ receptor polypeptide to a suspected agonist or antagonist in the presence and in the absence of a known $GABA_B$ receptor agonist. In the presence of a $GABA_B$ receptor agonist or antagonist, the amount of binding of the known agonist will be less than that seen upon contact of the receptor with the known agonist. In preferred embodiments, the cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide are recombinant cells that have been transfected with a vector or vectors comprising a polynucleotide(s) encoding one or both of the HG20 or the receptor polypeptide.

Optionally, control experiments can be run utilizing control cells that are the same as the cells above except that the control cells do not express functional $GABA_B$ receptors. The amount of binding of $GABA_B$ receptor agonist to the control cells should be significantly less than the amount of binding of agonist to cells expressing functional $GABA_B$ receptor. In this way, one can ensure that the binding of agonist in the above assays is actually due to binding of the agonist to $GABA_B$ receptors. Therefore, if the compound can block this binding, the compound is also likely to bind to $GABA_B$ receptors.

Another type of assay constituting part of the instant invention is for the identification of compounds which modulate the function of $GABA_B$ receptors which comprises determining if a compound upon contact with a receptor polypeptide comprising the extracellular region of the Gb1c receptor results in a signal generated

by activation or inhibition of the polypeptide. This can be done using detection systems appropriate to the cells or cell membranes bearing the polypeptide. A signal is an event indicating a functional response of the $GABA_B$ receptor. Note, as above, this assay expressly includes use of soluble proteins lacking transmembrane and intracellular regions, cells or membrane preparations bearing a receptor polypeptide comprising the extracellular region, fusion proteins comprising the extracellular Gb1c region and fragments of the Gb1c polypeptide of SEQ ID NO:2.

One skilled in the art is familiar with a variety of methods of determining and measuring the functional responses of G-protein coupled receptors such as the $GABA_B$ receptor; e.g., through the monitoring of changes in pigment distribution in melanophore cells, through the monitoring of changes in cAMP or calcium concentration or chemotaxis, through the monitoring of changes in membrane currents in *Xenopus* oocytes, through the monitoring of changes in calcium concentration measured using the aequorin assay, or through the monitoring of changes in inositol phosphate levels. Depending upon the cells in which heterodimers of HG20 and Gb1c are expressed, and thus the G-proteins with which the functional $GABA_B$ receptor is coupled, certain of such methods may be appropriate for measuring the functional responses of such functional $GABA_B$ receptors. It is well with the competence of one skilled in the art to select the appropriate method of measuring the functional response for a given experimental system. Note that while the functionality of Gb1c is attributed to heteromer formation, it remains possible that $GABA_B$ receptor monomers or homodimers are functional when in certain cellular environments.

Preferred embodiments of the present invention are wherein HG20 or a functional equivalent thereof is expressed along with the receptor polypeptide such that heterodimers of the receptor polypeptide and HG20 are allowed to form.

This assay as well as the assays above can further be carried out in the presence of a competitor, for instance, a $GABA_B$ receptor agonist.

Accordingly, the present invention includes assays by which $GABA_B$ receptor antagonists or inverse agonists may be identified by their ability to antagonize a functional response mediated by the $GABA_B$ receptor in cells. One such method comprises contacting cells expressing HG20 or a functional equivalent thereof and a receptor polypeptide comprising the extracellular region of the Gb1c receptor polypeptide with a suspected antagonist or inverse agonist and determining

whether a functional response follows. A $GABA_B$ receptor antagonist will antagonize $GABA_B$ receptor function whereas an inverse agonist will result in a function opposite of a $GABA_B$ receptor agonist (e.g., a decrease versus an increase in a specific function). A preferred embodiment of the present invention is the above assay wherein the cells are contacted with the suspected antagonist or inverse agonist in the presence of a known $GABA_B$ receptor agonist, and wherein the response is measured and compared to that of the known agonist. In this situation, the functional response of the known agonist would be less in the presence of a $GABA_B$ receptor antagonist or inverse agonist. Especially preferred embodiments comprise the further step of contacting a separate group of cells expressing the receptor polypeptide (the control group) with the suspected antagonist or inverse agonist; a lack of response in the control indicating the presence of an antagonist and a response in contradiction with (or in a direction opposite to) that of the known agonist indicating the presence of an inverse agonist.

In preferred embodiments, the cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide comprising the extracellular region of the Gb1c receptor polypeptide are recombinant cells that were transfected with a vector or vectors comprising a polynucleotide(s) encoding one or both of the HG20 or the receptor polypeptide.

The present invention further includes assays by which $GABA_B$ receptor agonists or allosteric modulators of agonists may be identified by their ability to promote a functional response mediated by the $GABA_B$ receptor in cells. One such method comprises contacting cells expressing HG20 or a functional equivalent thereof and a receptor polypeptide comprising the extracellular region of a Gb1c receptor polypeptide with a suspected agonist of Gb1c or an allosteric modulator of such and determining whether a functional response follows. Contact with a $GABA_B$ receptor agonist should stimulate $GABA_B$ receptor function. By contrast, contact with an allosteric modulator of an agonist should enhance stimulation brought on by an agonist. A preferred embodiment of the present invention is the above assay wherein the cells are contacted with the suspected agonist or allosteric modulator in the presence of a known $GABA_B$ receptor agonist, and wherein the response is measured and compared to that of the known agonist, the functional response being greater in the presence of a $GABA_B$ receptor agonist or allosteric modulator. Especially preferred embodiments comprise the further step of contacting a separate

group of cells expressing the receptor polypeptide (the control group) with the suspected agonist or allosteric modulator; a lack of response in the control indicating the presence of an allosteric modulator and a functional response indicating the presence of an agonist.

In preferred embodiments, the cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide are recombinant cells that were transfected with a vector or vectors comprising a polynucleotide(s) encoding one or both of the HG20 or the receptor polypeptide.

In particular embodiments, the functional response is selected from the group consisting of: modulation of the activity of an ion channel; changes in calcium concentration; changes in a signal from a reporter gene whose expression is controlled by a promoter that is induced by interaction of an agonist with the $GABA_B$ receptor; and changes in membrane currents. In particular embodiments, the change in membrane current is measured in *Xenopus* oocytes. In other embodiments, the change in membrane current is the modulation of an inwardly rectifying potassium current.

Particular types of functional assays that can be used to identify agonists or antagonists of $GABA_B$ receptors include transcription-based assays. Transcription-based assays involve the use of a reporter gene whose transcription is driven by an inducible promoter whose activity is regulated by a particular intracellular event such as, *e.g.*, changes in intracellular calcium levels that are caused by the interaction of a receptor with a ligand. Transcription-based assays are reviewed in Rutter *et al.*, 1998, *Chemistry & Biology* 5:R285-R290. The transcription-based assays of the present invention rely on the expression of reporter genes whose transcription is activated or repressed as a result of intracellular events that are caused by the interaction of an agonist such as baclofen with a heterodimer of HG20 and Gb1c where the heterodimer forms a functional $GABA_B$ receptor. In the above assays, compounds capable of effecting a greater signal from the reporter gene than $GABA_B$ receptor agonist are noted as more potent agonists than that employed and warrant further study. Conversely, if a suspected antagonist is utilized along with agonist, the inquiry would be whether the cells treated with the agonist emitted a greater signal than the cells treated with the agonist and the suspected antagonist, in which case the compound would be an antagonist of the $GABA_B$ receptor. Suitable

reporter genes for use in the above assays are green fluorescent protein (GFP), β-galactosidase, and luciferase.

A sensitive transcription based assay in accordance with the instant invention employs a plasmid encoding β-lactamase under the control of an inducible promoter. This plasmid is transfected into cells together with a plasmid encoding a receptor for which it is desired to identify modulators such as agonists or allosteric modulators of agonists. The inducible promoter on the β-lactamase is chosen so that it responds to at least one intracellular signal that is generated when an agonist or modulator of same binds to the receptor. Thus, following such binding of agonist or modulator to receptor, the level of β-lactamase in the transfected cells increases. This increase in β-lactamase is measured by treating the cells with a cell-permeable dye that is a substrate for cleavage by β-lactamase. The dye contains two fluorescent moieties. In the intact dye, the two fluorescent moieties are close enough to one another that fluorescence resonance energy transfer (FRET) can take place between them. Following cleavage of the dye into two parts by β-lactamase, the two fluorescent moieties are located on different parts, and thus can drift apart. This increases the distance between the fluorescent moieties, thus decreasing the amount of FRET that can occur between them. It is this decrease in FRET that is measured in the assay.

Through the use of the above assay wherein the inducible promoter that drives β-lactamase is activated by an intracellular signal generated by the interaction of the $GABA_B$ receptor and modulators thereof, one of ordinary skill in the art could identify modulators of $GABA_B$ receptors such as agonists, allosteric modulators of agonists, inverse agonists, and antagonists. To produce the $GABA_B$ receptor, a plasmid encoding HG20 and a plasmid encoding Gb1c are transfected into the cells. Cells endogenously expressing functional $GABA_B$ receptor are also suitable in the instant assay. The cells are exposed to the cell-permeable dye and then exposed to compounds suspected of being agonists of the $GABA_B$ receptor. Those compounds that cause a decrease in FRET are likely to actually be agonists of the $GABA_B$ receptor. To determine if the compounds that are identified as agonists are sufficiently potent agonists to warrant further study, one could compare the decrease in FRET caused by the compounds to the decrease in FRET that is caused by baclofen when baclofen is substituted for the compound in the assay.

Accordingly, the present invention includes a method for identifying modulators of the $GABA_B$ receptor comprising: (a) providing cells expressing (i) HG20 or a functional equivalent thereof, (ii) a receptor polypeptide comprising the extracellular region of Gb1c, and (iii) β-lactamase under the control of an inducible promoter that is activated by an intracellular signal generated by the interaction of $GABA_B$ receptor agonist and the $GABA_B$ receptor; (b) exposing the cells to a substrate of β-lactamase that is a cell-permeable dye that contains two fluorescent moieties where the two fluorescent moieties are on different parts of the dye and cleavage of the dye by β-lactamase allows the two fluorescent moieties to drift apart; and (c) measuring the amount of fluorescence resonance energy transfer (FRET) in the cells following contact with a suspected $GABA_B$ receptor modulator. In the instance wherein agonists are sought, the amount of FRET can be compared with that measured in the absence of the suspected agonist to obtain a value for the decrease in FRET caused by the suspected agonist.

A decrease in FRET resulting from a known $GABA_B$ receptor agonist can be determined by measuring the amount of FRET in the presence and in the absence of the agonist. By comparing the decrease in FRET caused by the suspected agonist to the decrease in FRET caused by the known agonist, one can identify modulators, and agonists in particular, of the $GABA_B$ receptor. An agonist would exhibit an equal or greater decrease in FRET than the known agonist.

In fact, by comparing the decrease in FRET caused by a suspected agonist to the decrease in FRET caused by a known $GABA_B$ receptor agonist such as baclofen, one can estimate how potent an agonist the suspect compound is. If the decrease in FRET caused by the compound is larger than the decrease in FRET caused by the agonist, then the compound is likely to be a more potent agonist than the known agonist employed. If the decrease in FRET caused by the compound is about the same as the decrease in FRET caused by the known agonist, then the compound is likely to be an agonist of about the same potency as that employed. If the decrease in FRET caused by the compound is less than the decrease in FRET caused by the known agonist, then the compound is likely to be a weaker agonist than that employed.

The above-described assay can be further modified to an assay for identifying antagonists of the $GABA_B$ receptor. This method comprises (a) providing cells expressing (i) HG20 or a functional equivalent thereof, (ii) a receptor

polypeptide comprising the extracellular region of Gb1c, and (iii) β-lactamase under the control of an inducible promoter that is activated by an intracellular signal generated by the interaction of $GABA_B$ receptor agonist and the $GABA_B$ receptor; (b) exposing the cells to a substrate of β-lactamase that is a cell-permeable dye that contains two fluorescent moieties where the two fluorescent moieties are on different parts of the dye and cleavage of the dye by β-lactamase allows the two fluorescent moieties to drift apart; and (c) measuring the amount of fluorescence resonance energy transfer (FRET) in the cells following contact with a suspected $GABA_B$ receptor antagonist. The amount of FRET can further be measured upon contact with a known $GABA_B$ receptor agonist and the suspected antagonist. This value can then be compared with the value obtained from contact with a known $GABA_B$ receptor agonist; a lesser amount of FRET in the cells contacted with agonist indicating the presence of an antagonist.

In a particular embodiment of the above-described methods, the inducible promoter that is activated by at least one intracellular signal generated by interaction of an agonist with the $GABA_B$ receptor is a promoter that is activated by changes in membrane currents, *e.g.*, changes in potassium currents or activation of the map kinase signaling pathway. In other embodiments, the inducible promoter is activated by the transcription factor NFAT, or is activated by a signal transduced by a chimeric Gq protein, or a signal generated by protein kinase C activity, or by changes in intracellular calcium levels.

The assays described above can be further modified to an additional assay for identifying antagonists of the $GABA_B$ receptor. Such modification would involve the use of β-lactamase under the control of a promoter that is repressed by at least one intracellular signal generated by interaction of a $GABA_B$ receptor agonist with the $GABA_B$ receptor and would also involve running the assay in the presence of said agonist. When the cells are exposed to compounds suspected of being antagonists of the $GABA_B$ receptor, β-lactamase will be induced, and FRET will decrease, if the compound tested is able to counteract the effect of the agonist, *i.e.*, if the compound tested is actually an antagonist.

In a particular embodiment, the inducible promoter that is repressed by at least one intracellular signal generated by interaction of a $GABA_B$ receptor agonist with the $GABA_B$ receptor is a promoter that is repressed by changes in potassium currents.

In particular embodiments of the above-described methods, β-lactamase is TEM-1 β-lactamase from *Escherichia coli.*

In other embodiments, the substrate of β-lactamase is CCF2/AM (Zlokarnik *et al.*, 1998, *Science* 279:84-88).

In particular embodiments of the above-described methods, the cells express a promiscuous G-protein, *e.g.*, Gα15 or Gα16. In other embodiments, the cells have been transfected with an expression vector that directs the expression of a G-protein subunit or subunits.

The assays described above could be modified to identify inverse agonists. In such assays, one would expect a decrease in β-lactamase activity where agonists would produce an increase.

