Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2109193 C 2003/04/08

(11)(21) 2 109 193

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 1992/06/09
(87) Date publication PCT/PCT Publication Date: 1992/12/23
(45) Date de délivrance/Issue Date: 2003/04/08
(85) Entrée phase nationale/National Entry: 1993/10/25
(86) N° demande PCT/PCT Application No.: GB 1992/001031
(87) N° publication PCT/PCT Publication No.: 1992/022652
(30) Priorité/Priority: 1991/06/11 (9112504.7) GB

(51) Cl.Int.^{5}/Int.Cl.^{5} C12N 15/12, C12N 15/85, C12N 5/10, G01N 33/566, C07K 14/705

(72) Inventeurs/Inventors:
HADINGHAM, KAREN L., GB;
LE BOURDELLES, BEATRICE, GB;
WHITING, PAUL J., GB;
WINGROVE, PETER B., GB

(73) Propriétaire/Owner:
MERCK SHARP & DOHME LIMITED, GB

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : SOUS-UNITES (α-2,α-3,α-5,α-6,β-2) DU RECEPTEUR A GABA-A ET CELLULES TRANSFECTEES LES EXPRIMANT

(54) Title: GABA-A RECEPTOR SUBUNITS (α-2,α-3,α-5,α-6,β-2) AND TRANSFECTED CELLS EXPRESSING THEM

(57) Abrégé/Abstract:
The present invention relates to a stably co-transfected eukaryotic cell line capable of expressing a $GABA_A$ receptor, particularly a human $GABA_A$ receptor, which receptor comprises at least one alpha, one beta and one gamma subunit; to the cloning of novel cDNA sequences encoding the α_2, α_3, α_5, α_6 and β_2 subunits of the human $GABA_A$ receptor; and to the use of the cell line in designing and developing $GABA_A$ receptor subtype-selective medicaments.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/12, C07K 13/00
C12P 21/02 2109193

A1

(11) International Publication Number: WO 92/22652

(43) International Publication Date: 23 December 1992 (23.12.92)

(21) International Application Number: PCT/GB92/01031

(22) International Filing Date: 9 June 1992 (09.06.92)

(30) Priority data:
9112504.7 11 June 1991 (11.06.91) GB

(71) Applicant *(for all designated States except US):* MERCK SHARP & DOHME LIMITED [GB/GB]; Hertford Road, Hoddesdon, Hertfordshire EN11 9BU (GB).

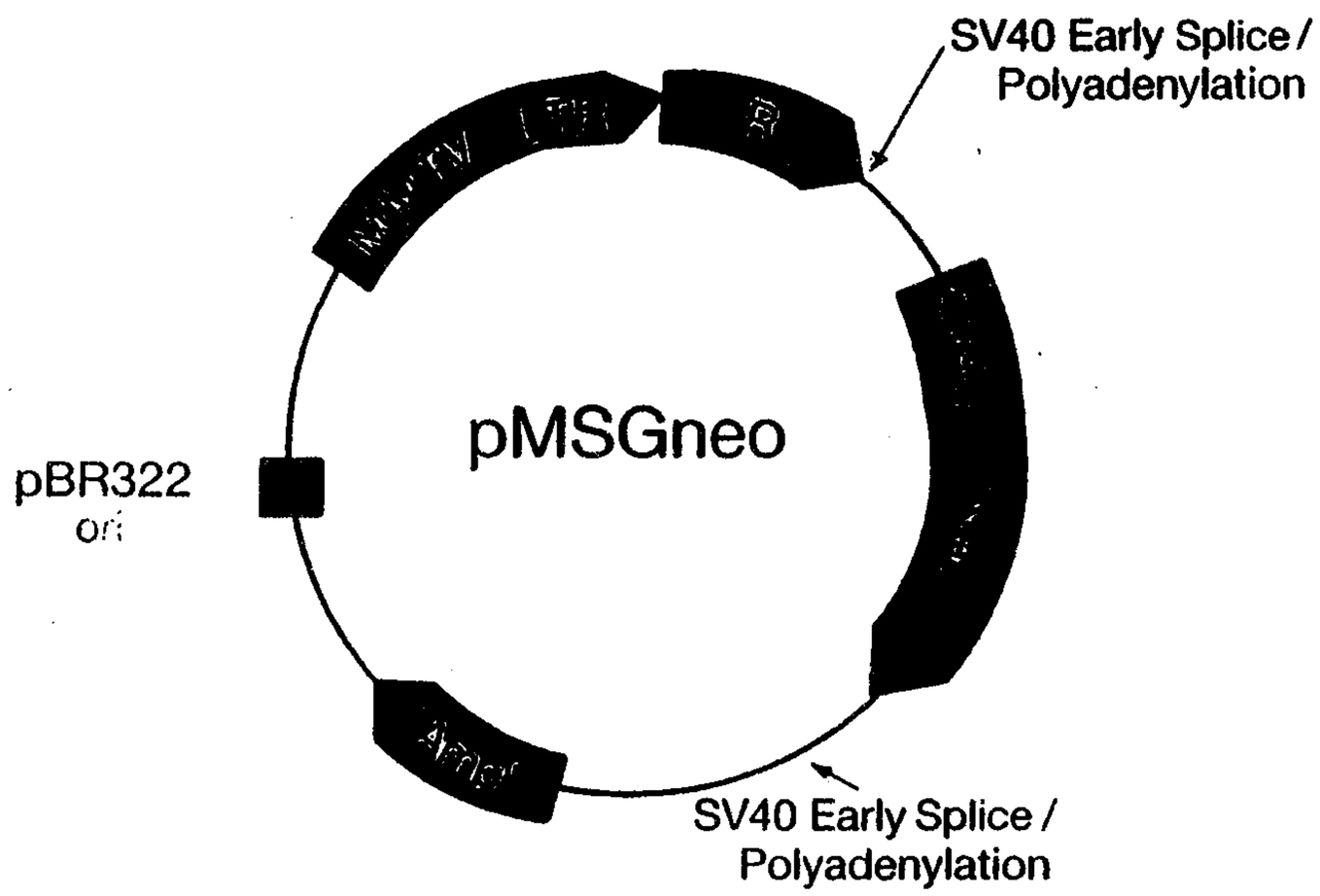
(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* HADINGHAM, Karen, Louise [GB/GB]; 83 Cambridge Road, Sawbridgeworth, Hertfordshire CM21 9BU (GB). LE BOURDELLES, Beatrice [FR/GB]; 100 Herons Wood, Harlow, Essex CM20 1RT (GB). WHITING, Paul, John [GB/GB]; Tithe Barn, Bentfield Green, Stansted Mountfitchet, Essex CM24 8HX (GB). WINGROVE, Peter, Baxter [GB/GB]; Fairgreen House, Sawbridgeworth, Hertfordshire CM21 9AJ (GB).

(74) Agent: THOMPSON, John; Merck & Co., Inc., European Patent Department, Terlings Park, Eastwick Road, Harlow, Essex CM20 2QR (GB).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC (European patent), MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, RU, SD, SE, SE (European patent), SN (OAPI patent), TD (OAPI patent), TG (OAPI patent), US.

Published
With international search report.

(54) Title: GABA-A RECEPTOR SUBUNITS (α-2, α-3, α-5, α-6, β-2) AND TRANSFECTED CELLS EXPRESSING THEM



(57) Abstract

The present invention relates to a stably co-transfected eukaryotic cell line capable of expressing a $GABA_A$ receptor, particularly a human $GABA_A$ receptor, which receptor comprises at least one alpha, one beta and one gamma subunit; to the cloning of novel cDNA sequences encoding the α_2, α_3, α_5, α_6 and β_2 subunits of the human $GABA_A$ receptor; and to the use of the cell line in designing and developing $GABA_A$ receptor subtype-selective medicaments.

GABA-A RECEPTOR SUBUNITS (α-2,α-3,α-5,α-6,β-2) AND TRANSFECTED CELLS EXPRESSING THEM

This invention concerns a cell line, and in particular relates to a stable cell line capable of expressing human or animal $GABA_A$ receptors. The invention further concerns the cloning of novel cDNA sequences encoding particular subunits of the human $GABA_A$ receptor. In addition, the invention relates to the use of the cell line in a screening technique for the design and development of subtype-specific medicaments.

Gamma-amino butyric acid (GABA) is a major inhibitory neurotransmitter in the central nervous system. It mediates fast synaptic inhibition by opening the chloride channel intrinsic to the $GABA_A$ receptor. This receptor comprises a multimeric protein of molecular size 230-270 kDa with specific binding sites for a variety of drugs including benzodiazepines, barbiturates and β-carbolines, in addition to sites for the agonist ligand GABA (for reviews see Stephenson, Biochem. J., 1988, 249, 21; Olsen and Tobin, Faseb J., 1990, 4, 1469; and Sieghart, Trends in Pharmacol. Sci., 1989, 10, 407).

Molecular biological studies demonstrate that the receptor is composed of several distinct types of subunit, which are divided into four classes (α, β, γ, and δ) based on their sequence similarities. To date, six types of α (Schofield et al., Nature (London), 1987, 328, 221; Levitan et al., Nature (London), 1988, 335, 76; Ymer et al., EMBO J., 1989, 8, 1665; Pritchett & Seeberg, J. Neurochem., 1990, 54, 802; Luddens et al., Nature (London), 1990, 346, 648; and Khrestchatisky et al., Neuron, 1989, 3, 745), three types of β (Ymer et al., EMBO J., 1989, 8, 1665), two types of γ (Ymer et al.,

EMBO J., 1990, 9, 3261; and Shivers et al., Neuron, 1989, 3, 327) and one δ subunit (Shivers et al., Neuron, 1989, 3, 327) have been identified.

The differential distribution of many of the subunits has been characterised by in situ hybridisation (Sequier et al., Proc. Natl. Acad. Sci. USA, 1988, 85, 7815; Malherbe et al., J. Neurosci., 1990, 10, 2330; and Shivers et al., Neuron, 1989, 3, 327) and this has permitted it to be speculated which subunits, by their co-localisation, could theoretically exist in the same receptor complex.

Various combinations of subunits have been co-transfected into cells to identify synthetic combinations of subunits whose pharmacology parallels that of bona fide $GABA_A$ receptors in vivo (Pritchett et al., Science, 1989, 245, 1389; Malherbe et al., J. Neurosci., 1990, 10, 2330; Pritchett and Seeberg, J. Neurochem., 1990, 54, 1802; and Luddens et al., Nature (London), 1990, 346, 648). This approach has revealed that, in addition to an α and β subunit, either γ_1 or γ_2 (Pritchett et al., Nature (London), 1989, 338, 582; Ymer et al., EMBO J., 1990, 9, 3261; and Malherbe et al., J. Neurosci., 1990, 10, 2330) or γ_3 (Herb et al., Proc. Natl. Acad. Sci. USA, 1992, 89, 1433; Knoflach et al., FEBS Lett., 1991, 293, 191; and Wilson-Shaw et al., FEBS Lett., 1991, 284, 211) is also generally required to confer benzodiazepine sensitivity, and that the benzodiazepine pharmacology of the expressed receptor is largely dependent on the identity of the α and γ subunits present. Receptors containing a δ subunit (i.e. αβδ) do not appear to bind benzodiazepines (Shivers et al., Neuron, 1989, 3, 327). Combinations of subunits have been identified which exhibit the pharmacological profile of a BZ_1 type receptor ($\alpha_1\beta_1\gamma_2$) and a BZ_2 type receptor ($\alpha_2\beta_1\gamma_2$ or

$\alpha_3\beta_1\gamma_2$, Pritchett et al., Nature (London), 1989, 338, 582), as well as two $GABA_A$ receptors with a novel pharmacology, $\alpha_5\beta_2\gamma_2$ (Pritchett and Seeberg, J. Neurochem., 1990, 54, 1802) and $\alpha_6\beta_2\gamma_2$ (Luddens et al., Nature (London), 1990, 346, 648). Although the pharmacology of these expressed receptors appears similar to that of those identified in brain tissue by radioligand binding, it has nonetheless not been shown that these receptor subunit combinations exist in vivo.

