

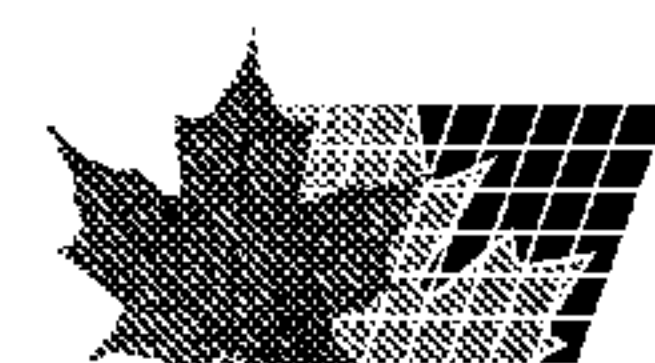


(86) Date de dépôt PCT/PCT Filing Date: 1993/11/11	(51) Cl.Int. ⁶ /Int.Cl. ⁶ C12N 15/12, A61K 38/17, C12N 5/10, C07K 14/705, C12Q 1/02, G01N 33/566
(87) Date publication PCT/PCT Publication Date: 1994/05/26	(72) Inventeurs/Inventors: LE BOURDELLES, BEATRICE, GB; MYERS, JANICE ANN, GB; WHITING, PAUL JOHN, GB
(45) Date de délivrance/Issue Date: 2004/01/20	(73) Propriétaire/Owner: MERCK SHARP & DOHME LIMITED, GB
(85) Entrée phase nationale/National Entry: 1995/05/03	(74) Agent: GOWLING LAFLEUR HENDERSON LLP
(86) N° demande PCT/PCT Application No.: GB 1993/002324	
(87) N° publication PCT/PCT Publication No.: 1994/011501	
(30) Priorités/Priorities: 1992/11/12 (9223769.2) GB; 1993/04/02 (9307026.6) GB	

(54) Titre : ADNc CODANT LA SOUS-UNITE DU RECEPTEUR 22A-NMDA HUMAIN ET DES ISOFORMES DE LA SOUS-UNITE DU RECEPTEUR R1-NMDA HUMAIN, ET LIGNEE CELLULAIRE TRANSFECTEE EXPRIMANT CES MOLECULES
(54) Title: cDNAs ENCODING HUMAN NMDA-22A RECEPTOR SUBUNIT AND ISOFORMS OF THE HUMAN NMDA-R1 RECEPTOR SUBUNIT, TRANSFECTED CELL LINE EXPRESSING THEM

(57) **Abrégé/Abstract:**

The present invention relates to a stably co-transfected eukaryotic cell line capable of expressing an N-methyl-D-aspartate (NMDA) receptor, particularly a human NMDA receptor, which receptor comprises at least one R1 subunit isoform, or at least one R1 subunit isoform and one or two R2 subunits; to the cloning of novel cDNA sequences encoding the human NMDA R2A subunit and various isoforms of the human NMDA R1 subunit; and to the use of the cell line in designing and developing NMDA receptor subtype-selective medicaments.