The assays described above could also be modified to identify allosteric modulators of $GABA_B$ receptor agonists. In such assays, one would not expect a response from the modulator in the absence of an agonist.

In particular embodiments of the above-described methods, HG20 is a polypeptide comprising an amino acid sequence selected from the group consisting of: SEQ ID NO:4; positions 9-941 of SEQ ID NO:4; positions 35-941 of SEQ ID NO:4; positions 36-941 of SEQ ID NO:4; positions 38-941 of SEQ ID NO:4; positions 39-941 of SEQ ID NO:4; positions 42-941 of SEQ ID NO:4; positions 44-941 of SEQ ID NO:4; positions 46-941 of SEQ ID NO:4; positions 52-941 of SEQ ID NO:4; positions 57-941 of SEQ ID NO:4; the amino acid sequence encoded by the DNA sequence deposited in GenBank Accession No. AF056085; the amino acid sequence encoded by the DNA sequence deposited in GenBank Accession No. AJ012188; and the amino acid sequence encoded by the DNA sequence deposited in GenBank Accession No. ASF074482.

It is important to note that in all of the above assays, HG20 (or gb2)can be substituted with a functional equivalent such as $GABA_BR2$, or GBR2.

In particular embodiments of the above-described methods, HG20 is a chimeric HG20 protein. By chimeric HG20 protein is meant a contiguous polypeptide sequence of HG20 fused in frame to a polypeptide sequence of a non-HG20 protein. For example, the N-terminal domain and seven transmembrane spanning domains of HG20 fused at the C-terminus in frame to a G protein is a chimeric HG20 protein. Another example of a chimeric HG20 protein is a polypeptide comprising the FLAG epitope fused in frame at the amino terminus of

amino acids 52-941 of SEQ ID NO:4. Especially preferred forms of chimeric HG20 proteins are those in which a non-HG20 polypeptide replaces a portion of the N-terminus of HG20. Chimeric Gb1c proteins or fragments of the $GABA_B$ receptor polypeptide Gb1c may also be used in the present invention. In particular embodiments, the chimeric Gb1c protein comprises the extracellular region of Gb1c with or without the signal sequence. In particularly preferred embodiments, a polypeptide comprising the extracellular region of Gb1c (amino acids 18-525 of SEQ ID NO:2) is fused to another non-Gb1c protein or portion thereof.

In preferred embodiments, the expression vector encoding HG20 comprises a nucleotide sequence selected from the group consisting of: positions 293-3,115 of SEQ ID NO:3; positions 317-3,115 of SEQ ID NO:3; positions 395-3,115 of SEQ ID NO:3; positions 398-3,115 of SEQ ID NO:3; positions 404-3,115 of SEQ ID NO:3; positions 407-3,115 of SEQ ID NO:3; positions 416-3,115 of SEQ ID NO:3; positions 422-3,115 of SEQ ID NO:3; positions 428-3,115 of SEQ ID NO:3; positions 446-3,115 of SEQ ID NO:3; and positions 461-3,115 of SEQ ID NO:3.

Some of the above-described methods can be modified to take advantage of other ways of assaying for agonist activity at the $GABA_B$ receptor. Agonists/antagonists/inverse agonists may affect the internalization or trafficking of functional $GABA_B$ receptors. For example, in the case of the β2-adrenergic receptor, agonist exposure results in receptor internalization. Therefore, receptor trafficking between intracellular pools and the cytoplasmic membrane may be considered an assay of agonist activity. It may be that $GABA_B$ receptor trafficking is modulated by agonists in a similar manner. It would then be possible to identify agonist activity by monitoring $GABA_B$ receptor trafficking. Such trafficking can be monitored by whole cell immunohistochemistry and confocal microscopy or by surface and intracellular receptor labeling and flow cytometry.

Furthermore, because the functional $GABA_B$ receptor may be a heterodimer, then agonists/antagonists/inverse agonists may be expected to alter the ratio of heterodimer to monomer. Hence the disruption or appearance of a heterodimer may be considered an additional screening assay. In this case, the monitoring of receptor dimerization or disappearance may be done by the appearance or disruption of FRET. Each of the monomers are labeled with a fluorophore such that close proximity would allow FRET to occur. Upon agonist binding, one might

see disruption of FRET, indicating disruption of dimers or increase in FRET indicating more dimerization in the course of agonist activation.

While the above-described methods are explicitly directed to testing whether "a" compound is an agonist or antagonist of the $GABA_B$ receptor, it will be clear to one skilled in the art that such methods can be adapted to test collections of compounds, *e.g.*, combinatorial libraries, collections of natural produces, *etc.*, to determine whether any members of such collections are activators or inhibitors of the $GABA_B$ receptor. Accordingly, the use of collections of compounds, or individual members of such collections, as the compound in the above-described methods is within the scope of the present invention.

Agonists and antagonists identified by the above-described methods are useful in the same manner as well-known agonists and antagonists of $GABA_B$ receptors. For example, (-)baclofen is a known agonist of $GABA_B$ receptors and, in racemic form, is a clinically useful muscle relaxant. Gabapentin has been sold since 1994 in the United States as a treatment for epilepsy. The therapeutic potential of $GABA_B$ receptor agonists and antagonists is well documented. For agonists, the therapeutic potential is said to include use as muscle relaxants and anti-asthmatics. For antagonists, the therapeutic potential is said to include use as antidepressants, anticonvulsants, nootropics, and anxiolytics.

Given the wide range of utility displayed by known agonists and antagonists of $GABA_B$ receptors, it is clear that those skilled in the art would consider the agonists and antagonists identified by the methods of the present invention to be pharmacologically useful. In addition, it is believed that such agonists and antagonists will also be useful in the treatment of epilepsy, neuropsychiatric disorders, and dementias.

The following non-limiting Examples are presented to better illustrate the invention.

EXAMPLE 1

Construction of GABAB Receptor Expression Constructs

The human gb1c isoform was obtained from human cerebellum cDNA

(Clontech) by nested PCR cloning using a High-fidelity HF-PCR cloning kit (Clontech). The first step of amplification used primers: hgb1cF-ECOR1 5'-GAT ATC GAA TTC GCC ACC ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC-3' (SEQ ID NO:5) and gb1RR 5'-gc cct tcc cct ctc cct ttc cct ccc-3' (SEQ ID NO:6). PCR conditions were: precycle denaturation of 94°C for 15s, followed by denaturation at 94°C for 15s and an annealing/extension step at 68°C for 4 min for 30 cycles. A final extension step was done at 72°C for 3 min. The second amplification used primers: hgb1cF-ECOR1 and hgb1ECOR1-R 5'-GCC GAG GAA TTC TCA CTT ATA AAG CAA ATG CAC TCG ACT CCC (SEQ ID NO:7). PCR conditions were as described above. PCR products were subcloned into PCR-TA cloning vector (Invitrogen), pT7 vector for expression in *Xenopus* oocytes, pcDNA3.1/Zeo(+) (Invitrogen) for transient high-level expression in COS-7 cells, or pIRES-bleomycin bicistronic vector (Clontech) for stable expression in HEK-293 cells.

EXAMPLE 2
Transient Expression for Radioligand binding

COS-7 (ATCC) cells were grown to ~70% confluence. 6μg of Glaxo 1c or Merck 1c plasmid DNA (Qiagen) were transfected into 1.2 million COS-7 cells using 36μl of lipofectamine as recommended by manufacturer (Gibco BRL).

EXAMPLE 3
Receptor Binding

Cells were harvested 72hr-post transfection. Crude membrane preparations were prepared as described in Belley *et al.*, 1999 *Bioorganic & Medicinal Chemistry* 7:2697-2704. ^{3}H-CGP71872 (a $GABA_B$ receptor antagonist) was custom synthesized by NEN. Saturation binding experiments were conducted using 25μg of membrane and increasing concentration of radiolabeled ^{3}H-CGP71872 ranging from ~0.05nM to 75nM. Non-specific binding was obtained in the presence of 1μM cold CGP71872. Each condition was made in duplicate. Incubation was carried out in 200μl of binding buffer (50mM Tris, pH7.4, 2.5mM $CaCl_2$, 1X protease inhibitor cocktail - Complete Tablet™ (Boehringer Mannheim)) with constant shaking at room temperature for 2hr. Brandel harvester was used to trap bound ligand

onto Whatman GF/B filters. Data analysis was performed using GraphPad Prism (San Diego).

EXAMPLE 4

Transient Expression and Generation of Gb1-Gb2 Stable HEK-293 Cell Lines

Receptor DNAs (6 μg total DNA/1.2 X 10^6 cells) were transfected into COS-7 or HEK-293 cells using 36 μl Lipofectamine reagent (Gibco BRL) according to the manufacturer's instructions. Stable gb2-expressing HEK-293 clones were selected by growth in puromycin (5μg/ml) containing media and dilution cloning. RNA was prepared from 95 clones using Trizol reagent (GibcoBRL) and 10 μg total RNA spotted by vacuum using a dot-blot apparatus onto BrightStar-Plus nylon membranes (Ambion). The blot was hybridized with a ^{32}P-labeled DNA fragment encoding the full-length gb2 receptor (10^6 cpm/ml) in Zip-Hyb solution (Ambion) for 10 h at 50 °C, and washed at 55 °C for 90 min in high-stringency wash buffer. Two high gb2 receptor RNA-expressing clones (gb2.46, and gb2.10) were analysed for cell surface gb2 receptor expression by indirect staining using gb2 antisera 1630.1-1630.2 and goat anti-rabbit antibodies coupled with Alexa-488 on a Becton Dickinson FACSVantage SE flow-cytometer configured to detect FITC fluorescence. Gb1 isoform DNAs were transfected into the gb2.10 clone as described, and bulk gb1c-gb2 stables were selected by growth in puromycin (5μg/ml) and phleomycin (50 μg/ml) containing media. Control stable pIRES-puromycin and pIRES-bleomycin vector expressing HEK-293 cell lines were generated also by growth in antibiotic selection media.

EXAMPLE 5

cAMP Functional Assays

cAMP determinations were made using a scintillation proximity assay (SPA) kit (Amersham, ONT). Briefly, HEK-293 cells were washed, detached, and 77,000-100,000 cells/well resuspended in Hank's Balanced Salt Solution containing 25 mM HEPES pH 7.4, 100 μM 4-(3-butoxy-4-methoxybenzyl)-2-imadazolidinone (Ro 20-1724, BIOMOL, PA) and incubated for 20 min at 37 °C. 2 μM forskolin and ligands (10^{-9}–10^{-3} M) were added and incubated for 30 min at 37 °C. Cells were

lysed by boiling and cAMP levels were determined by SPA according to the manufacturer's instructions. Data were analyzed by nonlinear least-squares regression using the computer-fitting program GraphPad Prism version 2.01 (San Diego).

EXAMPLE 6
In Vitro Receptor Expression

In vitro coupled transcription/translation reactions were performed in the presence of [^{35}S]methionine in the TNT® Coupled Reticulocyte Lysate system (Promega, Madison, WI) using pcDNA3.1 plasmids containing the gb1a, gb1b, and gb1c DNAs. Translation products were analysed by electrophoresis on 8-16% Tris-Glycine SDS gradient gels (Novex pre-cast gel system, San Diego, CA) under denaturing and reducing conditions. Gels were fixed, dried and exposed to Kodak X-AR film at -70 °C for 4 to 24 h.

WHAT IS CLAIMED:

1. An isolated human gamma-aminobutyric acid type B ($GABA_B$) receptor polypeptide comprising the amino acid sequence of SEQ ID NO:2.

2. An isolated human gamma-aminobutyric acid type B ($GABA_B$) receptor polypeptide comprising amino acids 18-899 of SEQ ID NO:2.

3. An isolated human gamma-aminobutyric acid type B ($GABA_B$) gb1c receptor polypeptide comprising a glutamine in place of a histidine at a position corresponding to position 21 of the $GABA_B$ receptor polypeptide of claim 1.

4. A receptor polypeptide comprising the extracellular portion of human gamma-aminobutyric acid type B ($GABA_B$) receptor polypeptide of SEQ ID NO:2.

5. A receptor polypeptide in accordance with claim 4 wherein the extracellular portion is fused with another protein or fragment thereof.

6. A receptor polypeptide in accordance with claim 4 wherein the receptor is a soluble $GABA_B$ receptor.

7. A receptor polypeptide in accordance with claim 4 comprising amino acids 1-525 of SEQ ID NO:2.

8. A receptor polypeptide in accordance with claim 4 comprising amino acids 18-525 of SEQ ID NO:2.

9. An isolated polynucleotide encoding a $GABA_B$ receptor polypeptide in accordance with claim 1.

10. A polynucleotide in accordance with claim 9 which is SEQ ID NO:1.

11. An isolated polynucleotide encoding a $GABA_B$ receptor polypeptide in accordance with claim 2.

12. A polynucleotide in accordance with claim 11 comprising nucleotides 52-2697 of SEQ ID NO:1.

13. An isolated polynucleotide encoding a receptor polypeptide in accordance with claim 4.

14. A polynucleotide in accordance with claim 13 comprising nucleotides 1-1575 of SEQ ID NO:1.

15. A polynucleotide in accordance with claim 13 comprising nucleotides 52-1575 of SEQ ID NO:1.

16. A vector comprising the polynucleotide of claim 10.

17. A vector comprising the polynucleotide of claim 12.

18. A vector comprising the polynucleotide of claim 14.

19. A vector comprising the polynucleotide of claim 15.

20. A host cell comprising the vector of claim 16.

21. A host cell comprising the vector of claim 17.

22. A host cell comprising the vector of claim 18.

23. A host cell comprising the vector of claim 19.

24. A method for the identification of modulators of gamma-aminobutyric acid type B ($GABA_B$) receptors which comprises:

(a) contacting a receptor polypeptide in accordance with claim 8 with a candidate compound, and

(b) determining if binding occurs; binding indicating the presence of a $GABA_B$ receptor modulator.

25. The method of claim 24 wherein the binding is measured in the presence of a competitor.

26. The method of claim 24 wherein the modulators are agonists or antagonists.

27. The method of claim 24 which comprises exposing cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide to a suspected agonist or antagonist in the presence of a known $GABA_B$ receptor agonist; the amount of binding of the known agonist being less in the presence of an agonist or antagonist of the $GABA_B$ receptor.