The present invention is concerned with the production of permanently transfected cells containing the $GABA_A$ receptor, which will be useful for screening for drugs which act on this receptor. The $GABA_A$ receptor has previously been expressed in Xenopus oocytes (Sïgel et al., Neuron, 1990, 5, 703-711) and in transiently transfected mammalian cells (Pritchett et al., Science, 1989, 245, 1389-1392). However, both of those systems involve transient expression and are unsuitable for screening purposes.

We have now achieved the stable expression of the receptor.

Accordingly, the present invention provides a stably co-transfected eukaryotic cell line capable of expressing a $GABA_A$ receptor, which receptor comprises at least one alpha, one beta and one gamma subunit.

This has been achieved by co-transfecting cells with three expression vectors, each harbouring cDNAs encoding for an α, β or γ $GABA_A$ receptor subunit. In a further aspect, therefore, the present invention provides a process for the preparation of a eukaryotic cell line capable of expressing a $GABA_A$ receptor, which comprises stably co-transfecting a eukaryotic host cell with at least three expression vectors, one such vector harbouring the cDNA sequence encoding for an alpha,

another such vector harbouring the cDNA sequence encoding for a beta, and a third such vector harbouring the cDNA sequence encoding for a gamma $GABA_A$ receptor subunit. The stable cell-line which is established expresses an $\alpha\beta\gamma$ $GABA_A$ receptor. Each receptor thereby expressed, comprising a unique combination of α, β and γ subunits, will be referred to hereinafter as a $GABA_A$ receptor "subunit combination". Pharmacological and electrophysiological data confirm that the recombinant $\alpha\beta\gamma$ receptor expressed by the cells of the present invention has the properties expected of a native $GABA_A$ receptor.

Expression of the $GABA_A$ receptor may be accomplished by a variety of different promoter-expression systems in a variety of different host cells. The eukaryotic host cells suitably include yeast, insect and mammalian cells. Preferably the eukaryotic cells which can provide the host for the expression of the receptor are mammalian cells. Suitable host cells include rodent fibroblast lines, for example mouse Ltk^-, Chinese hamster ovary (CHO) and baby hamster kidney (BHK); HeLa; and HEK293 cells. It is necessary to incorporate at least one α, one β and one γ subunit into the cell line in order to produce the required receptor. Within this limitation, the choice of receptor subunit combination is made according to the type of activity or selectivity which is being screened for. For example, benzodiazepines (designated BZ) represent one class of drugs which act upon the $GABA_A$ receptor. The presence of an α_1 subunit is specific for a class of benzodiazepines having the pharmacology designated BZ_1; whereas α_2 to α_5 define different pharmacological profiles, broadly designated as BZ_2. The type of β subunit is not critical in defining the class of benzodiazepine, although a β

subunit is required. The γ subunit is also important in defining BZ selectivity. It is likely that differentiation between α subunit selectivity is conferred by the identity of the particular γ subunit present.

In order to employ this invention most effectively for screening purposes, it is preferable to build up a library of cell lines, each with a different combination of subunits. Typically a library of 5 or 6 cell line types is convenient for this purpose. Preferred subunit combinations include: $\alpha_1\beta_1\gamma_2$; $\alpha_1\beta_2\gamma_2$; $\alpha_2\beta_1\gamma_1$; $\alpha_2\beta_1\gamma_2$; $\alpha_2\beta_1\gamma_3$; $\alpha_3\beta_1\gamma_2$; $\alpha_3\beta_1\gamma_3$; $\alpha_4\beta_1\gamma_2$; $\alpha_5\beta_1\gamma_2$; and $\alpha_6\beta_1\gamma_2$; especially $\alpha_1\beta_1\gamma_{2L}$.

The DNAs for the receptor subunits can be obtained from known sources, and are generally obtained as specific nucleotide sequences harboured by a standard cloning vector such as those described, for example, by Maniatis et al. in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, New York, 2nd edition, 1989. Preferably the cDNA sequences are derived from the human gene. However, for screening purposes, cDNAs from other species are also suitable, such as bovine or rat DNA. Known sources of $GABA_A$ receptor subunit cDNAs are as follows:

α_1 bovine) β_1 bovine)	Schofield et al., Nature, 1987, 328, 221-227.
α_1 human) β_1 human)	Schofield et al., FEBS Lett., 1989, 244, 361-364.
α_2 rat	Khrestchatisky et al., J. Neurochem., 1991, 56, 1717.

α_2 bovine)	Levitan et al., Nature, 1988, 335,
α_3 bovine)	76-79.
α_4 rat	Wisden et al., FEBS Lett., 1991, 289, 227.
α_4 bovine	Ymer et al., FEBS Lett., 1989, 258, 119-122.
α_5 rat	Pritchett and Seeburg, J. Neurochem., 1990, 54, 1802-1804.
α_6 rat)	Luddens et al., Nature, 1990, 346,
α_6 bovine)	648-651.
β_2 bovine)	Ymer et al., EMBO J., 1989, 8, 1665-1670.
β_2 rat)	
β_3 bovine)	
β_3 rat)	
γ_1 human)	Ymer et al., EMBO J., 1990, 9, 3261-3267.
γ_1 rat)	
γ_1 bovine)	
γ_2 human	Pritchett et al., Nature, 1989, 338, 582-585.
γ_2 bovine	Whiting et al., Proc. Natl. Acad. Sci. USA, 1990, 57, 9966-9970.
γ_3 rat	Herb et al., Proc. Natl. Acad. Sci. USA, 1992, 89, 1433; and Knoflach et al., FEBS Lett., 1991, 293, 191.

γ_3 mouse	Wilson-Shaw et al., FEBS Lett., 1991, 284, 211.
δ rat	Shivers et al., Neuron, 1989, 3, 327.

Certain cDNA sequences encoding various subunits of the human $GABA_A$ receptor have hitherto been unavailable. These include in particular the sequences encoding the α_2, α_3, α_5, α_6 and β_2 subunits, which nucleotide sequences are accordingly novel. We have now ascertained the cDNA sequences of the α_2, α_3, α_5, α_6 and β_2 subunits of the human $GABA_A$ receptor. These nucleotide sequences, together with the deduced amino acid sequences corresponding thereto, are depicted in Figures 2 to 6 of the accompanying drawings. The present invention accordingly provides in several additional aspects DNA molecules encoding the α_2, α_3, α_5, α_6 and β_2 subunits of the human $GABA_A$ receptor comprising all or a portion of the sequences depicted in Figures 2, 3, 4, 5 and 6 respectively, or substantially similar sequences.

The sequencing of the novel cDNA molecules in accordance with the invention can conveniently be carried out by the standard procedure described in accompanying Example 3; or may be accomplished by alternative molecular cloning techniques which are well known in the art, such as those described by Maniatis et al. in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, New York, 2nd edition, 1989.

In another aspect, the invention provides a recombinant expression vector comprising the nucleotide sequence of a $GABA_A$ receptor subunit together with additional sequences capable of directing the synthesis of the said $GABA_A$ receptor subunit in cultures of stably co-transfected eukaryotic cells.

The term "expression vectors" as used herein refers to DNA sequences that are required for the transcription of cloned copies of recombinant DNA sequences or genes and the translation of their mRNAs in an appropriate host. Such vectors can be used to express eukaryotic genes in a variety of hosts such as bacteria, blue-green algae, yeast cells, insect cells, plant cells and animal cells. Specifically designed vectors allow the shuttling of DNA between bacteria-yeast, bacteria-plant or bacteria-animal cells. An appropriately constructed expression vector should contain: an origin of replication for autonomous replication in host cells, selective markers, a limited number of useful restriction enzyme sites, a high copy number, and strong promoters. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and to initiate RNA synthesis. A strong promoter is one which causes mRNAs to be initiated at high frequency. Expression vectors may include, but are not limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses.

The term "cloning vector" as used herein refers to a DNA molecule, usually a small plasmid or bacteriophage DNA capable of self-replication in a host organism, and used to introduce a fragment of foreign DNA into a host cell. The foreign DNA combined with the vector DNA constitutes a recombinant DNA molecule which is derived from recombinant technology. Cloning vectors may include plasmids, bacteriophages, viruses and cosmids.

The recombinant expression vector in accordance with the invention may be prepared by inserting the nucleotide sequence of the chosen $GABA_A$ subunit into a suitable precursor expression vector (hereinafter

referred to as the "precursor vector") using conventional recombinant DNA methodology known from the art. The precursor vector may be obtained commercially, or constructed by standard techniques from known expression vectors. The precursor vector suitably contains a selection marker, typically an antibiotic resistance gene, such as the neomycin or ampicillin resistance gene. The precursor vector preferably contains a neomycin resistance gene, adjacent the SV40 early splicing and polyadenylation region; an ampicillin resistance gene; and an origin of replication, e.g. pBR322 ori. The vector also preferably contains an inducible promoter, such as MMTV-LTR (inducible with dexamethasone) or metallothionin (inducible with zinc), so that transcription can be controlled in the cell line of this invention. This reduces or avoids any problem of toxicity in the cells because of the chloride channel intrinsic to the $GABA_A$ receptor.

One suitable precursor vector is pMAMneo, available from Clontech Laboratories Inc. (Lee et al., Nature, 1981, 294, 228; and Sardet et al., Cell, 1989, 56, 271). Alternatively the precursor vector pMSGneo can be constructed from the vectors pMSG and pSV2neo as described in Example 1 herein.

The recombinant expression vector of the present invention is then produced by cloning the $GABA_A$ receptor subunit cDNA into the above precursor vector. The required receptor subunit cDNA is subcloned from the vector in which it is harboured, and ligated into a restriction enzyme site, e.g. the HindIII site, in the polylinker of the precursor vector, for example pMAMneo or pMSGneo, by standard cloning methodology known from the art, and in particular by techniques analogous to those described in Example 1, step (b) herein. Before

this subcloning, it is often advantageous, in order to improve expression, to modify the end of a subunit cDNA with additional 5' untranslated sequences, for example by modifying the 5' end of the γ_{2L} subunit DNA by addition of 5' untranslated region sequences from the α_1 subunit DNA.

One suitable expression vector of the present invention is illustrated in Fig. 1 of the accompanying drawings, in which R represents the nucleotide sequence of a given alpha, beta or gamma subunit of the $GABA_A$ receptor, and the remainder of the expression vector depicted therein is derived from the precursor vector pMSGneo and constructed as described in accompanying Example 1, steps (a) and (b).

For each cell line of the present invention, three such vectors will be necessary, one containing an α subunit, one containing a β subunit, and the third containing a γ subunit.

Cells are then co-transfected with the desired combination of three expression vectors. There are several commonly used techniques for transfection of eukaryotic cells in vitro. Calcium phosphate precipitation of DNA is most commonly used (Bachetti et al., Proc. Natl. Acad. Sci. USA, 1977, 74, 1590-1594; Maitland et al., Cell, 1977, 14, 133-141), and represents a favoured technique in the context of the present invention.