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

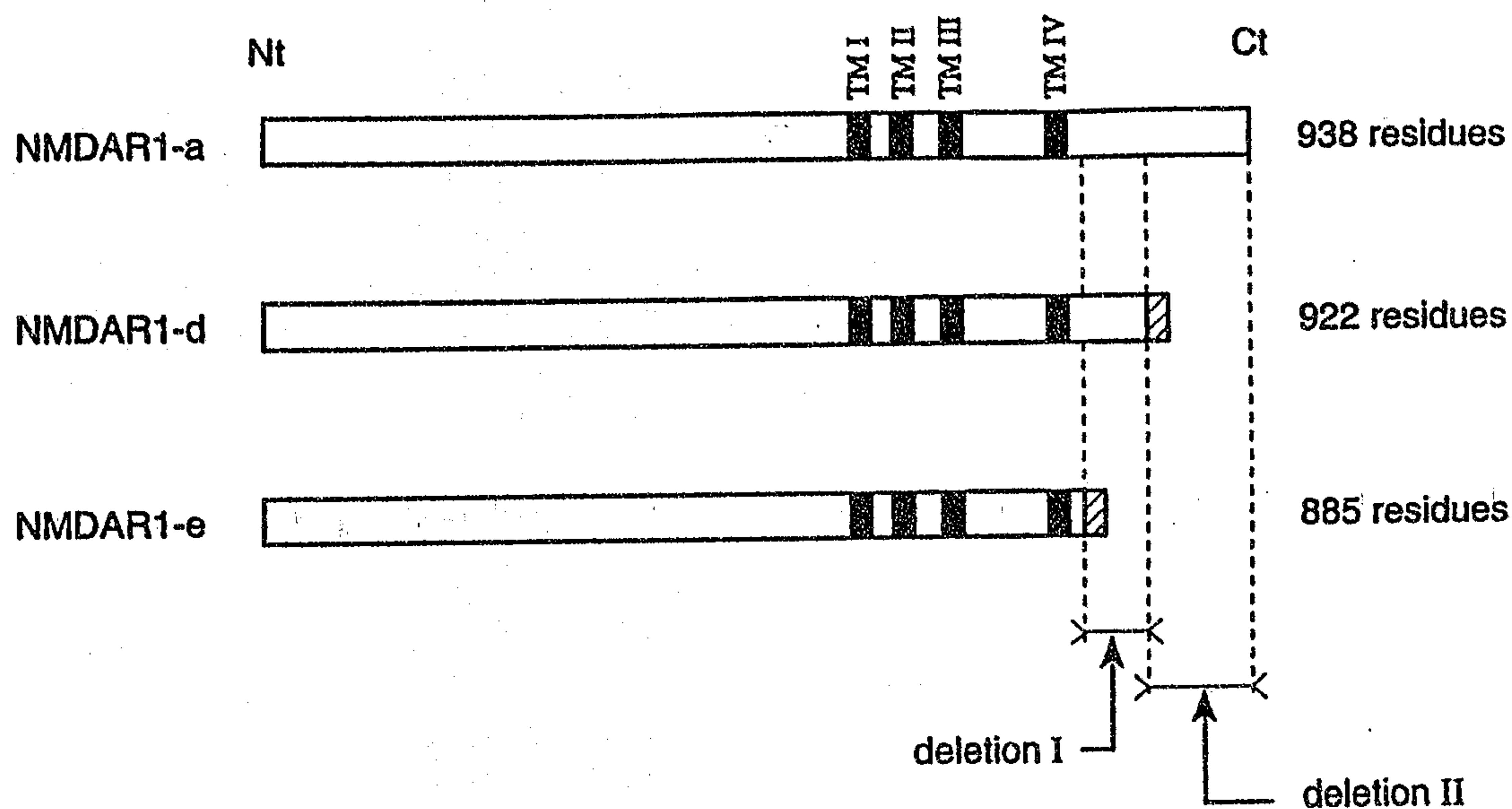
2148599



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12N 15/12, 5/10, G01N 33/68 A61K 37/02		A1	(11) International Publication Number: WO 94/11501 (43) International Publication Date: 26 May 1994 (26.05.94)
(21) International Application Number: PCT/GB93/02324 (22) International Filing Date: 11 November 1993 (11.11.93) (30) Priority data: 9223769.2 12 November 1992 (12.11.92) GB 9307026.6 2 April 1993 (02.04.93) 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) : LE BOURDELLES, Beatrice [FR/GB]; 6 Shearing Lower Road, Sawbridge-worth, Hertfordshire CM20 9LF (GB). MYERS, Janice, Ann [GB/GB]; 66 Beech Hall Road, Highams Park, Chingford, Essex E4 9NX (GB). WHITING, Paul, John [GB/GB]; Tithe Barn, Bentfield Green, Stansted Mount-fitchet, Essex CM24 8HY (GB).			(74) Agent: THOMPSON, John; Merck & Co., Inc., Terlings Park, Eastwick Road, Harlow, Essex CM20 2QR (GB). (81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT). Published With international search report.

(54) Title: cDNAs ENCODING HUMAN NMDA-22A RECEPTOR SUBUNIT AND ISOFORMS OF THE HUMAN NMDA-R1 RECEPTOR SUBUNIT, TRANSFECTED CELL LINE EXPRESSING THEM



(57) Abstract

The present invention relates to a stably co-transfected eukaryotic cell line capable of expressing an N-methyl-D-aspartate (NMDA) receptor, particularly a human NMDA receptor, which receptor comprises at least one R1 subunit isoform, or at least one R1 subunit isoform and one or two R2 subunits; to the cloning of novel cDNA sequences encoding the human NMDA R2A subunit and various isoforms of the human NMDA R1 subunit; and to the use of the cell line in designing and developing NMDA receptor subtype-selective medicaments.