28. The method of claim 25 wherein the $GABA_B$ receptor agonist is baclofen.

29. The method of claim 27 wherein the cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide are recombinant cells that were transfected with a vector or vectors comprising a polynucleotide(s) encoding one or both of the HG20 or the receptor polypeptide.

30. A method for the identification of compounds which modulate the function of gamma-aminobutyric acid type B ($GABA_B$) receptors which comprises determining whether a candidate compound upon contact with a receptor polypeptide of claim 8 results in a signal generated by activation or inhibition of the polypeptide.

31. The method of claim 30 wherein HG20 or a functional equivalent thereof is expressed along with the receptor polypeptide such that heterodimers of the receptor polypeptide and HG20 are allowed to form.

32. The method of claim 30 wherein the method is done in the presence of a competitor.

33. The method of claim 32 wherein the competitor is baclofen.

34. A method in accordance with claim 30 for the identification of $GABA_B$ receptor antagonists or inverse agonists which comprises:

(a) contacting cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide with a suspected antagonist or inverse agonist of the receptor polypeptide, and

(b) determining whether a functional response follows.

35. The method of claim 34 wherein the cells are contacted with the antagonist or inverse agonist in the presence of a known $GABA_B$ receptor agonist, and wherein the response is measured and compared to that of a control, the control comprising the known agonist; the functional response of the known agonist being less in the presence of a $GABA_B$ receptor antagonist or inverse agonist.

36. The method of claim 35 wherein the response is compared to that of a control, the control comprising contacting a separate group of cells with the antagonist or inverse agonist; a lack of response in the control indicating the presence of an antagonist and a response in contradiction with that of the known agonist indicating the presence of an inverse agonist.

37. The method of claim 34 wherein the cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide are recombinant cells that were transfected with a vector or vectors comprising a polynucleotide(s) encoding one or both of the HG20 or the receptor polypeptide.

38. A method in accordance with claim 30 for the identification of $GABA_B$ receptor agonists or allosteric modulators of agonists which comprises:

(a) contacting cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide with a suspected agonist of the receptor polypeptide or an allosteric modulator of such, and

(b) determining whether a functional response follows.

39. The method of claim 38 wherein the cells are contacted with the agonist or allosteric modulator in the presence of a known $GABA_B$ receptor agonist, and wherein the response is measured and compared to that of a control, the control comprising contact of the receptor polypeptide with the known agonist; the functional response being greater in the presence of a $GABA_B$ receptor agonist or allosteric modulator.

40. The method of claim 39 wherein the response is compared to that of a control, the control comprising contacting a separate group of cells with the agonist or allosteric modulator; a lack of response in the control indicating the presence of an allosteric modulator and a functional response indicating the presence of an agonist.

41. The method of claim 38 wherein the cells expressing HG20 or a functional equivalent thereof and the receptor polypeptide are recombinant cells that were transfected with a vector comprising a polynucleotide(s) encoding one or both of the HG20 or the receptor polypeptide.

MERCK hgb1C DNA AND PROTEIN SEQUENCE

```
1     atgttgctgctgctgctactggcgccactcttcctccgccccccgggcgcgggcggggcg
      tacaacgacgacgacgatgaccgcggtgagaaggaggcgggggggcccgcgcccgccccgc
      M  L  L  L  L  L  L  A  P  L  F  L  R  P  P  G  A  G  G  A

61    cagacccccaacgccacctcagaaggttgccagatcatacacccgccctgggaaggggggc
      gtctgggggttgcggtggagtcttccaacggtctagtatgtgggcgggacccttcccccg
      Q  T  P  N  A  T  S  E  G  C  Q  I  I  H  P  P  W  E  G  G

121   atcaggtaccggggcctgactcgggaccaggtgaaggctatcaacttcctgccagtggac
      tagtccatggccccggactgagccctggtccacttccgatagttgaaggacggtcacctg
      I  R  Y  R  G  L  T  R  D  Q  V  K  A  I  N  F  L  P  V  D

181   tatgagattgagtatgtgtgccgggggggagcgcgaggtggtggggcccaaggtccgcaag
      atactctaactcatacacacggccccccctcgcgctccaccaccccgggttccaggcgttc
      Y  E  I  E  Y  V  C  R  G  E  R  E  V  V  G  P  K  V  R  K

241   tgcctggccaacggctcctggacagatatggacacacccagccgctgtgtgaatcgaacg
      acggaccggttgccgaggacctgtctatacctgtgtgggtcggcgacacacttagcttgc
      C  L  A  N  G  S  W  T  D  M  D  T  P  S  R  C  V  N  R  T

301   ccacactcagaacggcgcgcagtgtacatcggggcactgtttcccatgagcgggggctgg
      ggtgtgagtcttgccgcgcgtcacatgtagccccgtgacaaagggtactcgccccgacc
      P  H  S  E  R  R  A  V  Y  I  G  A  L  F  P  M  S  G  G  W

361   ccaggggccaggcctgccagcccgcggtggagatggcgctggaggacgtgaatagccgc
      ggtcccccggtccggacggtcgggcgccacctctaccgcgacctcctgcacttatcggcg
      P  G  G  Q  A  C  Q  P  A  V  E  M  A  L  E  D  V  N  S  R

421   agggacatcctgccggactatgagctcaagctcatccaccacgacagcaagtgtgatcca
      tccctgtaggacggcctgatactcgagttcgagtaggtggtgctgtcgttcacactaggt
      R  D  I  L  P  D  Y  E  L  K  L  I  H  H  D  S  K  C  D  P

481   ggccaagccaccaagtacctatatgagctgctctacaacgaccctatcaagatcatcctt
      ccggttcggtggttcatggatatactcgacgagatgttgctgggatagttctagtaggaa
      G  Q  A  T  K  Y  L  Y  E  L  L  Y  N  D  P  I  K  I  I  L

541   atgcctggctgcagctctgtctccacgctggtggctgaggctgctaggatgtggaacctc
      tacggaccgacgtcgagacagaggtgcgaccaccgactccgacgatcctacaccttggag
      M  P  G  C  S  S  V  S  T  L  V  A  E  A  A  R  M  W  N  L
```

FIG.1A

```
601   attgtgctttcctatggctccagctcaccagccctgtcaaaccggcagcgtttccccact
      taacacgaaaggataccgaggtcgagtggtcgggacagtttggccgtcgcaaaggggtga
      I  V  L  S  Y  G  S  S  S  P  A  L  S  N  R  Q  R  F  P  T

661   ttcttccgaacgcacccatcagccacactccacaaccctacccgcgtgaaactctttgaa
      aagaaggcttgcgtgggtagtcggtgtgaggtgttgggatgggcgcactttgagaaactt
      F  F  R  T  H  P  S  A  T  L  H  N  P  T  R  V  K  L  F  E

721   aagtggggctggaagaagattgctaccatccagcagaccactgaggtcttcacttcgact
      ttcaccccgaccttcttctaacgatggtaggtcgtctggtgactccagaagtgaagctga
      K  W  G  W  K  K  I  A  T  I  Q  Q  T  T  E  V  F  T  S  T

781   ctggacgacctggaggaacgagtgaaggaggctggaattgagattactttccgccagagt
      gacctgctggacctccttgctcacttcctccgaccttaactctaatgaaaggcggtctca
      L  D  D  L  E  E  R  V  K  E  A  G  I  E  I  T  F  R  Q  S

841   ttcttctcagatccagctgtgcccgtcaaaaacctgaagcgccaggatgcccgaatcatc
      aagaagagtctaggtcgacacgggcagtttttggacttcgcggtcctacgggcttagtag
      F  F  S  D  P  A  V  P  V  K  N  L  K  R  Q  D  A  R  I  I

901   gtgggacttttctatgagactgaagcccggaaagtttttttgtgaggtgtacaaggagcgt
      caccctgaaaagatactctgacttcgggcctttcaaaaaacactccacatgttcctcgca
      V  G  L  F  Y  E  T  E  A  R  K  V  F  C  E  V  Y  K  E  R

961   ctctttgggaagaagtacgtctggttcctcattgggtggtatgctgacaattggttcaag
      gagaaacccttcttcatgcagaccaaggagtaacccaccatacgactgttaaccaagttc
      L  F  G  K  K  Y  V  W  F  L  I  G  W  Y  A  D  N  W  F  K

1021  atctacgacccttctatcaactgcacagtggatgagatgactgaggcggtggagggccac
      tagatgctgggaagatagttgacgtgtcacctactctactgactccgccacctcccggtg
      I  Y  D  P  S  I  N  C  T  V  D  E  M  T  E  A  V  E  G  H

1081  atcacaactgagattgtcatgctgaatcctgccaatacccgcagcatttccaacatgaca
      tagtgttgactctaacagtacgacttaggacggttatgggcgtcgtaaaggttgtactgt
      I  T  T  E  I  V  M  L  N  P  A  N  T  R  S  I  S  N  M  T

1141  tcccaggaatttgtggagaaactaaccaagcgactgaaaagacaccctgaggagacagga
      agggtccttaaacacctctttgattggttcgctgacttttctgtgggactcctctgtcct
      S  Q  E  F  V  E  K  L  T  K  R  L  K  R  H  P  E  E  T  G
```

FIG.1B

```
1201  ggcttccaggaggcaccgctggcctatgatgccatctgggccttggcactggccctgaac
      ccgaaggtcctccgtggcgaccggatactacggtagacccggaaccgtgaccgggacttg
      G  F  Q  E  A  P  L  A  Y  D  A  I  W  A  L  A  L  A  L  N

1261  aagacatctggaggaggcggccgttctggtgtgcgcctggaggacttcaactacaacaac
      ttctgtagacctcctccgccggcaagaccacacgcggacctcctgaagttgatgttgttg
      K  T  S  G  G  G  G  R  S  G  V  R  L  E  D  F  N  Y  N  N

1321  cagaccattaccgaccaaatctaccgggcaatgaactcttcgtcctttgagggtgtctct
      gtctggtaatggctggtttagatggcccgttacttgagaagcaggaaactcccacagaga
      Q  T  I  T  D  Q  I  Y  R  A  M  N  S  S  S  F  E  G  V  S

1381  ggccatgtggtgtttgatgccagcggctctcggatggcatggacgcttatcgagcagctt
      ccggtacaccacaaactacggtcgccgagagcctaccgtacctgcgaatagctcgtcgaa
      G  H  V  V  F  D  A  S  G  S  R  M  A  W  T  L  I  E  Q  L

1441  cagggtggcagctacaagaagattggctactatgacagcaccaaggatgatctttcctgg
      gtcccaccgtcgatgttcttctaaccgatgatactgtcgtggttcctactagaaaggacc
      Q  G  G  S  Y  K  K  I  G  Y  Y  D  S  T  K  D  D  L  S  W

1501  tccaaaacagataaatggattggagggtccccccccagctgaccagaccctggtcatcaag
      aggttttgtctatttacctaacctcccaggggggtcgactggtctgggaccagtagttc
      S  K  T  D  K  W  I  G  G  S  P  P  A  D  Q  T  L  V  I  K

1561  acattccgcttcctgtcacagaaactctttatctccgtctcagttctctccagcctgggc
      tgtaaggcgaaggacagtgtctttgagaaatagaggcagagtcaagagaggtcggacccg
      T  F  R  F  L  S  Q  K  L  F  I  S  V  S  V  L  S  S  L  G

1621  attgtcctagctgttgtctgtctgtcctttaacatctacaactcacatgtccgttatatc
      taacaggatcgacaacagacagacaggaaattgtagatgttgagtgtacaggcaatatag
      I  V  L  A  V  V  C  L  S  F  N  I  Y  N  S  H  V  R  Y  I

1681  cagaactcacagcccaacctgaacaacctgactgctgtgggctgctcactggctttagct
      gtcttgagtgtcgggttggacttgttggactgacgacacccgacgagtgaccgaaatcga
      Q  N  S  Q  P  N  L  N  N  L  T  A  V  G  C  S  L  A  L  A

1741  gctgtcttccccctggggctcgatggttaccacattgggaggaaccagtttcctttcgtc
      cgacagaaggggggaccccgagctaccaatggtgtaaccctccttggtcaaaggaaagcag
      A  V  F  P  L  G  L  D  G  Y  H  I  G  R  N  Q  F  P  F  V
```

FIG.1C

```
1801  tgccaggcccgcctctggctcctgggcctgggctttagtctgggctacggttccatgttc
      acggtccgggcggagaccgaggacccggacccgaaatcagacccgatgccaaggtacaag
      C  Q  A  R  L  W  L  L  G  L  G  F  S  L  G  Y  G  S  M  F

1861  accaagatttggtgggtccacacggtcttcacaaagaaggaagaaaagaaggagtggagg
      tggttctaaaccacccaggtgtgccagaagtgtttcttccttcttttcttcctcacctcc
      T  K  I  W  W  V  H  T  V  F  T  K  K  E  E  K  K  E  W  R

1921  aagactctggaaccctggaagctgtatgccacagtgggcctgctggtgggcatggatgtc
      ttctgagaccttgggaccttcgacatacggtgtcacccggacgaccacccgtacctacag
      K  T  L  E  P  W  K  L  Y  A  T  V  G  L  L  V  G  M  D  V

1981  ctcactctcgccatctggcagatcgtggaccctctgcaccggaccattgagacatttgcc
      gagtgagagcggtagaccgtctagcacctgggagacgtggcctggtaactctgtaaacgg
      L  T  L  A  I  W  Q  I  V  D  P  L  H  R  T  I  E  T  F  A

2041  aaggaggaacctaaggaagatattgacgtctctattctgccccagctggagcattgcagc
      ttcctccttggattccttctataactgcagagataagacggggtcgacctcgtaacgtcg
      K  E  E  P  K  E  D  I  D  V  S  I  L  P  Q  L  E  H  C  S

2101  tccaggaagatgaatacatggcttggcattttctatggttacaaggggctgctgctgctg
      aggtccttctacttatgtaccgaaccgtaaaagataccaatgttccccgacgacgacgac
      S  R  K  M  N  T  W  L  G  I  F  Y  G  Y  K  G  L  L  L  L

2161  ctgggaatcttccttgcttatgagaccaagagtgtgtccactgagaagatcaatgatcac
      gacccttagaaggaacgaatactctggttctcacacaggtgactcttctagttactagtg
      L  G  I  F  L  A  Y  E  T  K  S  V  S  T  E  K  I  N  D  H

2221  cgggctgtgggcatggctatctacaatgtggcagtcctgtgcctcatcactgctcctgtc
      gcccgacacccgtaccgatagatgttacaccgtcaggacacggagtagtgacgaggacag
      R  A  V  G  M  A  I  Y  N  V  A  V  L  C  L  I  T  A  P  V

2281  accatgattctgtccagccagcaggatgcagcctttgcctttgcctctcttgccatagtt
      tggtactaagacaggtcggtcgtcctacgtcggaaacggaaacggagagaacggtatcaa
      T  M  I  L  S  S  Q  Q  D  A  A  F  A  F  A  S  L  A  I  V

2341  ttctcctcctatatcactcttgttgtgctctttgtgcccaagatgcgcaggctgatcacc
      aagaggaggatatagtgagaacaacacgagaaacacgggttctacgcgtccgactagtgg
      F  S  S  Y  I  T  L  V  V  L  F  V  P  K  M  R  R  L  I  T
```