A small percentage of the host cells takes up the recombinant DNA. In a small percentage of those, the DNA will integrate into the host cell chromosome. Because the neomycin resistance gene will have been incorporated into these host cells, they can be selected by isolating the individual clones which will grow in the presence of neomycin. Each such clone is then tested to

identify those which will produce the receptor. This is achieved by inducing the production, for example with dexamethasone, and then detecting the presence of receptor by means of radioligand binding.

In a further aspect, the present invention provides protein preparations of $GABA_A$ receptor subunit combinations, especially human $GABA_A$ receptor subunit combinations, derived from cultures of stably transfected eukaryotic cells. The invention also provides preparations of membranes containing subunit combinations of the $GABA_A$ receptor, especially human $GABA_A$ receptor subunit combinations, derived from cultures of stably transfected eukaryotic cells. In particular, the protein preparation and membrane preparations according to the invention will suitably contain the $\alpha_1\beta_1\gamma_2$ subunit combinations of the human $GABA_A$ receptor, and will preferably contain a human $GABA_A$ receptor consisting of the $\alpha_1\beta_1\gamma_{2L}$ subunit combinations. In an especially preferred embodiment, the invention provides cell membranes containing a human $GABA_A$ receptor consisting of the $\alpha_1\beta_1\gamma_{2L}$ subunit combinations isolated from stably transfected mouse Ltk^- fibroblast cells.

The cell line, and the membrane preparations therefrom, according to the present invention have utility in screening and design of drugs which act upon the $GABA_A$ receptor, for example benzodiazepines, barbiturates, β-carbolines and neurosteroids. The present invention accordingly provides the use of the cell line described above, and membrane preparations derived therefrom, in screening for and designing medicaments which act upon the $GABA_A$ receptor. Of particular interest in this context are molecules capable of interacting selectively with $GABA_A$ receptors made up of varying subunit combinations. As will be readily

apparent, the cell line in accordance with the present invention, and the membrane preparations derived therefrom, provide ideal systems for the study of structure, pharmacology and function of the various $GABA_A$ receptor subtypes.

The following non-limiting Examples illustrate the present invention.

EXAMPLE 1

PREPARATION OF $\alpha_1\beta_1\gamma_{2L}$ TRANSFECTED CELLS

a) Construction of eukaryotic expression vector pMSGneo

The approx. 2500 base pair HindIII-EcoRI fragment of the vector pMSG (purchased from Pharmacia Biosystems Limited, Milton Keynes, United Kingdom), containing the gpt structural gene and SV40 polyadenylation signals was replaced by the approx. 2800 base pair HindIII-EcoRI fragment of pSV2neo (Southern, P.J. and Berg, P.J., Molecular and Applied Genetics, 1, 327-341, 1982) containing the neomycin resistance gene Neo^r and SV40 polyadenylation signals. The EcoRI and HindIII sites were then removed by restriction digesting, blunt ending with klenow polymerase, and religating. EcoRI and HindIII cloning sites were then inserted at the XhoI and SmaI sites of the polylinker by conventional techniques using EcoRI and HindIII linkers.

b) Cloning of subunit cDNAs into pMSGneo

Bovine α_1 and β_1 $GABA_A$ receptor cDNAs were obtained from the Molecular Neurobiology Unit, MRC Centre, Hills Road, Cambridge (Scholfield, P. et al. Nature, 328, 221-227, 1987). Bovine γ_2 cDNA was cloned

by the method of Whiting, P. et al. (Proc. Natl. Acad. Sci. USA, 87, 9966-9970, 1990). Bovine α_1 was subcloned from pbGRαsense by digestion with EcoRI, blunt ending the DNA with klenow polymerase, addition of HindIII linkers by ligation, digestion with HindIII and ligation into the HindIII site of pMSGneo. Bovine β_1 was subcloned from pbGRβsense by restriction digestion with EcoRI (partial digestion), klenow polymerase blunt ending, ligation of HindIII linkers, restriction digestion with HindIII and ligation into HindIII site of pMSGneo. Before subcloning into pMSGneo, the bovine γ_2 cDNA was modified from the published sequence as follows. The 5' untranslated region of the bovine α_1 cDNA (bases 60-200 of the published sequence) was added to the 5' end of the published γ_2 sequence by amplifying the α_1 untranslated region using polymerase chain reaction, and then subcloning the product into the 5' BamHI (site in the polylinker of the Bluescript* SK$^-$ cloning vector; Bluescript vector purchased from Stratagene, San Diego, U.S.A.) HindIII sites of the γ_2 cDNA. The modified γ_2 cDNA was then subcloned into pMSGneo by digestion with XbaI (site in the polylinker of the cloning vector), blunt ending with klenow polymerase, ligation of XhoI linkers, digestion with XhoI (site in the polylinker of the cloning vector), and ligation into XhoI site of pMSGneo.

c) Co-transfection of mouse Ltk$^-$ cells

Ltk$^-$ cells were obtained from the Salk Institute for Biological Studies, San Diego, California. Cells were grown at 37°C, 5-8% CO_2, in Modified Eagles Medium containing penicillin, streptomycin and 10% fetal calf serum. The expression vector harbouring the $GABA_A$ receptor subunit DNAs for co-transfection was prepared by

* Trademark

a standard protocol (Chen, C. and Okayama, H., BioTechniques, 6, 632-638, 1988). For co-transfection, Ltk$^-$ cells were plated in dishes (approx. 2x10^5 cells/dish) and grown overnight. The transfection was performed by calcium phosphate precipitation using a kit (purchased from 5 Prime -> 3 Prime Products, Westchester, Pennsylvania). Co-transfection was performed according to manufacturers' instructions, using 5μg of each subunit DNA construct per 10cm dish of cells. After 2 days in culture the cells were divided 1:8 into culture medium containing 1mg/ml neomycin [Geneticin* (obtainable from Gibco BRL, Paisley, Scotland, U.K.)]. After a further week the concentration was increased to 1.5mg/ml, and then 2mg/ml 1 week after that. Resistant clones of cells were isolated and subcloned using cloning cylinders. Subclones were analysed using radioligand binding: subclones were grown in 10cm culture dishes, and when confluent changed into culture medium containing 1μM dexamethasone (obtainable from Sigma Chemical Company, Poole, Dorset, United Kingdom). 3-5 days later the cells were harvested, membranes prepared and used for radioligand binding (see Example 2, step (a) below) using the benzodiazepine antagonist ^{3}H Ro15-1788 (obtained from New England Nuclear, Du Pont (U.K.) Ltd, Stevenage, United Kingdom). The clone expressing the highest amount of ^{3}H Ro15-1788 binding was subcloned from a single cell by limiting dilution. The resultant clonal population of cells described below is referred to as population A.

* Trademark

EXAMPLE 2

CHARACTERIZATION OF $\alpha_1\beta_1\gamma_{2L}$ TRANSFECTED CELLS

a) Radioligand binding

The nature of the recombinant $\alpha_1\beta_1\gamma_{2L}$ $GABA_A$ receptors prepared as described in Example 1 was addressed by characterization of the benzodiazepine (BZ) binding pharmacology, using the BZ antagonist ^3H Ro15-1788. For radioligand binding assays, cells which had been induced by culture in dexamethasone containing medium for 3-5 days were scraped off into 50mM Tris, pH7.5, 100mM NaCl in the form of Tris buffered saline (TBS) and pelleted (20,000rpm, Sorvall RC5C centrifuge). The cell pellet was resuspended in 50mM Tris, pH7.5, homogenised using an Ultra-Turrax* homogeniser and then pelleted as above. This was repeated once more, and the cells then resuspended in TBS (0.4ml per original 10cm dish of cells). Radioligand binding was performed in 0.1ml final volume TBS, containing 5-15 fmols of ^3H Ro15-1788 binding sites. After 1 hour incubation on ice the membranes were harvested onto filters using a Brandel cell harvester, washed with cold TBS, and bound radioactivity determined by scintillation counting. The recombinant $\alpha_1\beta_1\gamma_{2L}$ receptors bound ^3H Ro15-1788 with high affinity (K_D 0.4nM), at levels of up to 200fmols/10cm dish of cells. No binding was seen to either untransfected Ltk^- cells, or population A cells which had not been induced by addition of dexamethasone to the culture medium, confirming that the ^3H Ro15-1788 was binding to recombinant $\alpha_1\beta_1\gamma_2$ $GABA_A$ receptors. The ^3H Ro15-1788 binding was inhibited by flunitrazepam, CL218872, FG8205, βCCM, zolpidem and Ro15-4513,

* Trademark

confirming the BZ pharmacology of the recombinant receptor. Since it is established that only $GABA_A$ receptors containing an α, a β and a γ subunit exhibit BZ binding (Pritchett, D. et al., Nature, 338, 582-585, 1989) these data confirm the nature of the recombinant $\alpha_1\beta_1\gamma_2$ $GABA_A$ receptors expressed by population A cells.

b) Electrophysiology

The nature of the $GABA_A$ receptor expressed by population A cells has been extensively characterised by electrophysiological techniques, using whole cell patch clamp. Only cells induced by culture in the presence of dexamethasone showed responses to GABA. Concentration response curves to GABA gave a log EC_{50} of 5.2, and a Hill coefficient of 1.9. The response to GABA was potentiated by BZs flunitrazepam and CL218872, by the barbiturate pentobarbitone, and by the steroid alphaxalone. The response to GABA was antagonised by both bicuculline and picrotoxin. All these electrophysiological data confirm that the recombinant $GABA_A$ receptor expressed by population A cells has all of the properties expected of a bona fide $GABA_A$ receptor.

EXAMPLE 3

ISOLATION AND SEQUENCING OF cDNAS ENCODING HUMAN $GABA_A$ RECEPTOR α_2, α_3, α_5, α_6 & β_2 SUBUNITS

a) cDNA libraries

cDNAs were cloned from human foetal brain (α_2, α_3), hippocampal (α_5, β_2) and cerebellum (α_6) lambda bacteriophage cDNA libraries. All cDNA libraries were

constructed in the lambdaZAP vector, and were purchased from Stratagene (San Diego, California). For screening, the cDNA libraries were plated according to the manufacturer's instructions, at 40,000 pfu per 137 mm plate. Filter lifts were taken using Hybond*N filters (Amersham) according to the manufacturer's instructions.

b) Isolation of cDNA encoding human α_2 subunit

A bovine α_2 cDNA (obtained from E. Barnard, Molecular Neurobiology, University of Cambridge, Hills Road, Cambridge; Levitan et al., Nature, 1988, 335, 76) was labelled to high specific activity (>1.10^9 cpm/μg) with ^{32}P by random priming and used as a probe. Library filters (8 replica filters) were prehybridised for 3-6 hours at 42°C in 5x SSPE (1x SSPE is 0.18M NaCl, 0.01M Na_3PO_4 [pH7.4], 1mM EDTA), 5x Denhardt's solution, 100 μg/ml salmon sperm DNA, 0.1% sodium dodecyl sulphate (SDS), 30% formamide. Hybridisation was performed in the same buffer for 18 hours at 42°C, including 0.5-1.10^6 cpm ^{32}P-labelled probe per ml of hybridisation buffer. Filters were washed at 55°C in 5x SSPE (2x 15 minutes) and 1x SSPE (2x 15 minutes) and exposed to Kodak XAR*film for 1-3 days. Positive clones were plaque purified using standard techniques, and the Bluescript plasmid (Stratagene) "rescued" according to manufacturer's instructions. cDNA clones were sequenced on both strands by standard techniques using Sequenase*II enzyme (United States Biochemicals). The nucleotide sequence of the cDNA encoding the human $GABA_A$ receptor α_2 subunit, together with the deduced amino acid sequence corresponding thereto, is shown in Fig. 2 of the accompanying drawings.