- 1 -

CDNAs ENCODING HUMAN NMDA-22A RECEPTOR SUBUNIT AND ISOFORMS
OF THE HUMAN NMDA-R1 RECEPTOR SUBUNIT,
TRANSFECTED CELL LINE EXPRESSING THEM.

5 The present invention concerns a cell line, and
in particular relates to a stable cell line capable of
expressing human or animal N-methyl-D-aspartate (NMDA)
receptors. The invention also relates to complementary
DNAs (cDNAs) encoding novel human NMDA receptor subunits,
10 and concerns in particular the nucleotide and deduced
amino acid sequences of the human NMDA R2A receptor
subunit, and of various isoforms of the human NMDA R1
receptor subunit.

15 The NMDA receptor is the major excitatory amino
acid receptor that mediates glutamate transmission in the
central nervous system. This receptor has been
implicated in neuronal modulation, including long term
potentiation in the hippocampus (Collingridge and Singer,
TIPS, 1990, 11, 290). It is consequently believed to
20 play a key role in memory acquisition and learning.

 The integral channel of the NMDA receptor
allows Ca^{2+} to permeate as well as Na^{+} and K^{+} , and
presents at least seven pharmacologically distinct sites.
These include a glutamate binding site, a glycine binding
25 site, a polyamine site, a dizocilpine binding site, a
voltage dependent Mg^{2+} site, a Zn^{2+} binding site (Wong
and Kemp, Ann. Rev. Pharmacol. Toxicol., 1991, 31, 401)
and an ifenprodil binding site (Carter et al., J.
Pharmacol. Exp. Ther., 1988, 247, 1222).

30 The development of potent and selective NMDA
receptor antagonists which penetrate into the brain has
received considerable attention of late as an attractive
strategy for treating and/or preventing conditions which
are believed to arise from over-stimulation of
35 neurotransmitter release by excitatory amino acids. Such
conditions notably include neurodegenerative disorders
arising as a consequence of such pathological conditions

- 2 -

as stroke, hypoglycaemia, cerebral palsy, transient cerebral ischaemic attack, cerebral ischaemia during cardiac pulmonary surgery or cardiac arrest, perinatal asphyxia, epilepsy, Huntington's chorea, Alzheimer's disease, Amyotrophic Lateral Sclerosis, Parkinson's disease, Olivo-ponto-cerebellar atrophy, anoxia such as from drowning, spinal cord and head injury, and poisoning by exogenous and endogenous NMDA receptor agonists and neurotoxins, including environmental neurotoxins.

10 NMDA receptor antagonists may also be useful as anticonvulsant and antiemetic agents, as well as being of value in the prevention or reduction of dependence on dependence-inducing agents such as narcotics.

15 NMDA receptor antagonists have recently been shown to possess analgesic (see, for example, Dickenson and Aydar, Neuroscience Lett., 1991, 121, 263; Murray et al., Pain, 1991, 44, 179; and Woolf and Thompson, Pain, 1991, 44, 293) and anxiolytic (see, for example, US-5145866; and Kehne et al., Eur. J. Pharmacol., 1991, 193, 283) effects, and such compounds may accordingly be useful in the management of pain, depression and anxiety.

25 Compounds possessing functional antagonist properties for the NMDA receptor complex are stated in WO-A-91/19493 to be effective in the treatment of mood disorders, including major depression, bipolar disorder, dysthymia and seasonal affective disorder (cf. also Trullas and Skolnick, Eur. J. Pharmacol., 1990, 185, 1). Such compounds may consequently be of benefit in the treatment and/or prevention of those disorders.

30 The association of NMDA receptor antagonists with regulation of the dopaminergic system has recently been reported (see, for example, Werling et al., J. Pharmacol. Exp. Ther., 1990, 255, 40; Graham et al., Life Sciences, 1990, 47, PL-41; Hutson et al., Br. J. Pharmacol., 1991, 103, 2037; and Turski et al., Nature (London), 1991, 349, 414). This suggests that such compounds may thus be of assistance in the prevention

35

- 3 -

and/or treatment of disorders of the dopaminergic system such as schizophrenia and Parkinson's disease.