FIG.1D

2401 cgaggggaatggcagtcggaggcgcaggacaccatgaagacagggtcatcgaccaacaac
gctccccttaccgtcagcctccgcgtcctgtggtacttctgtcccagtagctggttgttg
R G E W Q S E A Q D T M K T G S S T N N

2461 aacgaggaggagaagtcccggctgttggagaaggagaaccgtgaactggaaaagatcatt
ttgctcctcctcttcagggccgacaacctcttcctcttggcacttgacctttctagtaa
N E E E K S R L L E K E N R E L E K I I

2521 gctgagaaagaggagcgtgtctctgaactgcgccatcaactccagtctcggcagcagctc
cgactctttctcctcgcacagagacttgacgcggtagttgaggtcagagccgtcgtcgag
A E K E E R V S E L R H Q L Q S R Q Q L

2581 cgctcccggcgccacccaccgacacccccagaaccctctgggggcctgcccaggggaccc
gcgagggccgcggtgggtggctgtgggggtcttgggagacccccggacgggtcccctggg
R S R R H P P T P P E P S G G L P R G P

2641 cctgagccccccgaccggcttagctgtgatgggagtcgagtgcatttgctttataagtga
ggactcgggggctggccgaatcgacactaccctcagctcacgtaaacgaaatattcact
P E P P D R L S C D G S R V H L L Y K *

FIG.1E

```
          1                                                   50
glaxo 1c  MLLLLLLAPL FLRPPGAGGA HTPNATSEGC QIIHPPWEGG IRYRGLTRDQ
merck 1c  MLLLLLLAPL FLRPPGAGGA QTPNATSEGC QIIHPPWEGG IRYRGLTRDQ

          51                                                  100
glaxo 1c  VKAINFLPVD YEIEYVCRGE REVVGPKVRK CLANGSWTDM DTPSRCVNRT
merck 1c  VKAINFLPVD YEIEYVCRGE REVVGPKVRK CLANGSWTDM DTPSRCVNRT

          101                                                 150
glaxo 1c  PHSERRAVYI GALFPMSGGW PGGQACQPAV EMALEDVNSR RDILPDYELK
merck 1c  PHSERRAVYI GALFPMSGGW PGGQACQPAV EMALEDVNSR RDILPDYELK

          151                                                 200
glaxo 1c  LIHHDSKCDP GQATKYLYEL LYNDPIKIIL MPGCSSVSTL VAEAARMWNL
merck 1c  LIHHDSKCDP GQATKYLYEL LYNDPIKIIL MPGCSSVSTL VAEAARMWNL

          201                                                 250
glaxo 1c  IVLSYGSSSP ALSNRQRFPT FFRTHPSATL HNPTRVKLFE KWGWKKIATI
merck 1c  IVLSYGSSSP ALSNRQRFPT FFRTHPSATL HNPTRVKLFE KWGWKKIATI

          251                                                 300
glaxo 1c  QQTTEVFTST LDDLEERVKE AGIEITFRQS FFSDPAVPVK NLKRQDARII
merck 1c  QQTTEVFTST LDDLEERVKE AGIEITFRQS FFSDPAVPVK NLKRQDARII

          301                                                 350
glaxo 1c  VGLFYETEAR KVFCEVYKER LFGKKYVWFL IGWYADNWFK IYDPSINCTV
merck 1c  VGLFYETEAR KVFCEVYKER LFGKKYVWFL IGWYADNWFK IYDPSINCTV

          351                                                 400
glaxo 1c  DEMTEAVEGH ITTEIVMLNP ANTRSISNMT SQEFVEKLTK RLKRHPEETG
merck 1c  DEMTEAVEGH ITTEIVMLNP ANTRSISNMT SQEFVEKLTK RLKRHPEETG

          401                                                 450
glaxo 1c  GFQEAPLAYD AIWALALALN KTSGGGGRSG VRLEDFNYNN QTITDQIYRA
merck 1c  GFQEAPLAYD AIWALALALN KTSGGGGRSG VRLEDFNYNN QTITDQIYRA

          451                                                 500
glaxo 1c  MNSSSFEGVS GHVVFDASGS RMAWTLIEQL QGGSYKKIGY YDSTKDDLSW
merck 1c  MNSSSFEGVS GHVVFDASGS RMAWTLIEQL QGGSYKKIGY YDSTKDDLSW
```

FIG.2A

```
          501                                        TM1          550
glaxo 1c  SKTDKWIGGS PPADQTLVIK TFRFLSQKLF ISVSVLSSLG IVLAVVCLSF
merck 1c  SKTDKWIGGS PPADQTLVIK TFRFLSQKLF ISVSVLSSLG IVLAVVCLSF

          551                        TM2                       600
glaxo 1c  NIYNSHVRYI QNSQPNLNNL TAVGCSLALA AVFPLGLDGY HIGRNQFPFV
merck 1c  NIYNSHVRYI QNSQPNLNNL TAVGCSLALA AVFPLGLDGY HIGRNQFPFV

          601        TM3                                    650
glaxo 1c  CQARLWLLGL GFSLGYGSMF TKIWWVHTVF TKKEEKKEWR KTLEPWKLYA
merck 1c  CQARLWLLGL GFSLGYGSMF TKIWWVHTVF TKKEEKKEWR KTLEPWKLYA

          651 TM4                                            700
glaxo 1c  TVGLLVGMDV LTLAIWQIVD PLHRTIETFA KEEPKEDIDV SILPQLEHCS
merck 1c  TVGLLVGMDV LTLAIWQIVD PLHRTIETFA KEEPKEDIDV SILPQLEHCS

          701        TM5                                    750
glaxo 1c  SRKMNTWLGI FYGYKGLLLL LGIFLAYETK SVSTEKINDH RAVGMAIYNV
merck 1c  SRKMNTWLGI FYGYKGLLLL LGIFLAYETK SVSTEKINDH RAVGMAIYNV

          751   TM6                                 TM7        800
glaxo 1c  AVLCLITAPV TMILSSQQDA AFAFASLAIV FSSYITLVVL FVPKMRRLIT
merck 1c  AVLCLITAPV TMILSSQQDA AFAFASLAIV FSSYITLVVL FVPKMRRLIT

          801                                                  850
glaxo 1c  RGEWQSEAQD TMKTGSSTNN NEEEKSRLLE KENRELEKII AEKEERVSEL
merck1c   RGEWQSEAQD TMKTGSSTNN NEEEKSRLLE KENRELEKII AEKEERVSEL

          851                                                  899
glaxo 1c  RHQLQSRQQL RSRRHPPTPP EPSGGLPRGP PEPPDRLSCD GSRVHLLYK
merck 1c  RHQLQSRQQL RSRRHPPTPP EPSGGLPRGP PEPPDRLSCD GSRVHLLYK
```

FIG.2B

```
1    ccgttctgag ccgagccgga accctagccc gagacggagc cggggcccgg gccggcgcca
61   ttgcgcgggc gccgcgggaa gaccttggcg cggggcggcg ggccgggcca ggccatgcgg
121  gccgagtgag ccggcgcccg cagcccgcgg cgcggcatgg cttccccgcg gaggtccggg
181  cagccagggc ggccgccgcc gccgccaccg ccgcccgcgc gcctgctact gctactgctg
241  ctgccgctgc tgctgcctct ggcgcccggg gcctggggct gggcgcgggg cgccccccgg
301  ccgccgccca gcagcccgcc gctctccatc atgggcctca tgccgctcac caaggaggtg
361  gccaagggca gcatcgggcg cggtgtgctc cccgccgtgg aactggccat cgagcagatc
421  cgcaacgagt cactcctgcg cccctacttc ctcgacctgc ggctctatga cacggagtgc
481  gacaacgcaa aagggttgaa agccttctac gatgcaataa aatacgggcc gaaccacttg
541  atggtgtttg gaggcgtctg tccatccgtc acatccatca ttgcagagtc cctccaaggc
601  tggaatctgg tgcagctttc ttttgctgca accacgcctg ttctagccga taagaaaaaa
661  tacccttatt tctttcggac cgtcccatca gacaatgcgg tgaatccagc cattctgaag
721  ttgctcaagc actaccagtg gaagcgcgtg ggcacgctga cgcaagacgt tcagaggttc
781  tctgaggtgc ggaatgacct gactggagtt ctgtatggcg aggacattga gatttcagac
841  accgagagct tctccaacga tccctgtacc agtgtcaaaa agctgaaggg gaatgatgtg
901  cggatcatcc ttggccagtt tgaccagaat atggcagcaa aagtgttctg ttgtgcatac
961  gaggagaaca tgtatggtag taaatatcag tggatcattc cgggctggta cgagccttct
1021 tggtgggagc aggtgcacac ggaagccaac tcatcccgct gcctccggaa gaatctgctt
1081 gctgccatgg agggctacat tggcgtggat ttcgagcccc tgagctccaa gcagatcaag
1141 accatctcag gaaagactcc acagcagtat gagagagagt acaacaacaa gcggtcaggc
1201 gtggggccca gcaagttcca cgggtacgcc tacgatggca tctgggtcat cgccaagaca
1261 ctgcagaggg ccatggagac actgcatgcc agcagccggc accagcggat ccaggacttc
1321 aactacacgg accacacgct gggcaggatc atcctcaatg ccatgaacga gaccaacttc
1381 ttcggggtca cgggtcaagt tgtattccgg aatggggaga gaatggggac cattaaattt
1441 actcaatttc aagacagcag ggaggtgaag gtgggagagt acaacgctgt ggccgacaca
1501 ctggagatca tcaatgacac catcaggttc caaggatccg aaccaccaaa agacaagacc
1561 atcatcctgg agcagctgcg gaagatctcc ctacctctct acagcatcct ctctgccctc
1621 accatcctcg ggatgatcat ggccagtgct tttctcttct tcaacatcaa gaaccggaat
1681 cagaagctca taaagatgtc gagtccatac atgaacaacc ttatcatcct tggagggatg
1741 ctctcctatg cttccatatt tctctttggc cttgatggat cctttgtctc tgaaaagacc
1801 tttgaaacac tttgcaccgt caggacctgg attctcaccg tgggctacac gaccgctttt
1861 ggggccatgt ttgcaaagac ctggagagtc cacgccatct tcaaaaatgt gaaaatgaag
1921 aagaagatca tcaaggacca gaaactgctt gtgatcgtgg ggggcatgct gctgatcgac
1981 ctgtgtatcc tgatctgctg gcaggctgtg gacccctgc gaaggacagt ggagaagtac
2041 agcatggagc cggacccagc aggacgggat atctccatcc gccctctcct ggagcactgt
2101 gagaacaccc atatgaccat ctggcttggc atcgtctatg cctacaaggg acttctcatg
2161 ttgttcggtt gtttcttagc ttgggagacc cgcaacgtca gcatccccgc actcaacgac
2221 agcaagtaca tcgggatgag tgtctacaac gtggggatca tgtgcatcat cggggccgct
2281 gtctccttcc tgacccggga ccagcccaat gtgcagttct gcatcgtggc tctggtcatc
2341 atcttctgca gcaccatcac cctctgcctg gtattcgtgc cgaagctcat caccctgaga
2401 acaaacccag atgcagcaac gcagaacagg cgattccagt tcactcagaa tcagaagaaa
```

FIG.3A

```
2461 gaagattcta aaacgtccac ctcggtcacc agtgtgaacc aagccagcac atcccgcctg
2521 gagggcctac agtcagaaaa ccatcgcctg cgaatgaaga tcacagagct ggataaagac
2581 ttggaagagg tcaccatgca gctgcaggac acaccagaaa agaccaccta cattaaacag
2641 aaccactacc aagagctcaa tgacatcctc aacctgggaa acttcactga gagcacagat
2701 ggaggaaagg ccattttaaa aaatcacctc gatcaaaatc cccagctaca gtggaacaca
2761 acagagccct ctcgaacatg caaagatcct atagaagata taaactctcc agaacacatc
2821 cagcgtcggc tgtccctcca gctccccatc ctccaccacg cctacctccc atccatcgga
2881 ggcgtggacg ccagctgtgt cagcccctgc gtcagcccca ccgccagccc ccgccacaga
2941 catgtgccac cctccttccg agtcatggtc tcgggcctgt aagggtgggg ggcctgggcc
3001 cggggcctcc cccgtgacag aaccacactg ggcagagggg tctgctgcag aaacactgtc
3061 ggctctggct gcgga
```