* Trademark

c) Isolation of cDNA encoding human α_3 subunit

A bovine α_3 cDNA (obtained from E. Barnard, Molecular Neurobiology, University of Cambridge, Hills Road, Cambridge; Levitan et al., Nature, 1988, 335, 76) was labelled to high specific activity with ^{32}P by random priming and used as a probe. Library filters were prehybridised for 3-6 hours at 55°C in 5x SSPE, 5x Denhardt's solution, 0.1% SDS, 100 μg/ml salmon sperm DNA, and hybridised for 18 hours, 55°C in the same buffer, containing 0.5-1x 10^6 cpm/ml of ^{32}P-labelled bovine α_3 cDNA as probe. Filters were washed and exposed to X-ray film as described above; cDNA clones were rescued and sequenced as described above. The longest α_3 cDNA clone was missing in approximately 100 bp of the 5' end of the coding region. This was obtained by PCR using as primers an oligonucleotide "anchor" primer derived from the T7 primer sequence of Bluescript vector (5'AGCGCGCGTAATACGACTCACTATAGGGCGAA3') and an oligonucleotide derived from sequence near the 5' end of the truncated α_3 cDNA, containing an internal Hpal site (5'CAGCATGAATTGTTAACCTCATTGTA3'). Oligonucleotides were synthesised on an Applied Biosystems 380B synthesiser. PCR was performed as described above, and a 300bp PCR product obtained which was double digested with Hpal and Kpnl and subcloned into the similarly cut truncated α_3 cDNA to yield a full length human α_3 cDNA. The cDNA was sequenced on both strands as described above. The nucleotide sequence of the cDNA encoding the human $GABA_A$ receptor α_3 subunit, together with the deduced amino acid sequence corresponding thereto, is shown in Fig. 3 of the accompanying drawings.

d) Isolation of cDNA encoding human α_5 subunit

A rat α_5 cDNA obtained by polymerase chain reaction (PCR) was used as a probe to screen the cDNA library. For PCR, sequences of the oligonucleotide primers were taken from the published α_5 sequences (Khrestchatisky et al., Neuron, 1989, 3, 745) and incorporated a Hind III site for subcloning purposes: 5' ATTATTCAAGCTTGCCATGGACAATGGAATGCTC3' (bp114-148); 5'GGTTTCCAGCTTACTTTGGAGAGGTAGC3' (bp1507-1535). PCR and subcloning of the PCR product into Bluescript SK-vector (Stratagene) for analysis was performed as described elsewhere (Whiting et al., Proc. Natl. Acad. Sci. USA, 1990, 87, 9966) except that rat brain cDNA was used as template. The rat α_5 cDNA was labelled with ^{32}P and used to screen the human hippocampal cDNA library, and positive α_5 clones rescued and sequenced as described for α_2 above. The nucleotide sequence of the cDNA encoding the human $GABA_A$ receptor α_5 subunit, together with the deduced amino acid sequence corresponding thereto, is shown in Fig. 4 of the accompanying drawings.

e) Isolation of cDNA encoding human α_6 subunit

A rat α_6 cDNA obtained by PCR was used as a probe to screen the cDNA library. PCR was performed as described above for α_5, using oligonucleotide primers derived from the published rat α_6 sequence (Luddens et al., Nature, 1990, 346, 648) incorporating an EcoRI site for subcloning purposes: 5'GAGGAAGAATTCAGGAGGGTGACCT3' (bp48-72); 5'GAAAATAACGAATTCCAGTGTCCAGCTTT3' (bp1376-1404). The rat α_6 cDNA clone isolated by PCR was labelled with ^{32}P and used to screen a human cerebellum cDNA library, as described above for α_2. Positive α_6 clones were purified, rescued and sequenced as described above. None of the cDNAs contained a complete coding

region. To obtain a full length cDNA 3 clones were joined together using convenient restriction sites. The nucleotide sequence of the cDNA encoding the human $GABA_A$ receptor α_6 subunit, together with the deduced amino acid sequence corresponding thereto, is shown in Fig. 5 of the accompanying drawings.

f) Isolation of cDNA encoding human β_2 subunit

Human β_2 cDNA was isolated using as a probe a short human β_2 cDNA obtained by PCR. PCR was performed as described above (except that the human cerebellum cDNA library was used as template), using oligonucleotide primers derived from the published rat β_2 sequence (Ymer et al., EMBO J., 1989, 8, 1665), incorporating EcoRI sites for subcloning purposes: 5' CAAAAGAATTCAGCTGAGAAAGCTGCTAATGC3' (bp1088-1119); 5' TCAGGCGAATTCTCTTTTGTGCCACATGTCGTTC3' (bp1331-1364). The human β_2 clone obtained by PCR was radiolabelled with ^{32}P and used to screen a human hippocampal cDNA library, as described above for α_2. The largest cDNA clone obtained lacked the 5' 500bp of the coding region of the β_2 subunit. This was obtained by PCR using as primers an oligonucleotide "anchor" primer derived from the T7 primer sequence of the Bluescript vector (5' AGCGCGCGTAATACGACTCACTATAGGGCGAA3'), and an oligonucleotide derived from sequence near the 5' end of the truncated β_2 cDNA, containing a Kpnl site (5' CATCCAGTGGGTACCTCCTTAGGT3'). PCR was performed as described above, and a 700bp PCR product obtained which was digested with kpnl and subcloned into the truncated cDNA clone (also Kpnl digested) to yield a full length human β_2 cDNA. The nucleotide sequence of the cDNA encoding the human $GABA_A$ receptor β_2 subunit, together

with the deduced amino acid sequence corresponding thereto, is shown in Fig. 6 of the accompanying drawings.

WHAT IS CLAIMED:

1. A stably co-transfected rodent fibroblast cell line capable of expressing a human $GABA_A$ receptor, which receptor comprises at least one alpha, at least one beta and at least one gamma subunit.

2. A stably co-transfected rodent fibroblast cell line as claimed in claim 1 wherein the rodent fibroblast cell line is a mouse Ltk^- cell line.

3. A stably co-transfected cell line as claimed in claim 1 or claim 2 wherein the alpha subunit is the α_2 subunit of the human $GABA_A$ receptor encoded by a cDNA molecule comprising the sequence depicted in Figure 2 herein.

4. A stably co-transfected cell line as claimed in claim 1 or claim 2 wherein the alpha subunit is the α_3 subunit of the human $GABA_A$ receptor encoded by a cDNA molecule comprising the sequence depicted in Figure 3 herein.

5. A stably co-transfected cell line as claimed in claim 1 or claim 2 wherein the alpha subunit is the α_5 subunit of the human $GABA_A$ receptor encoded by a cDNA molecule comprising the sequence depicted in Figure 4 herein.

6. A stably co-transfected cell line as claimed in claim 1 or claim 2 wherein the alpha subunit is the α_6 subunit of the human $GABA_A$ receptor encoded by a cDNA molecule comprising the sequence depicted in Figure 5 herein.

7. A stably co-transfected cell line as claimed in claim 1 or claim 2 wherein the beta subunit is the β_2 subunit of the human $GABA_A$ receptor encoded by a cDNA molecule comprising the sequence depicted in Figure 6 herein.

8. A DNA molecule encoding the α_2 subunit of the human $GABA_A$ receptor comprising all or a portion of the sequence depicted in Figure 2 herein, or a modified human sequence.

9. A DNA molecule encoding the α_3 subunit of the human $GABA_A$ receptor comprising all or a portion of the sequence depicted in Figure 3 herein, or a modified human sequence.

10. A DNA molecule encoding the α_5 subunit of the human $GABA_A$ receptor comprising all or a portion of the sequence depicted in Figure 4 herein, or a modified human sequence.

11. A DNA molecule encoding the α_6 subunit of the human $GABA_A$ receptor comprising all or a portion of the sequence depicted in Figure 5 herein, or a modified human sequence.

12. A DNA molecule encoding the β_2 subunit of the human $GABA_A$ receptor comprising all or a portion of the sequence depicted in Figure 6 herein, or a modified human sequence.

13. A protein preparation of human $GABA_A$ receptor subunit combinations derived from a culture of a stably co-transfected rodent fibroblast cell line.

14. A membrane preparation containing subunit combinations of the human $GABA_A$ receptor derived from a culture of a stably co-transfected rodent fibroblast cell line.

15. A preparation as claimed in claim 13 or claim 14 wherein said rodent fibroblast cell line is a mouse Ltk^- cell line.

2109193

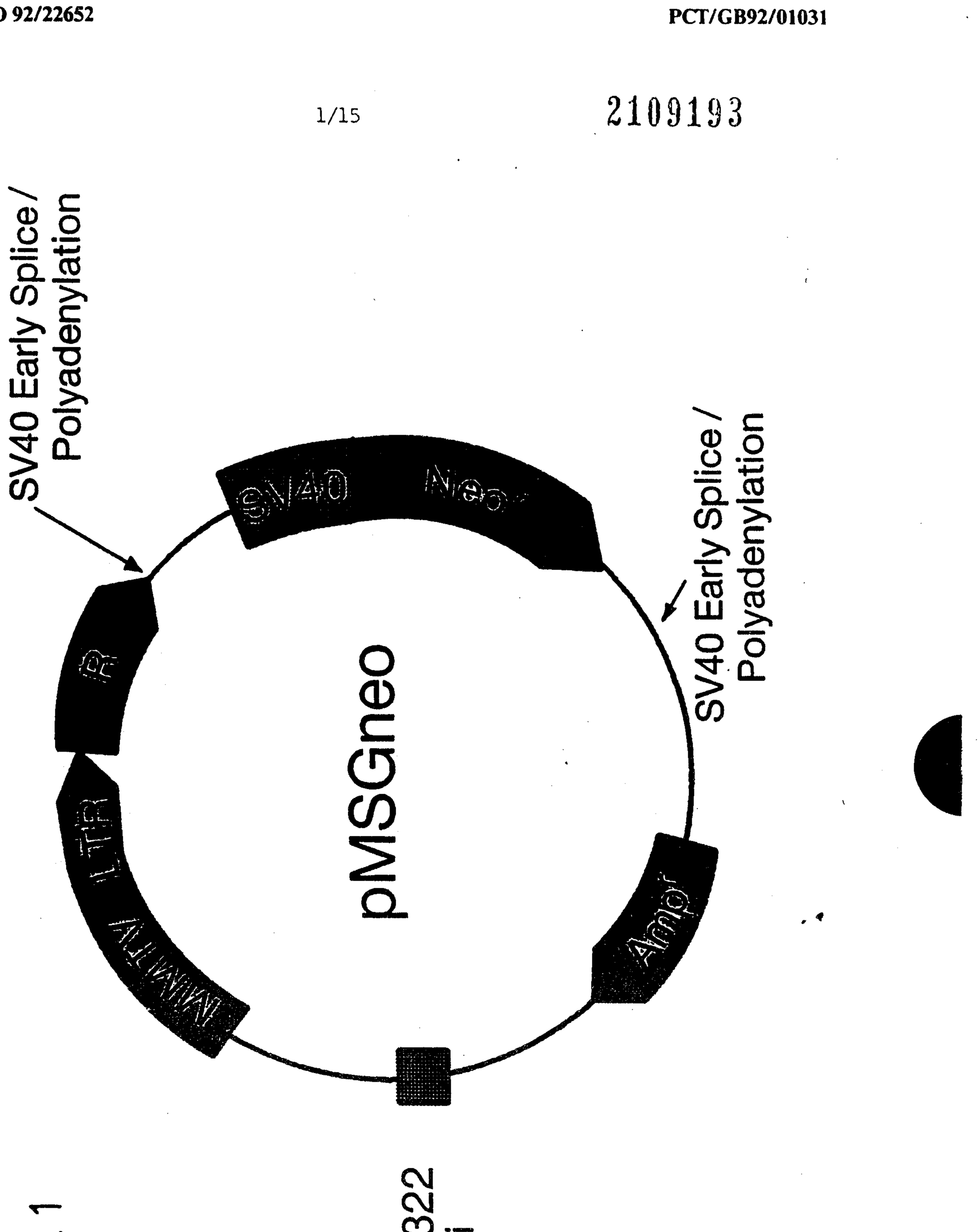


FIG. 1

2109193

FIGURE 2

```
        10          20          30          40          50          60          70
CCTAGCGCTC CTCTCCGGCT TCCACCAGCC CATCGCTCCA CGCTCTCTTG GCTGCTGCAG TCTCGGTCTC

        80          90         100         110         120         130         140
TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCCCAA

      150         160         170         180         190         200         210
GTTTCCTATC TCGTCAAGAT CAGGGCAAAA GAAGAAAACA CCGAATTCTG CTTGCCGTTT CAGAGCGGCG

           219         228         237         246         255         264
       >
GTG ATG AAG ACA AAA TTG AAC ATC TAC AAC ATC GAG TTC CTG CTT TTT GTT TTC
    MET Lys Thr Lys Leu Asn Ile Tyr Asn Ile Glu Phe Leu Leu Phe Val Phe

           273         282         291         300         309         318
TTG GTG TGG GAC CCT GCC AGG TTG GTG CTG GCT AAC ATC CAA GAA GAT GAG GCT
Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile Gln Glu Asp Glu Ala

           327         336         345         354         363         372
AAA AAT AAC ATT ACC ATC TTT ACG AGA ATT CTT GAC AGA CTT CTG GAT GGT TAC
Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile Leu Asp Arg Leu Leu Asp Gly Tyr

           381         390         399         408         417         426
GAT AAT CGG CTT AGA CCA GGA CTG GGA GAC AGT ATT ACT GAA GTC TTC ACT AAC
Asp Asn Arg Leu Arg Pro Gly Leu Gly Asp Ser Ile Thr Glu Val Phe Thr Asn

           435         444         453         462         471         480
ATC TAC GTG ACC AGT TTT GGC CCT GTC TCA GAT ACA GAT ATG GAA TAT ACA ATT
Ile Tyr Val Thr Ser Phe Gly Pro Val Ser Asp Thr Asp MET Glu Tyr Thr Ile

           489         498         507         516         525         534
GAT GTT TTC TTT CGA CAA AAA TGG AAA GAT GAA CGT TTA AAA TTT AAA GGT CCT
Asp Val Phe Phe Arg Gln Lys Trp Lys Asp Glu Arg Leu Lys Phe Lys Gly Pro

           543         552         561         570         579         588
ATG AAT ATC CTT CGA CTA AAC AAT TTA ATG GCT AGC AAA ATC TGG ACT CCA GAT
MET Asn Ile Leu Arg Leu Asn Asn Leu MET Ala Ser Lys Ile Trp Thr Pro Asp

           597         606         615         624         633         642
ACC TTT TTT CAC AAT GGG AAG AAA TCA GTA GCT CAT AAT ATG ACA ATG CCA AAT
Thr Phe Phe His Asn Gly Lys Lys Ser Val Ala His Asn MET Thr MET Pro Asn

           651         660         669         678         687         696
AAG TTG CTT CGA ATT CAG GAT GAT GGG ACT CTG CTG TAT ACC ATG AGG CTT ACA
Lys Leu Leu Arg Ile Gln Asp Asp Gly Thr Leu Leu Tyr Thr MET Arg Leu Thr

           705         714         723         732         741         750
GTT CAA GCT GAA TGC CCA ATG CAC TTG GAG GAT TTC CCA ATG GAT GCT CAT TCA
Val Gln Ala Glu Cys Pro MET His Leu Glu Asp Phe Pro MET Asp Ala His Ser
```

FIGURE 2 (CONTINUED)

```
      759         768         777         786         795         804
TGT CCT CTG AAA TTT GGC AGC TAT GCA TAT ACA ACT TCA GAG GTC ACT TAT ATT
Cys Pro Leu Lys Phe Gly Ser Tyr Ala Tyr Thr Thr Ser Glu Val Thr Tyr Ile

      813         822         831         840         849         858
TGG ACT TAC AAT GCA TCT GAT TCA GTA CAG GTT GCT CCT GAT GGC TCT AGG TTA
Trp Thr Tyr Asn Ala Ser Asp Ser Val Gln Val Ala Pro Asp Gly Ser Arg Leu

      867         876         885         894         903         912
AAT CAA TAT GAC CTG CTG GGC CAA TCA ATC GGA AAG GAG ACA ATT AAA TCC AGT
Asn Gln Tyr Asp Leu Leu Gly Gln Ser Ile Gly Lys Glu Thr Ile Lys Ser Ser

      921         930         939         948         957         966
ACA GGT GAA TAT ACT GTA ATG ACA GCT CAT TTC CAC CTG AAA AGA AAA ATT GGG
Thr Gly Glu Tyr Thr Val MET Thr Ala His Phe His Leu Lys Arg Lys Ile Gly

      975         984         993        1002        1011        1020
TAT TTT GTG ATT CAA ACC TAT CTG CCT TGC ATC ATG ACT GTC ATT CTC TCC CAA
Tyr Phe Val Ile Gln Thr Tyr Leu Pro Cys Ile MET Thr Val Ile Leu Ser Gln

     1029        1038        1047        1056        1065        1074
GTT TCA TTC TGG CTT AAC AGA GAA TCT GTG CCT GCA AGA ACT GTG TTT GGA GTA
Val Ser Phe Trp Leu Asn Arg Glu Ser Val Pro Ala Arg Thr Val Phe Gly Val

     1083        1092        1101        1110        1119        1128
ACA ACT GTC CTA ACA ATG ACA ACT CTA AGC ATC AGT GCT CGG AAT TCT CTC CCC
Thr Thr Val Leu Thr MET Thr Thr Leu Ser Ile Ser Ala Arg Asn Ser Leu Pro

     1137        1146        1155        1164        1173        1182
AAA GTG GCT TAT GCA ACT GCC ATG GAC TGG TTT ATT GCT GTT TGT TAT GCA TTT
Lys Val Ala Tyr Ala Thr Ala MET Asp Trp Phe Ile Ala Val Cys Tyr Ala Phe

     1191        1200        1209        1218        1227        1236
GTG TTC TCT GCC CTA ATT GAA TTT GCA ACT GTT AAT TAC TTC ACC AAA AGA GGA
Val Phe Ser Ala Leu Ile Glu Phe Ala Thr Val Asn Tyr Phe Thr Lys Arg Gly

     1245        1254        1263        1272        1281        1290
TGG ACT TGG GAT GGG AAG AGT GTA GTA AAT GAC AAG AAA AAA GAA AAG GCT TCC
Trp Thr Trp Asp Gly Lys Ser Val Val Asn Asp Lys Lys Lys Glu Lys Ala Ser

     1299        1308        1317        1326        1335        1344
GTT ATG ATA CAG AAC AAC GCT TAT GCA GTG GCT GTT GCC AAT TAT GCC CCG AAT
Val MET Ile Gln Asn Asn Ala Tyr Ala Val Ala Val Ala Asn Tyr Ala Pro Asn

     1353        1362        1371        1380        1389        1398
CTT TCA AAA GAT CCA GTT CTC TCC ACC ATC TCC AAG AGT GCA ACC ACG CCA GAA
Leu Ser Lys Asp Pro Val Leu Ser Thr Ile Ser Lys Ser Ala Thr Thr Pro Glu

     1407        1416        1425        1434        1443        1452
CCC AAC AAG AAG CCA GAA AAC AAG CCA GCT GAA GCA AAG AAA ACT TTC AAC AGT
Pro Asn Lys Lys Pro Glu Asn Lys Pro Ala Glu Ala Lys Lys Thr Phe Asn Ser

     1461        1470        1479        1488        1497        1506
```

FIGURE 2 (CONTINUED)


```
GTT AGC AAA ATT GAC AGA ATG TCC AGA ATA GTT TTT CCA GTT TTG TTT GGT ACC
Val Ser Lys Ile Asp Arg MET Ser Arg Ile Val Phe Pro Val Leu Phe Gly Thr

      1515        1524        1533        1542        1551        1560

TTT AAT TTA GTT TAC TGG GCT ACA TAT TTA AAC AGA GAA CCT GTA TTA GGG GTC
Phe Asn Leu Val Tyr Trp Ala Thr Tyr Leu Asn Arg Glu Pro Val Leu Gly Val

      1569        1579        1589        1599        1609        1619        1629
                   >
AGT CCT TGA ATTGAGACCC ATGTTATCTT TGGGATGTAT AGCAACATTA AATTTGGTTT GTTTTGCTAT
Ser Pro  .

      1639        1649        1659        1669        1679        1689        1699

GTACAGTCTG ACTAATAACT GCTAATTTGT GATCCAACAT GTACAGTATG TATATAGTGA CATAGCTTAC

      1709        1719        1729        1739        1749        1759        1769

CAGTAGACCT TTAATGGAGA CATGCATTTG CTAACTCATG GAACTGCAGA CAGAAAGCAC TCCATGCGAA

      1779        1789        1799        1809        1819        1829        1839

AACAGCCATT GCCTTTTTTA AAGATTTACC CTAGGACCTG ATTTAAAGTG AATTTCAAGT GACCTGATTA

      1849        1859        1869        1879        1889        1899        1909

ATTTCCTATT CTTCCAAATG AGATGAAAAT GGGGATCCTG TACAACCCTT TGTGGACCCT TTTGGTTTAG

      1919        1929        1939        1949        1959        1969        1979

CTCTTAAGTA GGGGTATTTT CTACTGTTGC TTAATTATGA TGGAAGATAA CATTGTCATT CCTAGATGAA

      1989        1999        2009        2019        2029        2039        2049

TCCTTTGAAG TAACAAACAT TGTATCTGAC ATCAGCTCTG TTCATGAGTG CTCAGAGTCC CTGCTAATGT

      2059        2069        2079        2089        2099        2109        2119

AATTGGAAGC TTGGTACACA TAAGAAAAAC TAGAGATTTG AAATCTAGCT ATGAATTACT CTATATAGTA

      2129        2139        2149        2159        2169        2179        2189

TCTATAGCCA TGTACATATT ACAGCATGAC AAGCTCGAAA TAATTATGAG TCAGCCCGAA AGATGTTAAT
```