FIG.3B

```
MASPR RSGQP GRPPP PPPPP ARLLL LLLLP LLLPL APGAW GWARG APRPP PSSPP LSIMG   60
LMPLT KEVAK GSIGR GVLPA VELAI EQIRN ESLLR PYFLD LRLYD TECDN AKGLK AFYDA  120
IKYGP NHLMV FGGVC PSVTS IIAES LQGWN LVQLS FAATT PVLAD KKKYP YFFRT VPSDN  180
AVNPA ILKLL KHYQW KRVGT LTQDV QRFSE VRNDL TGVLY GEDIE ISDTE SFSND PCTSV  240
KKLKG NDVRI ILGQF DQNMA AKVFC CAYEE NMYGS KYQWI IPGWY EPSWW EQVHT EANSS  300
RCLRK NLLAA MEGYI GVDFE PLSSK QIKTI SGKTP QQYER EYNNK RSGVG PSKFH GYAYD  360
GIWVI AKTLQ RAMET LHASS RHQRI QDFNY TDHTL GRIIL NAMNE TNFFG VTGQV VFRNG  420
ERMGT IKFTQ FQDSR EVKVG EYNAV ADTLE IINDT IRFQG SEPPK DKTII LEQLR KISLP  480
LYSIL SALTI LGMIM ASAFL FFNIK NRNQK LIKMS SPYMN NLIIL GGMLS YASIF LFGLD  540
GSFVS EKTFE TLCTV RTWIL TVGYT TAFGA MFAKT WRVHA IFKNV KMKKK IIKDQ KLLVI  600
VGGML LIDLC ILICW QAVDP LRRTV EKYSM EPDPA GRDIS IRPLL EHCEN THMTI WLGIV  660
YAYKG LLMLF GCFLA WETRN VSIPA LNDSK YIGMS VYNVG IMCII GAAVS FLTRD QPNVQ  720
FCIVA LVIIF CSTIT LCLVF VPKLI TLRTN PDAAT QNRRF QFTQN QKKED SKTST SVTSV  780
NQAST SRLEG LQSEN HRLRM KITEL DKDLE EVTMQ LQDTP EKTTY IKQNH YQELN DILNL  840
GNFTE STDGG KAILK NHLDQ NPQLQ WNTTE PSRTC KDPIE DINSP EHIQR RLSLQ LPILH  900
HAYLP SIGGV DASCV SPCVS PTASP RHRHV PPSFR VMVSG L                        941
```

FIG.4

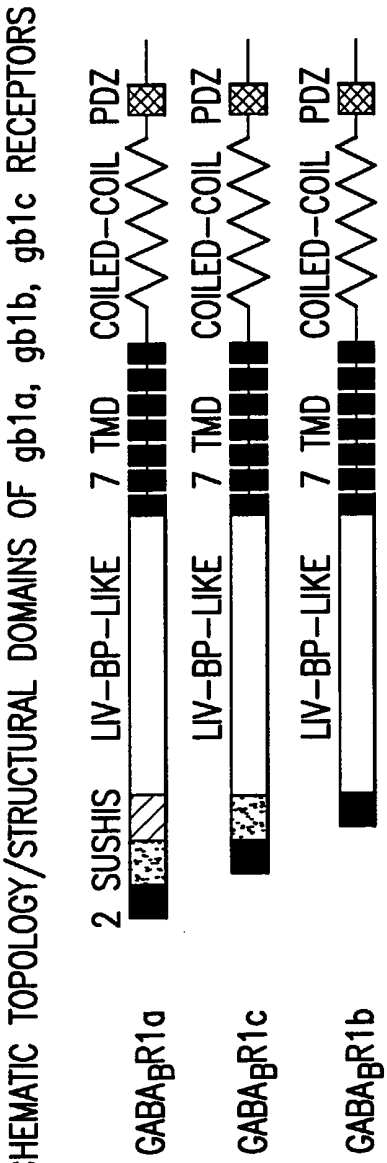

SCHEMATIC TOPOLOGY/STRUCTURAL DOMAINS OF gb1a, gb1b, gb1c RECEPTORS
2 SUSHIS
LIV-BP-LIKE
7 TMD
COILED-COIL
PDZ
GABA$_B$R1a
GABA$_B$R1c
GABA$_B$R1b

FIG. 5A

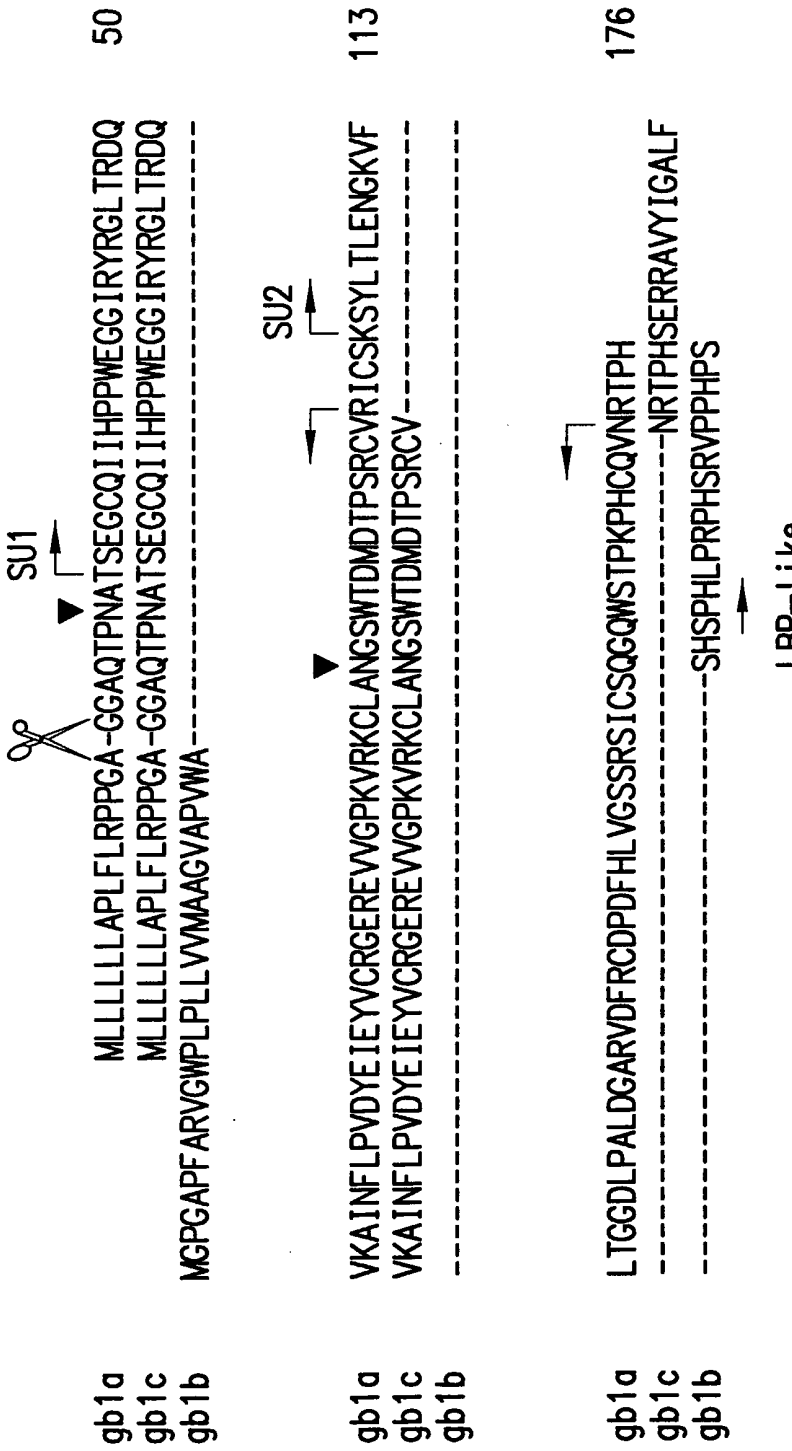

SU1
SU2
LBP-like
gb1a MLLLLLLLAPLFLRPPGA-GGAQTPNATSEGCQIIHPPWEGGIRYRGLTRDQ 50
gb1c MLLLLLLAPLFLRPPGA-GGAQTPNATSEGCQIIHPPWEGGIRYRGLTRDQ
gb1b MGPGAPFARVGWPLPLLVVMAAGVAPVWA----------
gb1a VKAINFLPVDYEIEYVCRGEREVVGPKVRKCLANGSWTDMDTPSRCVRICSKSYLTLENGKVF 113
gb1c VKAINFLPVDYEIEYVCRGEREVVGPKVRKCLANGSWTDMDTPSRCV----------
gb1b ----------
gb1a LTGGDLPALDGARVDFRCDPDFHLVGSSRSICSQGQWSTPKPHCQVNRTPH 176
gb1c ----------NRTPHSERRAVYIGALF
gb1b ----------SHSPHLPRPHSRVPPHPS

FIG. 5B

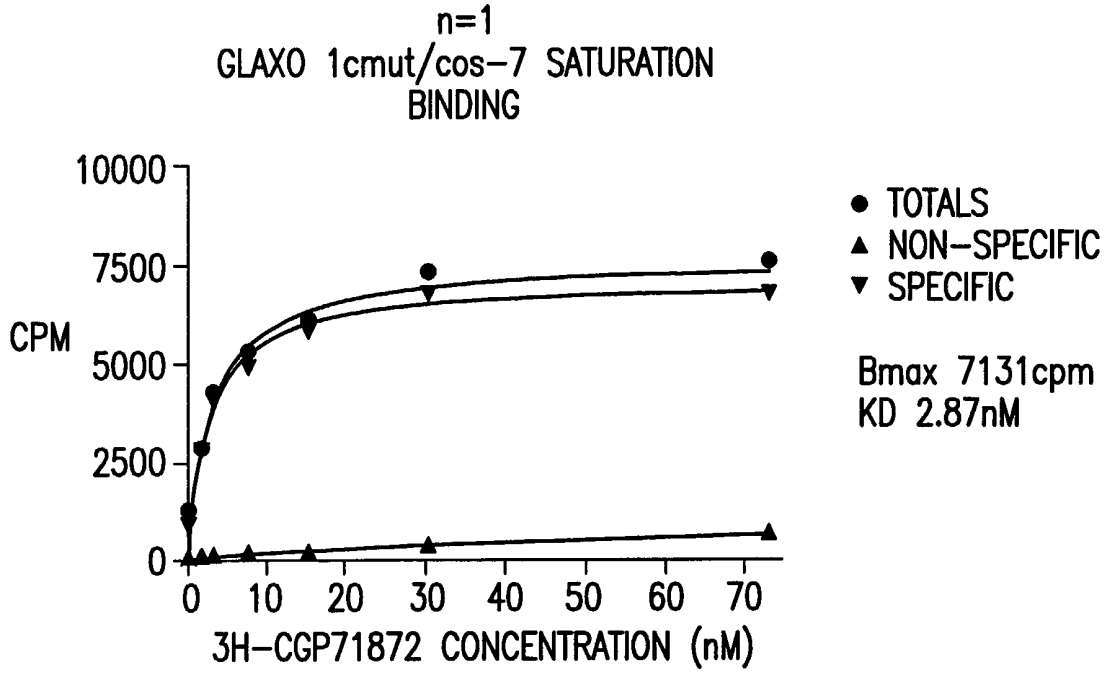

n=1
GLAXO 1cmut/cos-7 SATURATION
BINDING
10000
7500
5000
2500
0
CPM
0 10 20 30 40 50 60 70
3H-CGP71872 CONCENTRATION (nM)
TOTALS
NON-SPECIFIC
SPECIFIC
Bmax 7131cpm
KD 2.87nM

FIG. 6A

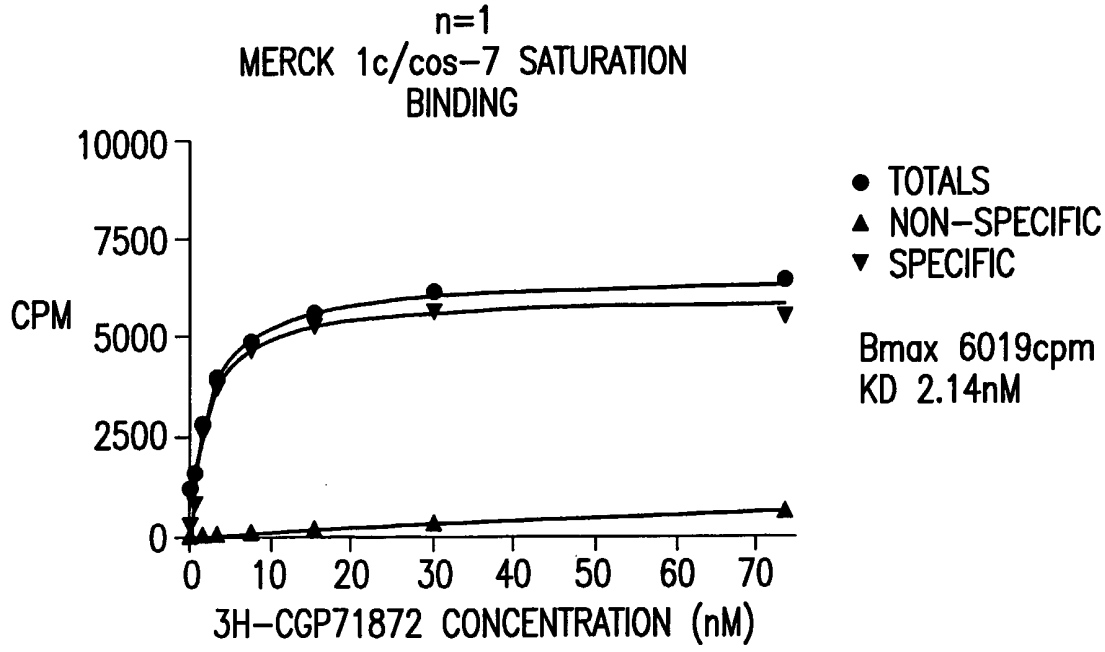

n=1
MERCK 1c/cos-7 SATURATION
BINDING
10000
7500
5000
2500
0
CPM
0 10 20 30 40 50 60 70
3H-CGP71872 CONCENTRATION (nM)
TOTALS
NON-SPECIFIC
SPECIFIC
Bmax 6019cpm
KD 2.14nM

FIG. 6B

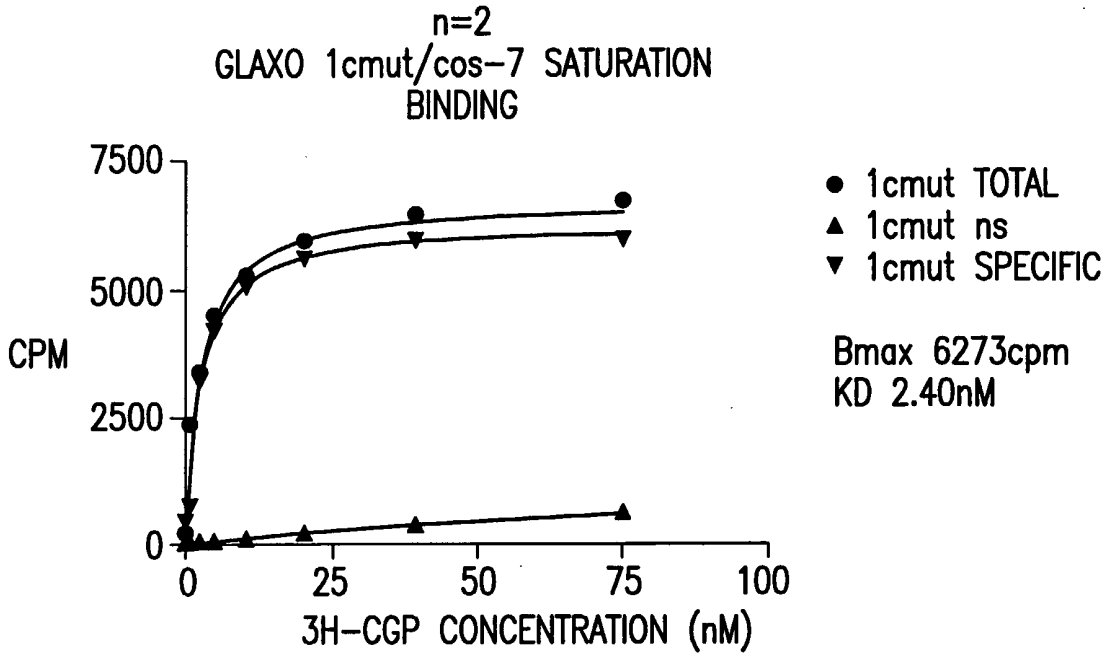

n=2
GLAXO 1cmut/cos-7 SATURATION BINDING
7500
5000
2500
0
CPM
0
25
50
75
100
3H-CGP CONCENTRATION (nM)
1cmut TOTAL
1cmut ns
1cmut SPECIFIC
Bmax 6273cpm
KD 2.40nM

FIG. 6C

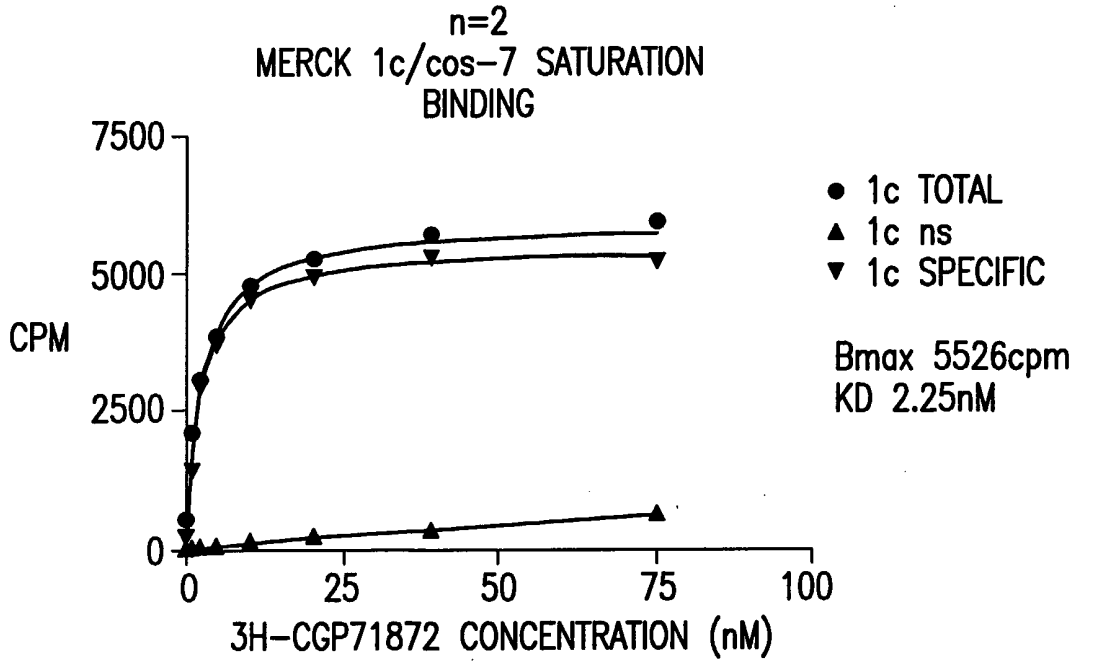

n=2
MERCK 1c/cos-7 SATURATION BINDING
7500
5000
2500
0
CPM
0
25
50
75
100
3H-CGP71872 CONCENTRATION (nM)
1c TOTAL
1c ns
1c SPECIFIC
Bmax 5526cpm
KD 2.25nM

FIG. 6D

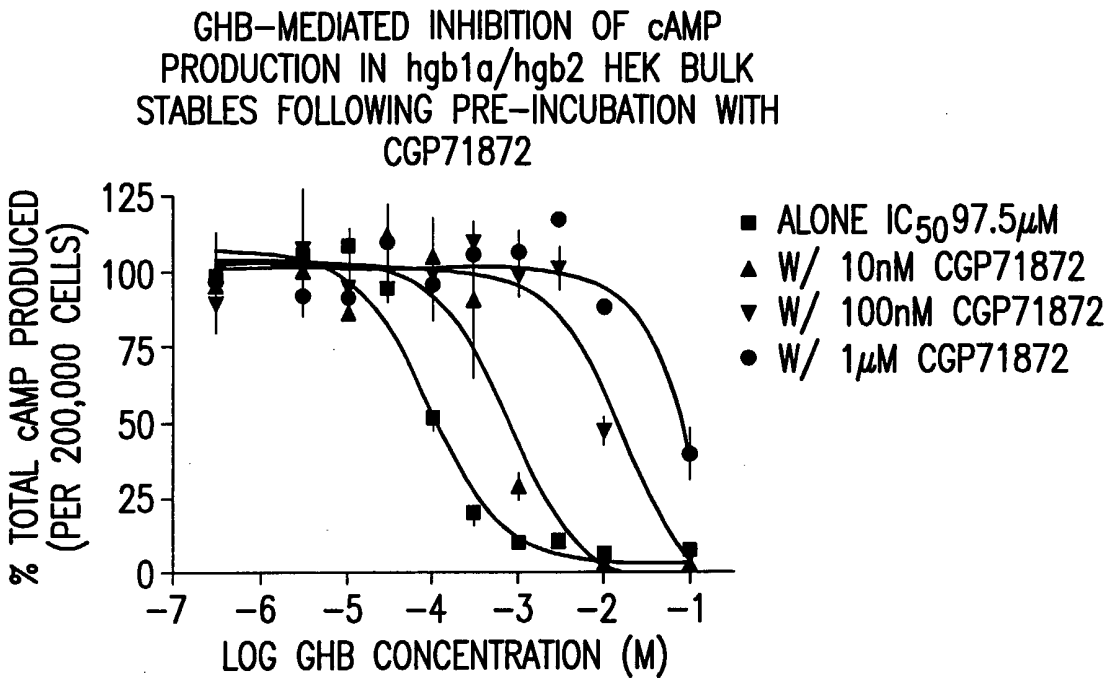

GHB-MEDIATED INHIBITION OF cAMP PRODUCTION IN hgb1a/hgb2 HEK BULK STABLES FOLLOWING PRE-INCUBATION WITH CGP71872
% TOTAL cAMP PRODUCED (PER 200,000 CELLS)
125
100
75
50
25
0
-7
-6
-5
-4
-3
-2
-1
LOG GHB CONCENTRATION (M)
ALONE IC50 97.5μM
W/ 10nM CGP71872
W/ 100nM CGP71872
W/ 1μM CGP71872

FIG. 7A

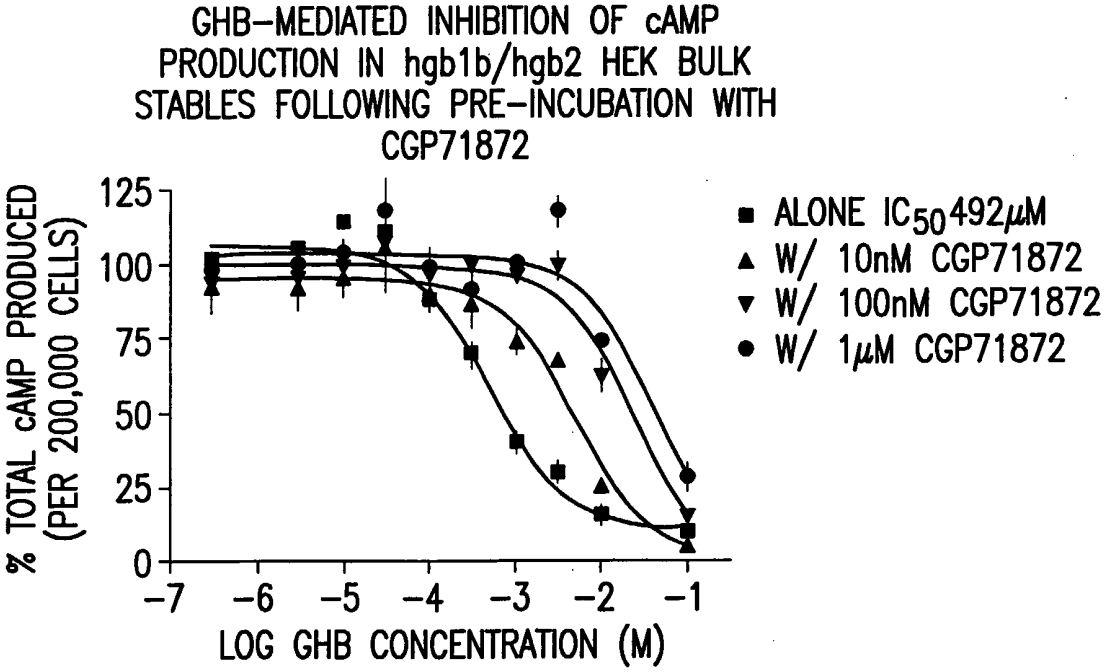

GHB-MEDIATED INHIBITION OF cAMP PRODUCTION IN hgb1b/hgb2 HEK BULK STABLES FOLLOWING PRE-INCUBATION WITH CGP71872
% TOTAL cAMP PRODUCED (PER 200,000 CELLS)
125
100
75
50
25
0
-7
-6
-5
-4
-3
-2
-1
LOG GHB CONCENTRATION (M)
ALONE IC50 492μM
W/ 10nM CGP71872
W/ 100nM CGP71872
W/ 1μM CGP71872

FIG. 7B

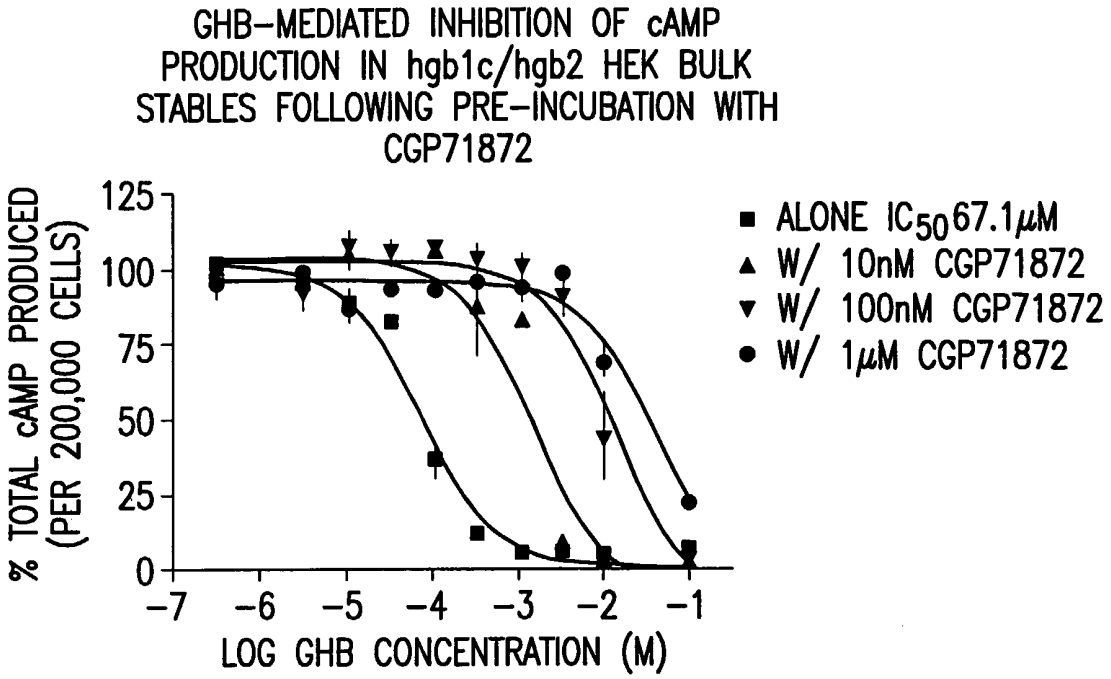

GHB-MEDIATED INHIBITION OF cAMP PRODUCTION IN hgb1c/hgb2 HEK BULK STABLES FOLLOWING PRE-INCUBATION WITH CGP71872
% TOTAL cAMP PRODUCED (PER 200,000 CELLS)
125
100
75
50
25
0
-7
-6
-5
-4
-3
-2
-1
LOG GHB CONCENTRATION (M)
ALONE IC50 67.1μM
W/ 10nM CGP71872
W/ 100nM CGP71872
W/ 1μM CGP71872