FIGURE 3

2109193

```
        10          20          30          40          50          60          70
GAATTCCCTT GTTTCAGTTC ATTCATCCTT CTCTCCTTTC CGCTCAGACT GTAGAGCTCG GTCTCTCCAA

        80          89          98         107         116         125
GTTTGTGCCT AAGAAG ATG ATA ATC ACA CAA ACA AGT CAC TGT TAC ATG ACC AGC
                  MET Ile Ile Thr Gln Thr Ser His Cys Tyr MET Thr Ser

       134         143         152         161         170         179
CTT GGG ATT CTT TTC CTG ATT AAT ATT CTC CCT GGA ACC ACT GGT CAA GGG GAA
Leu Gly Ile Leu Phe Leu Ile Asn Ile Leu Pro Gly Thr Thr Gly Gln Gly Glu

       188         197         206         215         224         233
TCA AGA CGA CAA GAA CCC GGG GAC TTT GTG AAG CAG GAC ATT GGC GGG CTG TCT
Ser Arg Arg Gln Glu Pro Gly Asp Phe Val Lys Gln Asp Ile Gly Gly Leu Ser

       242         251         260         269         278         287
CCT AAG CAT GCC CCA GAT ATT CCT GAT GAC AGC ACT GAC AAC ATC ACT ATC TTC
Pro Lys His Ala Pro Asp Ile Pro Asp Asp Ser Thr Asp Asn Ile Thr Ile Phe

       296         305         314         323         332         341
ACC AGA ATC TTG GAT CGT CTT CTG GAC GGC TAT GAC AAC CGG CTG CGA CCT GGG
Thr Arg Ile Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly

       350         359         368         377         386         395
CTT GGA GAT GCA GTG ACT GAA GTG AAG ACT GAC ATC TAC GTG ACC AGT TTT GGC
Leu Gly Asp Ala Val Thr Glu Val Lys Thr Asp Ile Tyr Val Thr Ser Phe Gly

       404         413         422         431         440         449
CCT GTG TCA GAC ACT GAC ATG GAG TAC ACT ATT GAT GTA TTT TTT CGG CAG ACA
Pro Val Ser Asp Thr Asp MET Glu Tyr Thr Ile Asp Val Phe Phe Arg Gln Thr

       458         467         476         485         494         503
TGG CAT GAT GAA AGA CTG AAA TTT GAT GGC CCC ATG AAG ATC CTT CCA CTG AAC
Trp His Asp Glu Arg Leu Lys Phe Asp Gly Pro MET Lys Ile Leu Pro Leu Asn

       512         521         530         539         548         557
AAT CTC CTG GCT AGT AAG ATC TGG ACA CCG GAC ACC TTC TTC CAC AAT GGC AAG
Asn Leu Leu Ala Ser Lys Ile Trp Thr Pro Asp Thr Phe Phe His Asn Gly Lys

       566         575         584         593         602         611
AAA TCA GTG GCT CAT AAC ATG ACC ACG CCC AAC AAG CTG CTC AGA TTG GTG GAC
Lys Ser Val Ala His Asn MET Thr Thr Pro Asn Lys Leu Leu Arg Leu Val Asp

       620         629         638         647         656         665
AAC GGA ACC CTC CTC TAT ACA ATG AGG TTA ACA ATT CAT GCT GAG TGT CCC ATG
Asn Gly Thr Leu Leu Tyr Thr MET Arg Leu Thr Ile His Ala Glu Cys Pro MET

       674         683         692         701         710         719
CAT TTG GAA GAT TTT CCC ATG GAT GTG CAT GCC TGC CCA CTG AAG TTT GGA AGC
His Leu Glu Asp Phe Pro MET Asp Val His Ala Cys Pro Leu Lys Phe Gly Ser

       728         737         746         755         764         773
```

FIGURE 3 (CONTINUED)

TAT GCC TAT ACA ACA GCT GAA GTG GTT TAT TCT TGG ACT CTC GGA AAG AAC AAA
Tyr Ala Tyr Thr Thr Ala Glu Val Val Tyr Ser Trp Thr Leu Gly Lys Asn Lys

782 791 800 809 818 827

TCC GTG GAA GTG GCA CAG GAT GGT TCT CGC TTG AAC CAG TAT GAC CTT TTG GGC
Ser Val Glu Val Ala Gln Asp Gly Ser Arg Leu Asn Gln Tyr Asp Leu Leu Gly

836 845 854 863 872 881

CAT GTT GTT GGG ACA GAG ATA ATC CGG TCT AGT ACA GGA GAA TAT GTC GTC ATG
His Val Val Gly Thr Glu Ile Ile Arg Ser Ser Thr Gly Glu Tyr Val Val MET

890 899 908 917 926 935

ACA ACC CAC TTC CAT CTC AAG CGA AAA ATT GGC TAC TTT GTG ATC CAG ACC TAC
Thr Thr His Phe His Leu Lys Arg Lys Ile Gly Tyr Phe Val Ile Gln Thr Tyr

944 953 962 971 980 989

TTG CCA TGT ATC ATG ACT GTC ATT CTG TCA CAA GTG TCG TTC TGG CTC AAC AGA
Leu Pro Cys Ile MET Thr Val Ile Leu Ser Gln Val Ser Phe Trp Leu Asn Arg

998 1007 1016 1025 1034 1043

GAG TCT GTT CCT GCC CGT ACA GTC TTT GGT GTC ACC ACT GTG CTT ACC ATG ACC
Glu Ser Val Pro Ala Arg Thr Val Phe Gly Val Thr Thr Val Leu Thr MET Thr

1052 1061 1070 1079 1088 1097

ACC TTG AGT ATC AGT GCC AGA AAT TCC TTA CCT AAA GTG GCA TAT GCG ACG GCC
Thr Leu Ser Ile Ser Ala Arg Asn Ser Leu Pro Lys Val Ala Tyr Ala Thr Ala

1106 1115 1124 1133 1142 1151

ATG GAC TGG TTC ATA GCC GTC TGT TAT GCC TTT GTA TTT TCT GCA CTG ATT GAA
MET Asp Trp Phe Ile Ala Val Cys Tyr Ala Phe Val Phe Ser Ala Leu Ile Glu

1160 1169 1178 1187 1196 1205

TTT GCC ACT GTC AAC TAT TTC ACC AAG CGG AGT TGG GCT TGG GAA GGC AAG AAG
Phe Ala Thr Val Asn Tyr Phe Thr Lys Arg Ser Trp Ala Trp Glu Gly Lys Lys

1214 1223 1232 1241 1250 1259

GTG CCA GAG GCC CTG GAG ATG AAG AAG AAA ACA CCA GCA GCC CCA GCA AAG AAA
Val Pro Glu Ala Leu Glu MET Lys Lys Lys Thr Pro Ala Ala Pro Ala Lys Lys

1268 1277 1286 1295 1304 1313

ACC AGC ACT ACC TTC AAC ATC GTG GGG ACC ACC TAT CCC ATC AAC CTG GCC AAG
Thr Ser Thr Thr Phe Asn Ile Val Gly Thr Thr Tyr Pro Ile Asn Leu Ala Lys

1322 1331 1340 1349 1358 1367

GAC ACT GAA TTT TCC ACC ATC TCC AAG GGC GCT GCT CCC AGT GCC TCC TCA ACC
Asp Thr Glu Phe Ser Thr Ile Ser Lys Gly Ala Ala Pro Ser Ala Ser Ser Thr

1376 1385 1394 1403 1412 1421

CCA ACA ATC ATT GCT TCA CCC AAG GCC ACC TAC GTG CAG GAC AGC CCG ACT GAG
Pro Thr Ile Ile Ala Ser Pro Lys Ala Thr Tyr Val Gln Asp Ser Pro Thr Glu

1430 1439 1448 1457 1466 1475

FIGURE 3 (CONTINUED)


```
ACC AAG ACC TAC AAC AGT GTC AGC AAG GTT GAC AAA ATT TCC CGC ATC ATC TTT
Thr Lys Thr Tyr Asn Ser Val Ser Lys Val Asp Lys Ile Ser Arg Ile Ile Phe

      1484        1493        1502        1511        1520        1529

CCT GTG CTC TTT GCC ATA TTC AAT CTG GTC TAT TGG GCC ACA TAT GTC AAC CGG
Pro Val Leu Phe Ala Ile Phe Asn Leu Val Tyr Trp Ala Thr Tyr Val Asn Arg

      1538        1547        1556        1565        1575        1585

GAG TCA GCT ATC AAG GGC ATG ATC CGC AAA CAG TAG ATAGTGGCAG TGCAGCAACC
Glu Ser Ala Ile Lys Gly MET Ile Arg Lys Gln  .

      1595        1605        1615        1625        1635

AGAGCACTGT ATACCCCGTG AAGCATCCAG GCACCCAAAC CCCGGGGCTC CCC
```

2109193

FIGURE 4

```
        10         20         30         40         50         60         70
GAATTCCCCC CTTGCAGGCC GAGCCGGGGC CCTGCGCCCT CCCCCTCCGC CCAGCTCGGC CAAGGGCGCA

        80         90        100        110        120        130        140
TTTGCTGAGC GTCTGGCGGC CTCTACCGGA GCACCTCTGC AGAGGGCCGA TCCTCCAGCC CAGAGACGAC

       150        160        170        180        190        200        210
ATGTGGCGCT CGGGCGAGTG CCTTGCAGAG AGAGGAGTAG CTTGCTGGCT TTGAACGCGT GGCGTGGCAG

       220        230        240        250        260        270        280
ATATTTCAGA AAGCTTCAAG AACAAGCTGG AGAAGGGAAG AGTTATTCCT CCATATTCAC CTGCTTCAAC

       290        300        309        318        327        336
                   >
TACTATTCTT ATTGGGA ATG GAC AAT GGA ATG TTC TCT GGT TTT ATC ATG ATC AAA
                   MET Asp Asn Gly MET Phe Ser Gly Phe Ile MET Ile Lys

       345        354        363        372        381        390
AAC CTC CTT CTC TTT TGT ATT TCC ATG AAC TTA TCC AGT CAC TTT GGC TTT TCA
Asn Leu Leu Leu Phe Cys Ile Ser MET Asn Leu Ser Ser His Phe Gly Phe Ser

       399        408        417        426        435        444
CAG ATG CCA ACC AGT TCA GTG AAA GAT GAG ACC AAT GAC AAC ATC ACG ATA TTT
Gln MET Pro Thr Ser Ser Val Lys Asp Glu Thr Asn Asp Asn Ile Thr Ile Phe

       453        462        471        480        489        498
ACC AGG ATC TTG GAT GGG CTC TTG GAT GGC TAC GAC AAC AGA CTT CGG CCC GGG
Thr Arg Ile Leu Asp Gly Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly

       507        516        525        534        543        552
CTG GGA GAG CGC ATC ACT CAG GTG AGG ACC GAC ATC TAC GTC ACC AGC TTC GGC
Leu Gly Glu Arg Ile Thr Gln Val Arg Thr Asp Ile Tyr Val Thr Ser Phe Gly

       561        570        579        588        597        606
CCG GTG TCC GAC ACG GAA ATG GAG TAC ACC ATA GAC GTG TTT TTC CGA CAA AGC
Pro Val Ser Asp Thr Glu MET Glu Tyr Thr Ile Asp Val Phe Phe Arg Gln Ser

       615        624        633        642        651        660
TGG AAA GAT GAA AGG CTT CGG TTT AAG GGG CCC ATG CAG CGC CTC CCT CTC AAC
Trp Lys Asp Glu Arg Leu Arg Phe Lys Gly Pro MET Gln Arg Leu Pro Leu Asn

       669        678        687        696        705        714
AAC CTC CTT GCC AGC AAG ATC TGG ACC CCA GAC ACG TTC TTC CAC AAC GGG AAG
Asn Leu Leu Ala Ser Lys Ile Trp Thr Pro Asp Thr Phe Phe His Asn Gly Lys

       723        732        741        750        759        768
AAG TCC ATC GCT CAC AAC ATG ACC ACG CCC AAC AAG CTG CTG CGG CTG GAG GAC
Lys Ser Ile Ala His Asn MET Thr Thr Pro Asn Lys Leu Leu Arg Leu Glu Asp
```