FIG. 7C

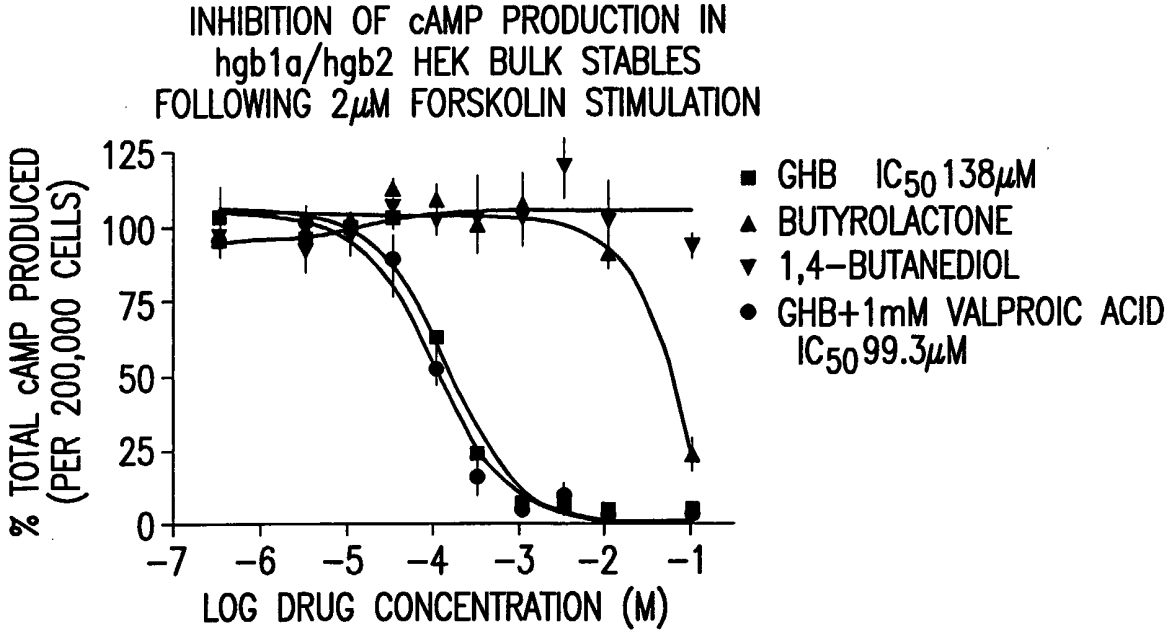

INHIBITION OF cAMP PRODUCTION IN hgb1a/hgb2 HEK BULK STABLES FOLLOWING 2μM FORSKOLIN STIMULATION
% TOTAL cAMP PRODUCED (PER 200,000 CELLS)
125
100
75
50
25
0
-7
-6
-5
-4
-3
-2
-1
LOG DRUG CONCENTRATION (M)
GHB IC50 138μM
BUTYROLACTONE
1,4-BUTANEDIOL
GHB+1mM VALPROIC ACID IC50 99.3μM

FIG. 7D

SEQUENCE LISTING

<110> Merck Frosst Canada & Co.

<120> NOVEL GB1C ISOFORM AND NUCLEOTIDES ENCODING SAME

<130> 8426-1413PCT KPM/ch

<140> 60/207,031
<141> 2000-05-25

<160> 9

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2700
<212> DNA
<213> Homo sapiens

<400> 1
atgttgctgc tgctgctact ggcgccactc ttcctccgcc ccccgggcgc gggcggggcg 60
cagaccccca acgccacctc agaaggttgc cagatcatac acccgccctg ggaagggggc 120
atcaggtacc ggggcctgac tcgggaccag gtgaaggcta tcaacttcct gccagtggac 180
tatgagattg agtatgtgtg ccgggggggag cgcgaggtgg tggggcccaa ggtccgcaag 240
tgcctggcca acggctcctg gacagatatg gacacaccca gccgctgtgt gaatcgaacg 300
ccacactcag aacggcgcgc agtgtacatc ggggcactgt ttcccatgag cgggggctgg 360
ccaggggcc aggcctgcca gcccgcggtg gagatggcgc tggaggacgt gaatagccgc 420
agggacatcc tgccggacta tgagctcaag ctcatccacc acgacagcaa gtgtgatcca 480
ggccaagcca ccaagtacct atatgagctg ctctacaacg accctatcaa gatcatcctt 540
atgcctggct gcagctctgt ctccacgctg gtggctgagg ctgctaggat gtggaacctc 600
attgtgcttt cctatggctc cagctcacca gccctgtcaa accggcagcg tttccccact 660
ttcttccgaa cgcacccatc agccacactc cacaacccta cccgcgtgaa actctttgaa 720
aagtggggct ggaagaagat tgctaccatc cagcagacca ctgaggtctt cacttcgact 780
ctggacgacc tggaggaacg agtgaaggag gctggaattg agattacttt ccgccagagt 840
ttcttctcag atccagctgt gcccgtcaaa aacctgaagc gccaggatgc ccgaatcatc 900
gtgggacttt tctatgagac tgaagcccgg aaagtttttt gtgaggtgta caaggagcgt 960
ctctttggga agaagtacgt ctggttcctc attgggtggt atgctgacaa ttggttcaag 1020
atctacgacc cttctatcaa ctgcacagtg gatgagatga ctgaggcggt ggagggccac 1080
atcacaactg agattgtcat gctgaatcct gccaataccc gcagcatttc caacatgaca 1140
tcccaggaat ttgtggagaa actaaccaag cgactgaaaa gacaccctga ggagacagga 1200
ggcttccagg aggcaccgct ggcctatgat gccatctggg ccttggcact ggccctgaac 1260
aagacatctg gaggaggcgg ccgttctggt gtgcgcctgg aggacttcaa ctacaacaac 1320
cagaccatta ccgaccaaat ctaccgggca atgaactctt cgtcctttga gggtgtctct 1380
ggccatgtgg tgtttgatgc cagcggctct cggatggcat ggacgcttat cgagcagctt 1440
cagggtggca gctacaagaa gattggctac tatgacagca ccaaggatga tctttcctgg 1500
tccaaaacag ataaatggat tggagggtcc cccccagctg accagaccct ggtcatcaag 1560
acattccgct tcctgtcaca gaaactcttt atctccgtct cagttctctc cagcctgggc 1620
attgtcctag ctgttgtctg tctgtccttt aacatctaca actcacatgt ccgttatatc 1680
cagaactcac agcccaacct gaacaacctg actgctgtgg gctgctcact ggctttagct 1740
gctgtcttcc ccctggggct cgatggttac cacattggga ggaaccagtt tcctttcgtc 1800
tgccaggccc gcctctggct cctgggcctg ggctttagtc tgggctacgg ttccatgttc 1860
accaagattt ggtgggtcca cacggtcttc acaaagaagg aagaaaagaa ggagtggagg 1920
aagactctgg aaccctggaa gctgtatgcc acagtgggcc tgctggtggg catggatgtc 1980
ctcactctcg ccatctggca gatcgtggac cctctgcacc ggaccattga gacatttgcc 2040

```
aaggaggaac ctaaggaaga tattgacgtc tctattctgc cccagctgga gcattgcagc   2100
tccaggaaga tgaatacatg gcttggcatt ttctatggtt acaaggggct gctgctgctg   2160
ctgggaatct tccttgctta tgagaccaag agtgtgtcca ctgagaagat caatgatcac   2220
cgggctgtgg gcatggctat ctacaatgtg gcagtcctgt gcctcatcac tgctcctgtc   2280
accatgattc tgtccagcca gcaggatgca gcctttgcct ttgcctctct tgccatagtt   2340
ttctcctcct atatcactct tgttgtgctc tttgtgccca agatgcgcag gctgatcacc   2400
cgaggggaat ggcagtcgga ggcgcaggac accatgaaga cagggtcatc gaccaacaac   2460
aacgaggagg agaagtcccg gctgttggag aaggagaacc gtgaactgga aaagatcatt   2520
gctgagaaag aggagcgtgt ctctgaactg cgccatcaac tccagtctcg gcagcagctc   2580
cgctcccggc gccacccacc gacaccccca gaaccctctg ggggcctgcc caggggaccc   2640
cctgagcccc ccgaccggct tagctgtgat gggagtcgag tgcatttgct ttataagtga   2700

<210> 2
<211> 899
<212> PRT
<213> Homo sapiens

<400> 2
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95
Val Asn Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala
            100                 105                 110
Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro
        115                 120                 125
Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu
    130                 135                 140
Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro
145                 150                 155                 160
Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile
                165                 170                 175
Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala
            180                 185                 190
Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser
        195                 200                 205
Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr
    210                 215                 220
His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu
225                 230                 235                 240
Lys Trp Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val
                245                 250                 255
Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly
            260                 265                 270
Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro
        275                 280                 285
Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe
    290                 295                 300
```

```
Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg
305                 310                 315                 320
Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp
                325                 330                 335
Asn Trp Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu
            340                 345                 350
Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu
        355                 360                 365
Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe
    370                 375                 380
Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly
385                 390                 395                 400
Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala
                405                 410                 415
Leu Ala Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg
            420                 425                 430
Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr
        435                 440                 445
Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val
    450                 455                 460
Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu
465                 470                 475                 480
Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp
                485                 490                 495
Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro
            500                 505                 510
Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys
        515                 520                 525
Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala
    530                 535                 540
Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile
545                 550                 555                 560
Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser
                565                 570                 575
Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile
            580                 585                 590
Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu
        595                 600                 605
Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp
    610                 615                 620
Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg
625                 630                 635                 640
Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val
                645                 650                 655
Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu
            660                 665                 670
His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile
        675                 680                 685
Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met
    690                 695                 700
Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu
705                 710                 715                 720
Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys
                725                 730                 735
Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val
            740                 745                 750
```

```
Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln
        755                 760                 765
Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr
    770                 775                 780
Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr
785                 790                 795                 800
Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser
                805                 810                 815
Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu
            820                 825                 830
Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser
        835                 840                 845
Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg
    850                 855                 860
His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro
865                 870                 875                 880
Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu
                885                 890                 895
Leu Tyr Lys
```

<210> 3
<211> 3075
<212> DNA
<213> Homo sapiens

<400> 3

```
ccgttctgag ccgagccgga accctagccc gagacggagc cggggcccgg gccggcgcca     60
ttgcgcgggc gccgcgggaa gaccttggcg cggggcggcg ggccgggcca ggccatgcgg    120
gccgagtgag ccggcgcccg cagcccgcgg cgcggcatgg cttccccgcg gaggtccggg    180
cagccagggc ggccgccgcc gccgccaccg ccgcccgcgc gcctgctact gctactgctg    240
ctgccgctgc tgctgcctct ggcgcccggg gcctggggct gggcgcgggg cgccccccgg    300
ccgccgccca gcagcccgcc gctctccatc atgggcctca tgccgctcac caaggaggtg    360
gccaagggca gcatcgggcg cggtgtgctc cccgccgtgg aactggccat cgagcagatc    420
cgcaacgagt cactcctgcg cccctacttc ctcgacctgc ggctctatga cacggagtgc    480
gacaacgcaa aagggttgaa agccttctac gatgcaataa aatacgggcc gaaccacttg    540
atggtgtttg gaggcgtctg tccatccgtc acatccatca ttgcagagtc cctccaaggc    600
tggaatctgg tgcagctttc ttttgctgca accacgcctg ttctagccga taagaaaaaa    660
tacccttatt tctttcggac cgtcccatca gacaatgcgg tgaatccagc cattctgaag    720
ttgctcaagc actaccagtg gaagcgcgtg ggcacgctga cgcaagacgt tcagaggttc    780
tctgaggtgc ggaatgacct gactggagtt ctgtatggcg aggacattga gatttcagac    840
accgagagct tctccaacga tccctgtacc agtgtcaaaa agctgaaggg gaatgatgtg    900
cggatcatcc ttggccagtt tgaccagaat atggcagcaa aagtgttctg ttgtgcatac    960
gaggagaaca tgtatggtag taaatatcag tggatcattc cgggctggta cgagccttct   1020
tggtgggagc aggtgcacac ggaagccaac tcatcccgct gcctccggaa gaatctgctt   1080
gctgccatgg agggctacat tggcgtggat ttcgagcccc tgagctccaa gcagatcaag   1140
accatctcag gaaagactcc acagcagtat gagagagagt acaacaacaa gcggtcaggc   1200
gtggggccca gcaagttcca cgggtacgcc tacgatggca tctgggtcat cgccaagaca   1260
ctgcagaggg ccatggagac actgcatgcc agcagccggc accagcggat ccaggacttc   1320
aactacacgg accacacgct gggcaggatc atcctcaatg ccatgaacga gaccaacttc   1380
ttcggggtca cgggtcaagt tgtattccgg aatggggaga gaatggggac cattaaattt   1440
actcaatttc aagacagcag ggaggtgaag gtgggagagt acaacgctgt ggccgacaca   1500
ctggagatca tcaatgacac catcaggttc caaggatccg aaccaccaaa agacaagacc   1560
atcatcctgg agcagctgcg gaagatctcc ctacctctct acagcatcct ctctgccctc   1620
accatcctcg ggatgatcat ggccagtgct tttctcttct tcaacatcaa gaaccggaat   1680
cagaagctca taaagatgtc gagtccatac atgaacaacc ttatcatcct tggagggatg   1740
```

```
ctctcctatg cttccatatt tctctttggc cttgatggat cctttgtctc tgaaaagacc   1800
tttgaaacac tttgcaccgt caggacctgg attctcaccg tgggctacac gaccgctttt   1860
ggggccatgt ttgcaaagac ctggagagtc cacgccatct tcaaaaatgt gaaaatgaag   1920
aagaagatca tcaaggacca gaaactgctt gtgatcgtgg ggggcatgct gctgatcgac   1980
ctgtgtatcc tgatctgctg gcaggctgtg gacccccctgc gaaggacagt ggagaagtac   2040
agcatggagc cggacccagc aggacgggat atctccatcc gccctctcct ggagcactgt   2100
gagaacaccc atatgaccat ctggcttggc atcgtctatg cctacaaggg acttctcatg   2160
ttgttcggtt gtttcttagc ttgggagacc cgcaacgtca gcatccccgc actcaacgac   2220
agcaagtaca tcgggatgag tgtctacaac gtggggatca tgtgcatcat cggggccgct   2280
gtctccttcc tgacccggga ccagcccaat gtgcagttct gcatcgtggc tctggtcatc   2340
atcttctgca gcaccatcac cctctgcctg gtattcgtgc cgaagctcat caccctgaga   2400
acaaacccag atgcagcaac gcagaacagg cgattccagt tcactcagaa tcagaagaaa   2460
gaagattcta aaacgtccac ctcggtcacc agtgtgaacc aagccagcac atcccgcctg   2520
gagggcctac agtcagaaaa ccatcgcctg cgaatgaaga tcacagagct ggataaagac   2580
ttggaagagg tcaccatgca gctgcaggac acaccagaaa agaccaccta cattaaacag   2640
aaccactacc aagagctcaa tgacatcctc aacctgggaa acttcactga gagcacagat   2700
ggaggaaagg ccattttaaa aaatcacctc gatcaaaatc cccagctaca gtggaacaca   2760
acagagccct ctcgaacatg caaagatcct atagaagata taaactctcc agaacacatc   2820
cagcgtcggc tgtccctcca gctccccatc ctccaccacg cctacctccc atccatcgga   2880
ggcgtggacg ccagctgtgt cagcccctgc gtcagcccca ccgccagccc ccgccacaga   2940
catgtgccac cctccttccg agtcatggtc tcgggcctgt aagggtgggg ggcctgggcc   3000
cggggcctcc cccgtgacag aaccacactg ggcagagggg tctgctgcag aaacactgtc   3060
ggctctggct gcgga                                                    3075
```

<210> 4
<211> 941
<212> PRT
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 4

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

```
            180                 185                 190
Tyr Gln Trp Lys Arg Val Gly Thr Leu Thr Gln Asp Val Gln Arg Phe
        195                 200                 205
Ser Glu Val Arg Asn Asp Leu Thr Gly Val Leu Tyr Gly Glu Asp Ile
    210                 215                 220
Glu Ile Ser Asp Thr Glu Ser Phe Ser Asn Asp Pro Cys Thr Ser Val
225                 230                 235                 240
Lys Lys Leu Lys Gly Asn Asp Val Arg Ile Ile Leu Gly Gln Phe Asp
                245                 250                 255
Gln Asn Met Ala Ala Lys Val Phe Cys Cys Ala Tyr Glu Glu Asn Met
            260                 265                 270
Tyr Gly Ser Lys Tyr Gln Trp Ile Ile Pro Gly Trp Tyr Glu Pro Ser
        275                 280                 285
Trp Trp Glu Gln Val His Thr Glu Ala Asn Ser Ser Arg Cys Leu Arg
    290                 295                 300
Lys Asn Leu Leu Ala Ala Met Glu Gly Tyr Ile Gly Val Asp Phe Glu
305                 310                 315                 320
Pro Leu Ser Ser Lys Gln Ile Lys Thr Ile Ser Gly Lys Thr Pro Gln
                325                 330                 335
Gln Tyr Glu Arg Glu Tyr Asn Asn Lys Arg Ser Gly Val Gly Pro Ser
            340                 345                 350
Lys Phe His Gly Tyr Ala Tyr Asp Gly Ile Trp Val Ile Ala Lys Thr
        355                 360                 365
Leu Gln Arg Ala Met Glu Thr Leu His Ala Ser Ser Arg His Gln Arg
    370                 375                 380
Ile Gln Asp Phe Asn Tyr Thr Asp His Thr Leu Gly Arg Ile Ile Leu
385                 390                 395                 400
Asn Ala Met Asn Glu Thr Asn Phe Phe Gly Val Thr Gly Gln Val Val
                405                 410                 415
Phe Arg Asn Gly Glu Arg Met Gly Thr Ile Lys Phe Thr Gln Phe Gln
            420                 425                 430
Asp Ser Arg Glu Val Lys Val Gly Glu Tyr Asn Ala Val Ala Asp Thr
        435                 440                 445
Leu Glu Ile Ile Asn Asp Thr Ile Arg Phe Gln Gly Ser Glu Pro Pro
    450                 455                 460
Lys Asp Lys Thr Ile Ile Leu Glu Gln Leu Arg Lys Ile Ser Leu Pro
465                 470                 475                 480
Leu Tyr Ser Ile Leu Ser Ala Leu Thr Ile Leu Gly Met Ile Met Ala
                485                 490                 495
Ser Ala Phe Leu Phe Phe Asn Ile Lys Asn Arg Asn Gln Lys Leu Ile
            500                 505                 510
Lys Met Ser Ser Pro Tyr Met Asn Asn Leu Ile Ile Leu Gly Gly Met
        515                 520                 525
Leu Ser Tyr Ala Ser Ile Phe Leu Phe Gly Leu Asp Gly Ser Phe Val
    530                 535                 540
Ser Glu Lys Thr Phe Glu Thr Leu Cys Thr Val Arg Thr Trp Ile Leu
545                 550                 555                 560
Thr Val Gly Tyr Thr Thr Ala Phe Gly Ala Met Phe Ala Lys Thr Trp
                565                 570                 575
Arg Val His Ala Ile Phe Lys Asn Val Lys Met Lys Lys Lys Ile Ile
            580                 585                 590
Lys Asp Gln Lys Leu Leu Val Ile Val Gly Gly Met Leu Leu Ile Asp
        595                 600                 605
Leu Cys Ile Leu Ile Cys Trp Gln Ala Val Asp Pro Leu Arg Arg Thr
    610                 615                 620
Val Glu Lys Tyr Ser Met Glu Pro Asp Pro Ala Gly Arg Asp Ile Ser
```

625 630 635 640
Ile Arg Pro Leu Leu Glu His Cys Glu Asn Thr His Met Thr Ile Trp
645 650 655
Leu Gly Ile Val Tyr Ala Tyr Lys Gly Leu Leu Met Leu Phe Gly Cys
660 665 670
Phe Leu Ala Trp Glu Thr Arg Asn Val Ser Ile Pro Ala Leu Asn Asp
675 680 685
Ser Lys Tyr Ile Gly Met Ser Val Tyr Asn Val Gly Ile Met Cys Ile
690 695 700
Ile Gly Ala Ala Val Ser Phe Leu Thr Arg Asp Gln Pro Asn Val Gln
705 710 715 720
Phe Cys Ile Val Ala Leu Val Ile Ile Phe Cys Ser Thr Ile Thr Leu
725 730 735
Cys Leu Val Phe Val Pro Lys Leu Ile Thr Leu Arg Thr Asn Pro Asp
740 745 750
Ala Ala Thr Gln Asn Arg Arg Phe Gln Phe Thr Gln Asn Gln Lys Lys
755 760 765
Glu Asp Ser Lys Thr Ser Thr Ser Val Thr Ser Val Asn Gln Ala Ser
770 775 780
Thr Ser Arg Leu Glu Gly Leu Gln Ser Glu Asn His Arg Leu Arg Met
785 790 795 800
Lys Ile Thr Glu Leu Asp Lys Asp Leu Glu Glu Val Thr Met Gln Leu
805 810 815
Gln Asp Thr Pro Glu Lys Thr Thr Tyr Ile Lys Gln Asn His Tyr Gln
820 825 830
Glu Leu Asn Asp Ile Leu Asn Leu Gly Asn Phe Thr Glu Ser Thr Asp
835 840 845
Gly Gly Lys Ala Ile Leu Lys Asn His Leu Asp Gln Asn Pro Gln Leu
850 855 860
Gln Trp Asn Thr Thr Glu Pro Ser Arg Thr Cys Lys Asp Pro Ile Glu
865 870 875 880
Asp Ile Asn Ser Pro Glu His Ile Gln Arg Arg Leu Ser Leu Gln Leu
885 890 895
Pro Ile Leu His His Ala Tyr Leu Pro Ser Ile Gly Gly Val Asp Ala
900 905 910
Ser Cys Val Ser Pro Cys Val Ser Pro Thr Ala Ser Pro Arg His Arg
915 920 925
His Val Pro Pro Ser Phe Arg Val Met Val Ser Gly Leu
930 935 940

<210> 5
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 5
gatatcgaat tcgccaccat gttgctgctg ctgctactgg cgccactc 48

<210> 6
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 6
gcccttcccc tctccctttc cctccc 26

<210> 7
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 7
gccgaggaat tctcacttat aaagcaaatg cactcgactc cc 42

<210> 8
<211> 899
<212> PRT
<213> Homo sapiens

<400> 8
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15
Ala Gly Gly Ala His Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95
Val Asn Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala
100 105 110
Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro
115 120 125
Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu
130 135 140
Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro
145 150 155 160
Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile
165 170 175
Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala
180 185 190
Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser
195 200 205
Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr
210 215 220
His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu
225 230 235 240
Lys Trp Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val
245 250 255
Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly

```
            260                 265                 270
Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro
        275                 280                 285
Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe
    290                 295                 300
Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg
305                 310                 315                 320
Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp
                325                 330                 335
Asn Trp Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu
            340                 345                 350
Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu
        355                 360                 365
Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe
    370                 375                 380
Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly
385                 390                 395                 400
Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala
                405                 410                 415
Leu Ala Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg
            420                 425                 430
Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr
        435                 440                 445
Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val
    450                 455                 460
Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu
465                 470                 475                 480
Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp
                485                 490                 495
Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro
            500                 505                 510
Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys
        515                 520                 525
Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala
    530                 535                 540
Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile
545                 550                 555                 560
Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser
                565                 570                 575
Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile
            580                 585                 590
Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu
        595                 600                 605
Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp
    610                 615                 620
Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg
625                 630                 635                 640
Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val
                645                 650                 655
Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu
            660                 665                 670
His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile
        675                 680                 685
Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met
    690                 695                 700
Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu
```

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

<210> 9
<211> 2700
<212> DNA
<213> Homo sapiens

<400> 9

```
tacaacgacg acgacgatga ccgcggtgag aaggaggcgg ggggcccgcg cccgccccgc     60
gtctgggggt tgcggtggag tcttccaacg gtctagtatg tgggcgggac ccttcccccg    120
tagtccatgg ccccggactg agccctggtc cacttccgat agttgaagga cggtcacctg    180
atactctaac tcatacacac ggcccccctc gcgctccacc accccgggtt ccaggcgttc    240
acggaccggt tgccgaggac ctgtctatac ctgtgtgggt cggcgacaca cttagcttgc    300
ggtgtgagtc ttgccgcgcg tcacatgtag ccccgtgaca aagggtactc gcccccgacc    360
ggtccccgg tccggacggt cgggcgccac ctctaccgcg acctcctgca cttatcggcg    420
tccctgtagg acggcctgat actcgagttc gagtaggtgg tgctgtcgtt cacactaggt    480
ccggttcggt ggttcatgga tatactcgac gagatgttgc tgggatagtt ctagtaggaa    540
tacggaccga cgtcgagaca gaggtgcgac caccgactcc gacgatccta caccttggag    600
taacacgaaa ggataccgag gtcgagtggt cgggacagtt tggccgtcgc aaaggggtga    660
aagaaggctt gcgtgggtag tcggtgtgag gtgttgggat gggcgcactt tgagaaactt    720
ttcaccccga ccttcttcta acgatggtag gtcgtctggt gactccagaa gtgaagctga    780
gacctgctgg acctccttgc tcacttcctc cgaccttaac tctaatgaaa ggcggtctca    840
aagaagagtc taggtcgaca cgggcagttt ttggacttcg cggtcctacg ggcttagtag    900
caccctgaaa agatactctg acttcgggcc tttcaaaaaa cactccacat gttcctcgca    960
gagaaaccct tcttcatgca gaccaaggag taacccacca tacgactgtt aaccaagttc   1020
tagatgctgg gaagatagtt gacgtgtcac ctactctact gactccgcca cctcccggtg   1080
tagtgttgac tctaacagta cgacttagga cggttatggg cgtcgtaaag gttgtactgt   1140
agggtcctta aacacctctt tgattggttc gctgactttt ctgtgggact cctctgtcct   1200
ccgaaggtcc tccgtggcga ccggatacta cggtagaccc ggaaccgtga ccgggacttg   1260
ttctgtagac ctcctccgcc ggcaagacca cacgcggacc tcctgaagtt gatgttgttg   1320
gtctggtaat ggctggttta gatggcccgt tacttgagaa gcaggaaact cccacagaga   1380
ccggtacacc acaaactacg gtcgccgaga gcctaccgta cctgcgaata gctcgtcgaa   1440
```

```
gtcccaccgt cgatgttctt ctaaccgatg atactgtcgt ggttcctact agaaaggacc   1500
aggttttgtc tatttaccta acctcccagg gggggtcgac tggtctggga ccagtagttc   1560
tgtaaggcga aggacagtgt ctttgagaaa tagaggcaga gtcaagagag gtcggacccg   1620
taacaggatc gacaacagac agacaggaaa ttgtagatgt tgagtgtaca ggcaatatag   1680
gtcttgagtg tcgggttgga cttgttggac tgacgacacc cgacgagtga ccgaaatcga   1740
cgacagaagg gggaccccga gctaccaatg gtgtaaccct ccttggtcaa aggaaagcag   1800
acggtccggg cggagaccga ggacccggac ccgaaatcag acccgatgcc aaggtacaag   1860
tggttctaaa ccacccaggt gtgccagaag tgtttcttcc ttcttttctt cctcacctcc   1920
ttctgagacc ttgggacctt cgacatacgg tgtcacccgg acgaccaccc gtacctacag   1980
gagtgagagc ggtagaccgt ctagcacctg ggagacgtgg cctggtaact ctgtaaacgg   2040
ttcctccttg gattccttct ataactgcag agataagacg gggtcgacct cgtaacgtcg   2100
aggtccttct acttatgtac cgaaccgtaa aagataccaa tgttccccga cgacgacgac   2160
gacccttaga aggaacgaat actctggttc tcacacaggt gactcttcta gttactagtg   2220
gcccgacacc cgtaccgata gatgttacac cgtcaggaca cggagtagtg acgaggacag   2280
tggtactaag acaggtcggt cgtcctacgt cggaaacgga aacggagaga acggtatcaa   2340
aagaggagga tatagtgaga acaacacgag aaacacgggt tctacgcgtc cgactagtgg   2400
gctcccctta ccgtcagcct ccgcgtcctg tggtacttct gtcccagtag ctggttgttg   2460
ttgctcctcc tcttcagggc cgacaacctc ttcctcttgg cacttgacct tttctagtaa   2520
cgactctttc tcctcgcaca gagacttgac gcggtagttg aggtcagagc cgtcgtcgag   2580
gcgagggccg cggtgggtgg ctgtgggggt cttgggagac ccccggacgg gtcccctggg   2640
ggactcgggg ggctggccga atcgacacta ccctcagctc acgtaaacga aatattcact   2700
```