FIGURE 4 (CONTINUED)


```
          777                 786                 795                 804                 813                 822
GAC GGC ACC CTG CTC TAC ACC ATG CGC TTG ACC ATC TCT GCA GAG TGC CCC ATG
Asp Gly Thr Leu Leu Tyr Thr MET Arg Leu Thr Ile Ser Ala Glu Cys Pro MET

          831                 840                 849                 858                 867                 876
CAG CTT GAG GAC TTC CCG ATG GAT GCG CAC GCT TGC CCT CTG AAA TTT GGC AGC
Gln Leu Glu Asp Phe Pro MET Asp Ala His Ala Cys Pro Leu Lys Phe Gly Ser

          885                 894                 903                 912                 921                 930
TAT GCG TAC CCT AAT TCT GAA GTC GTT TAC GTC TGG ACC AAC GGC TCC ACC AAG
Tyr Ala Tyr Pro Asn Ser Glu Val Val Tyr Val Trp Thr Asn Gly Ser Thr Lys

          939                 948                 957                 966                 975                 984
TCG GTG GTG GTG GCG GAA GAT GGC TCC AGA CTG AAC CAG TAC CAC CTG ATG GGG
Ser Val Val Val Ala Glu Asp Gly Ser Arg Leu Asn Gln Tyr His Leu MET Gly

          993                1002                1011                1020                1029                1038
CAG ACG GTG GGC ACT GAG AAC ATC AGC ACC AGC ACA GGC GAA TAC ACA ATC ATG
Gln Thr Val Gly Thr Glu Asn Ile Ser Thr Ser Thr Gly Glu Tyr Thr Ile MET

         1047                1056                1065                1074                1083                1092
ACA GCT CAC TTC CAC CTG AAA AGG AAG ATT GGC TAC TTT GTC ATC CAG ACC TAC
Thr Ala His Phe His Leu Lys Arg Lys Ile Gly Tyr Phe Val Ile Gln Thr Tyr

         1101                1110                1119                1128                1137                1146
CTT CCC TGC ATA ATG ACC GTG ATC TTA TCA CAG GTG TCC TTT TGG CTG AAC CGG
Leu Pro Cys Ile MET Thr Val Ile Leu Ser Gln Val Ser Phe Trp Leu Asn Arg

         1155                1164                1173                1182                1191                1200
GAA TCA GTC CCA GCC AGG ACA GTT TTT GGG GTC ACC ACG GTG CTG ACC ATG ACG
Glu Ser Val Pro Ala Arg Thr Val Phe Gly Val Thr Thr Val Leu Thr MET Thr

         1209                1218                1227                1236                1245                1254
ACC CTC AGC ATC AGC GCC AGG AAC TCT CTG CCC AAA GTG GCC TAC GCC ACC GCC
Thr Leu Ser Ile Ser Ala Arg Asn Ser Leu Pro Lys Val Ala Tyr Ala Thr Ala

         1263                1272                1281                1290                1299                1308
ATG GAC TGG TTC ATA GCT GTG TGC TAT GCC TTC GTC TTC TCG GCG CTG ATA GAG
MET Asp Trp Phe Ile Ala Val Cys Tyr Ala Phe Val Phe Ser Ala Leu Ile Glu

         1317                1326                1335                1344                1353                1362
TTT GCC ACG GTC AAT TAC TTT ACC AAG AGA GGC TGG GCC TGG GAT GGC AAA AAA
Phe Ala Thr Val Asn Tyr Phe Thr Lys Arg Gly Trp Ala Trp Asp Gly Lys Lys

         1371                1380                1389                1398                1407                1416
GCC TTG GAA GCA GCC AAG ATC AAG AAA AAG CGT GAA GTC ATA CTA AAT AAG TCA
Ala Leu Glu Ala Ala Lys Ile Lys Lys Lys Arg Glu Val Ile Leu Asn Lys Ser

         1425                1434                1443                1452                1461                1470
ACA AAC GCT TTT ACA ACT GGG AAG ATG TCT CAC CCC CCA AAC ATT CCG AAG GAA
Thr Asn Ala Phe Thr Thr Gly Lys MET Ser His Pro Pro Asn Ile Pro Lys Glu

         1479                1488                1497                1506                1515                1524
```

FIGURE 4 (CONTINUED)

2109193

CAG ACC CCA GCA GGG ACG TCG AAT ACA ACC TCA GTC TCA GTA AAA CCC TCT GAA
Gln Thr Pro Ala Gly Thr Ser Asn Thr Thr Ser Val Ser Val Lys Pro Ser Glu

1533 1542 1551 1560 1569 1578

GAG AAG ACT TCT GAA AGC AAA AAG ACT TAC AAC AGT ATC AGC AAA ATT GAC AAA
Glu Lys Thr Ser Glu Ser Lys Lys Thr Tyr Asn Ser Ile Ser Lys Ile Asp Lys

1587 1596 1605 1614 1623 1632

ATG TCC CGA ATC GTA TTC CCA GTC TTG TTC GGC ACT TTC AAC TTA GTT TAC TGG
MET Ser Arg Ile Val Phe Pro Val Leu Phe Gly Thr Phe Asn Leu Val Tyr Trp

1641 1650 1659 1668 1677 1693

GCA ACG TAT TTG AAT AGG GAG CCG GTG ATA AAA GGA GCC GCC TCT CCA AAA >TAACCGGCCA
Ala Thr Tyr Leu Asn Arg Glu Pro Val Ile Lys Gly Ala Ala Ser Pro Lys

1703 1713 1723 1733 1743 1753 1763

CACTCCCAAA CTCCAAGACA GCCATACTTC CAGCGAAATG GTACCAAGGA GAGGTTTTGC TCACAGGGAC

1773 1783 1793 1803 1813 1823 1833

TCTCCATATG TGAGCACTAT CTTTCAGGAA ATTTTTGCAT GTTTAATAAT ATGTACAAAT AATATTGCCT

1843 1853 1863 1873 1883 1893 1903

TGATGTTTCT ATATGTAACT TCAGATGTTT CCAAGATGTC CCATTGATAA TTCGAGCAAA CAACTTTCTG

1913 1923 1933 1943 1953 1963 1973

GAAAAACAGG ATACGATGAC TGACACTCAG ATGCCCAGTA TCATACGTTG ATAGTTTACA AACAAGATAC

1983 1993 2003 2013 2023 2033 2043

GTATATTTTT AACTGCTTCA AGTGTTACCT AACAATGTTT TTTATACTTC AAATGTCATT TCATACAAAT

2053 2063 2073 2083 2093 2103 2113

TTTCCCAGTG AATAAATATT TTAGGAAACT CTCCATGATT ATTAGAAGAC CAACTATATT GCGAGAAACA

2123 2133 2143 2153 2163 2173 2183

GAGATCATAA AGAGCACGTT TTCCATTATG AGGAAACTTG GACATTTATG TACAAAATGA ATTGCCTTTG

2193 2203 2213 2223 2233 2243 2253

ATAATTCTTA CTGTTCTGAA ATTAGGAAAG TACTTGCATG ATCTTACACG AAGAAATAGA ATAGGCAAAC

2263 2273 2283 2293 2303

TTTTATGTAG GCAGATTAAT AACAGAAATA CATCATATGT TAGATACACA AAATATT

FIGURE 5

2109193

```
        10          20          29          38          47          56
                                >
AATTCTGCAT TTCAGTGCAC TGCAGG ATG GCG TCA TCT CTG CCC TGG CTG TGC ATT
                             MET Ala Ser Ser Leu Pro Trp Leu Cys Ile

        65          74          83          92          101         110
ATT CTG TGG CTA GAA AAT GCC CTA GGG AAA CTC GAA GTT GAA GGC AAC TTC TAC
Ile Leu Trp Leu Glu Asn Ala Leu Gly Lys Leu Glu Val Glu Gly Asn Phe Tyr

        119         128         137         146         155         164
TCA GAA AAC GTC AGT CGG ATC CTG GAC AAC TTG CTT GAA GGC TAT GAC AAT CGG
Ser Glu Asn Val Ser Arg Ile Leu Asp Asn Leu Leu Glu Gly Tyr Asp Asn Arg

        173         182         191         200         209         218
CTG CGG CCG GGA TTT GGA GGT GCT GTC ACT GAA GTC AAA ACA GAC ATT TAT GTG
Leu Arg Pro Gly Phe Gly Gly Ala Val Thr Glu Val Lys Thr Asp Ile Tyr Val

        227         236         245         254         263         272
ACC AGT TTT GGG CCC GTG TCA GAT GTG GAG ATG GAG TAT ACG ATG GAT GTT TTT
Thr Ser Phe Gly Pro Val Ser Asp Val Glu MET Glu Tyr Thr MET Asp Val Phe

        281         290         299         308         317         326
TTT CGC CAG ACC TGG ACT GAT GAG AGG TTG AAG TTT GGG GGG CCA ACT GAG ATT
Phe Arg Gln Thr Trp Thr Asp Glu Arg Leu Lys Phe Gly Gly Pro Thr Glu Ile

        335         344         353         362         371         380
CTG AGT CTG AAT AAT TTG ATG GTC AGT AAA ATC TGG ACG CCT GAC ACC TTT TTC
Leu Ser Leu Asn Asn Leu MET Val Ser Lys Ile Trp Thr Pro Asp Thr Phe Phe

        389         398         407         416         425         434
AGA AAT GGT AAA AAG TCC ATT GCT CAC AAC ATG ACA ACT CCT AAT AAA CTC TTC
Arg Asn Gly Lys Lys Ser Ile Ala His Asn MET Thr Thr Pro Asn Lys Leu Phe

        443         452         461         470         479         488
AGA ATA ATG CAG AAT GGA ACC ATT TTA TAC ACC ATG AGG CTT ACC ATC AAT GCT
Arg Ile MET Gln Asn Gly Thr Ile Leu Tyr Thr MET Arg Leu Thr Ile Asn Ala

        497         506         515         524         533         542
GAC TGT CCC ATG AGG CTG GTT AAC TTT CCT ATG GAT GGG CAT GCT TGT CCA CTC
Asp Cys Pro MET Arg Leu Val Asn Phe Pro MET Asp Gly His Ala Cys Pro Leu

        551         560         569         578         587         596
AAG TTT GGG AGC TAT GCT TAT CCC AAA AGT GAA ATC ATA TAT ACG TGG AAA AAA
Lys Phe Gly Ser Tyr Ala Tyr Pro Lys Ser Glu Ile Ile Tyr Thr Trp Lys Lys

        605         614         623         632         641         650
GGA CCA CTT TAC TCA GTA GAA GTC CCA GAA GAA TCT TCA AGC CTT CTC CAG TAT
Gly Pro Leu Tyr Ser Val Glu Val Pro Glu Glu Ser Ser Ser Leu Leu Gln Tyr

        659         668         677         686         695         704
GAT CTG ATT GGA CAA ACA GTA TCT AGT GAG ACA ATT AAA TCT AAC ACA GGT GAA
Asp Leu Ile Gly Gln Thr Val Ser Ser Glu Thr Ile Lys Ser Asn Thr Gly Glu
```

FIGURE 5 (CONTINUED)

```
          713         722         731         740         749         758
TAC GTT ATA ATG ACA GTT TAC TTC CAC TTG CAA AGG AAG ATG GGC TAC TTC ATG
Tyr Val Ile MET Thr Val Tyr Phe His Leu Gln Arg Lys MET Gly Tyr Phe MET

          767         776         785         794         803         812
ATA CAG ATA TAC ACT CCT TGC ATT ATG ACA GTC ATT CTT TCC CAG GTG TCT TTC
Ile Gln Ile Tyr Thr Pro Cys Ile MET Thr Val Ile Leu Ser Gln Val Ser Phe

          821         830         839         848         857         866
TGG ATT AAT AAG GAG TCC GTC CCA GCA AGA ACT GTT CTT GGG ATC ACC ACT GTT
Trp Ile Asn Lys Glu Ser Val Pro Ala Arg Thr Val Leu Gly Ile Thr Thr Val

          875         884         893         902         911         920
TTA ACT ATG ACC ACT TTG AGC ATC AGT GCC CGG CAC TCT TTG CCA AAA GTG TCA
Leu Thr MET Thr Thr Leu Ser Ile Ser Ala Arg His Ser Leu Pro Lys Val Ser

          929         938         947         956         965         974
TAT GCC ACT GCC ATG GAT TGG TTC ATA GCT GTT TGC TTT GCA TTC GTC TTC TCT
Tyr Ala Thr Ala MET Asp Trp Phe Ile Ala Val Cys Phe Ala Phe Val Phe Ser

          983         992        1001        1010        1019        1028
GCT CTT ATC GAG TTC GCA GCT GTC AAC TAC TTT ACC AAT CTT CAG ACA CAG AAG
Ala Leu Ile Glu Phe Ala Ala Val Asn Tyr Phe Thr Asn Leu Gln Thr Gln Lys

         1037        1046        1055        1064        1073        1082
GCG AAA AGG AAG GCA CAG TTT GCA GCC CCA CCC ACA GTG ACA ATA TCA AAA GCT
Ala Lys Arg Lys Ala Gln Phe Ala Ala Pro Pro Thr Val Thr Ile Ser Lys Ala

         1091        1100        1109        1118        1127        1136
ACT GAA CCT TTG GAA GCT GAG ATT GTT TTG CAT CCT GAC TCC AAA TAT CAT CTG
Thr Glu Pro Leu Glu Ala Glu Ile Val Leu His Pro Asp Ser Lys Tyr His Leu

         1145        1154        1163        1172        1181        1190
AAG AAA AGG ATC ACT TCT CTG TCT TTG CCA ATA GTT TCA TCT TCC GAG GCC AAT
Lys Lys Arg Ile Thr Ser Leu Ser Leu Pro Ile Val Ser Ser Ser Glu Ala Asn

         1199        1208        1217        1226        1235        1244
AAA GTG CTC ACG AGA GCG CCC ATC TTA CAA TCA ACA CCT GTC ACA CCC CCA CCA
Lys Val Leu Thr Arg Ala Pro Ile Leu Gln Ser Thr Pro Val Thr Pro Pro Pro

         1253        1262        1271        1280        1289        1298
CTC CCG CCA GCC TTT GGA GGC ACC AGT AAA ATA GAC CAG TAT TCT CGA ATT CTC
Leu Pro Pro Ala Phe Gly Gly Thr Ser Lys Ile Asp Gln Tyr Ser Arg Ile Leu

         1307        1316        1325        1334        1343        1352
TTC CCA GTT GCA TTT GCA GGA TTC AAC CTT GTG TAC TGG GTA GTT TAT CTT TCC
Phe Pro Val Ala Phe Ala Gly Phe Asn Leu Val Tyr Trp Val Val Tyr Leu Ser

         1361        1370        1379        1388        1398        1408
                                             >
AAA GAT ACA ATG GAA GTG AGT AGC AGT GTT GAA TAG CTTTTCCAGG ACAACCTGAA
Lys Asp Thr MET Glu Val Ser Ser Ser Val Glu  .
```

FIGURE 5

```
        10          20          30          40          50          60          70
GAATTCCGCG CGGGGAAGGG AAGAAGAGGA CGAGGTGGCG CAGAGACCGC GGGAGAACAC AGTGCCTCCG

        80          90          100         110         120         130         140
GAGGAAATCT GCTCGGTCCC CGGCAGCCGC GCTTCCCCTT TGATGTTTTG GTACGCCGTG GCCATGCGCC

       150         160         170         180         190         200         210
TCACATTAGA ATTACTGCAC TGGGCAGACT AAGTTGGATC TCCTCTCTTC AGTGAAACCC TCAATTCCAT

       220         230         239         248         257         266
                  >
CAAAAACTAA AGGG ATG TGG AGA GTG CGG AAA AGG GGC TAC TTT GGG ATT TGG TCC
                M   W   R   V   R   K   R   G   Y   F * G   I   W   S

       275         284         293         302         311         320
TTC CCC TTA ATA ATC GCC GCT GTC TGT GCG CAG AGT GTC AAT GAC CCT AGT AAT
F   P   L   I   I   A   A   V   C   A   Q   S   V   N   D   P   S   N

       329         338         347         356         365         374
ATG TCG CTG GTT AAA GAG ACG GTG GAT AGA CTC CTG AAA GGC TAT GAC ATT CGT
M   S   L   V   K   E   T   V   D   R   L   L   K   G   Y   D   I   R

       383         392         401         410         419         428
CTG AGA CCA GAT TTT GGA GGT CCC CCC GTG GCT GTG GGG ATG AAC ATT GAC ATT
L   R   P   D   F   G   G   P   P   V   A   V   G   M   N   I   D   I

       437         446         455         464         473         482
GCC AGC ATC GAT ATG GTT TCT GAA GTC AAT ATG GAT TAT ACC TTG ACA ATG TAC
A   S   I   D   M   V   S   E   V   N   M   D   Y   T   L   T   M   Y

       491         500         509         518         527         536
TTT CAA CAA GCC TGG AGA GAT AAG AGG CTG TCC TAT AAT GTA ATA CCT TTA AAC
F   Q   Q   A   W   R   D   K   R   L   S   Y   N   V   I   P   L   N

       545         554         563         572         581         590
TTG ACT CTG GAC AAC AGA GTG GCA GAC CAG CTC TGG GTG CCT GAT ACC TAT TTC
L   T   L   D   N   R   V   A   D   Q   L   W   V   P   D   T   Y   F

       599         608         617         626         635         644
CTG AAC GAT AAG AAG TCA TTT GTG CAC GGA GTG ACT GTT AAG AAC CGC ATG ATT
L   N   D   K   K   S   F   V   H   G   V   T   V   K   N   R   M   I

       653         662         671         680         689         698
CGC CTG CAT CCT GAT GGC ACC GTC CTT TAT GGA CTC AGA ATC ACA ACC ACA GCT
R   L   H   P   D   G   T   V   L   Y   G   L   R   I   T   T   T   A

       707         716         725         734         743         752
```

FIGURE 6 (CONTINUED)

```
GCC TGC ATG ATG GAC CTA AGG AGG TAC CCA CTG GAT GAA CAA AAC TGC ACC TTG
A   C   M   M   D   L   R   R   Y   P   L   D   E   Q   N   C   T   L

        761             770             779             788             797             806

GAA ATT GAG AGC TAT GGA TAC ACA ACT GAT GAC ATT GAG TTT TAC TGG CGT GGC
E   I   E   S   Y   G   Y   T   T   D   D   I   E   F   Y   W   R   G

        815             824             833             842             851             860

GAT GAT AAT GCA GTA ACA GGA GTA ACG AAA ATT GAA CTT CCA CAG TTC TCT ATT
D   D   N   A   V   T   G   V   T   K   I   E   L   P   Q   F   S   I

        869             878             887             896             905             914

GTA GAT TAC AAA CTT ATC ACC AAG AAG GTT GTT TTT TCC ACA GGT TCC TAT CCC
V   D   Y   K   L   I   T   K   K   V   V   F   S   T   G   S   Y   P

        923             932             941             950             959             968

AGG TTA TCC CTC AGC TTT AAG CTT AAG AGA AAC ATT GGC TAC TTT ATC CTG CAA
R   L   S   L   S   F   K   L   K   R   N   I   G   Y   F   I   L   Q

        977             986             995            1004            1013            1022

ACA TAC ATG CCT TCC ATC CTG ATT ACC ATC CTC TCC TGG GTC TCC TTC TGG ATT
T   Y   M   P   S   I   L   I   T   I   L   S   W   V   S   F   W   I

       1031            1040            1049            1058            1067            1076

AAT TAC GAT GCT TCA GCT GCA AGG GTG GCA TTA GGA ATC ACA ACT GTC CTC ACA
N   Y   D   A   S   A   A   R   V   A   L   G   I   T   T   V   L   T

       1085            1094            1103            1112            1121            1130

ATG ACC ACA ATC AAC ACC CAC CTC CGG GAA ACT CTC CCT AAA ATC CCC TAT GTG
M   T   T   I   N   T   H   L   R   E   T   L   P   K   I   P   Y   V

       1139            1148            1157            1166            1175            1184

AAG GCC ATT GAC ATG TAC CTG ATG GGG TGC TTT GTC TTC GTT TTC ATG GCC CTT
K   A   I   D   M   Y   L   M   G   C   F   V   F   V   F   M   A   L

       1193            1202            1211            1220            1229            1238

CTG GAA TAT GCC CTA GTC AAC TAC ATC TTC TTT GGG AGG GGG CCC CAA CGC CAA
L   E   Y   A   L   V   N   Y   I   F   F   G   R   G   P   Q   R   Q

       1247            1256            1265            1274            1283            1292

AAG AAA GCA GCT GAG AAG GCT GCC AGT GCC AAC AAT GAG AAG ATG CGC CTG GAT
K   K   A   A   E   K   A   A   S   A   N   N   E   K   M   R   L   D

       1301            1310            1319            1328            1337            1346

GTC AAC AAG ATG GAC CCC CAT GAG AAC ATC TTA CTG AGC ACT CTC GAG ATA AAA
V   N   K   M   D   P   H   E   N   I   L   L   S   T   L   E   I   K

       1355            1364            1373            1382            1391            1400

AAT GAA ATG GCC ACA TCT GAG GCT GTG ATG GGA CTT GGA GAC CCC AGA AGC ACA
N   E   M   A   T   S   E   A   V   M   G   L   G   D   P   R   S   T

       1409            1418            1427            1436            1445            1454

ATG CTA GCC TAT GAT GCC TCC AGC ATC CAG TAT CGG AAA GCT GGG TTG CCC AGG
M   L   A   Y   D   A   S   S   I   Q   Y   R   K   A   G   L   P   R
```

FIGURE 6 (CONTINUED)

```
      1463        1472        1481        1490        1499        1508
CAT AGT TTT GGC CGA AAT GCT CTG GAA CGA CAT GTG GCG CAA AAG AAA AGT CGC
H   S   F   G   R   N   A   L   E   R   H   V   A   Q   K   K   S   R

      1517        1526        1535        1544        1553        1562
CTG AGG AGA CGC GCC TCC CAA CTG AAA ATC ACC ATC CCT GAC TTG ACT GAT GTG
L   R   R   R   A   S   Q   L   K   I   T   I   P   D   L   T   D   V

      1571        1580        1589        1598        1607        1616
AAT GCC ATA GAT CGG TGG TCC CGC ATA TTC TTC CCA GTG GTT TTT TCC TTC TTC
N   A   I   D   R   W   S   R   I   F   F   P   V   V   F   S   F   F

      1625        1634        1643            1656        1666        1676
AAC ATC GTC TAT TGG CTT TAT TAT GTG AAC> TAAAACATGG CCTCCCACTG GAAGCAAGGA
N   I   V   Y   W   L   Y   Y   V   N

      1686        1696        1706        1716        1726        1736        1746
CTAGATTCCT CCTCAAACCA GTTGTACAGC CTGATGTAGG ACTTGGAAAA CACATCAATC CAGGACAAAA

      1756        1766        1776        1786        1796        1806        1816
GTGACGCTAA AATACCTTAG TTGCTGGCCT ATCCTGTGGT CCATTTCATA CCATTTGGGT TGCTTCTGCT

      1826        1836        1846        1856        1866
AAGTAATGAA TACACTAAGG TCCTTGTGGT TTTCCAGTTA AAACGCAAGT
```