It has also been reported recently (see Lauritzen et al., Journal of Cerebral Blood Flow and Metabolism, 1991, vol. 11, suppl. 2, Abstract XV-4) that NMDA receptor antagonists block cortical spreading depression (CSD), which may thus be of clinical importance since CSD is a possible mechanism of migraine. The class of substituted 2-amino-4-phosphonomethylalk-3-ene carboxylic acids and esters described in EP-A-0420806, which are stated to be selective NMDA antagonists, are alleged thereby to be of potential utility in the treatment of inter alia migraine.

Excitatory amino acid receptor antagonists, including inter alia antagonists of NMDA receptors, are alleged in EP-A-0432994 to be of use in suppressing emesis.

Recent reports in the literature have also suggested a link between the neurotoxicity of certain viruses and the deleterious effects of these viruses on an organism caused by the potentiation of neurotransmission via excitatory amino acid receptors. Antagonists of NMDA receptors may therefore be effective in controlling the manifestations of neuroviral diseases such as measles, rabies, tetanus (cf. Bagetta et al., Br. J. Pharmacol., 1990, 101, 776) and AIDS (cf. Lipton et al., Society for Neuroscience Abstracts, 1990, 16, 128.11).

NMDA antagonists have, moreover, been shown to have an effect on the neuroendocrine system (see, for example, van den Pol et al., Science, 1990, 250, 1276; and Urbanski, Endocrinology, 1990, 127, 2223), and such compounds may therefore also be effective in the control of seasonal breeding in mammals.

A cDNA, encoding a subunit of the rat NMDA receptor and designated NMDA R1, has been cloned by expression cloning (Moriyoshi et al., Nature (London),

- 4 -

1991, 354, 31). When expressed in Xenopus oocytes, this cDNA exhibits the electrophysiological and pharmacological properties expected of an authentic NMDA receptor, although the levels of expression are extremely low. More recently, the existence of several discrete isoforms of the rat NMDA R1 receptor subunit, generated by alternative RNA splicing, has been reported (Sugihara et al., BBRC, 1992, 185, 826). Using both low stringency hybridization and polymerase chain reaction methodologies, four additional rodent NMDA receptor subunit cDNAs have been cloned: ϵ 1 or NMDA R2A; ϵ 2 or NMDA R2B; ϵ 3 or NMDA R2C; and ϵ 4 or NMDA R2D (see Monyer et al., Science, 1992, 256, 1217; Kutsuwada et al., Nature (London), 1992, 358, 36; Ikeda et al., FEBS Lett., 1992, 313, 34; and Ishii et al., J. Biol. Chem., 1993, 268, 2836). Co-expression in Xenopus oocytes or transiently transfected cells of the NMDA R1 subunit, with any one of the R2A, R2B, R2C or R2D subunits referred to above, gives rise to a more robust NMDA receptor than that constituted by the NMDA R1 subunit alone (Monyer et al., Science, 1992, 256, 1217). Moreover, these four resulting putative NMDA receptors (R1/R2A; R1/R2B; R1/R2C; and R1/R2D) are observed to be pharmacologically and electrophysiologically distinguishable. These data support the hypothesis that a family of NMDA receptor subtypes with distinct pharmacological profiles may exist in the brain through combination of different subunits.

Any of a variety of procedures may be used to molecularly clone human NMDA receptor cDNA. These methods include, but are not limited to, direct functional expression of the human NMDA receptor cDNAs following the construction of a human NMDA receptor containing cDNA library in an appropriate expression vector system. Another method is to screen a human NMDA receptor containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a labelled

- 5 -

oligonucleotide probe designed from the amino acid sequence of the purified NMDA receptor protein or from the DNA sequence of known NMDA receptor cDNAs. The preferred method consists of screening a human NMDA receptor containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a ^{32}P -labelled cDNA oligonucleotide-primed fragment of rodent NMDA receptor subunit cDNA. The preferred human cDNA library is a commercially available human hippocampal cDNA library.

It is readily apparent to those skilled in the art that other types of libraries, as well as libraries constructed from other brain regions, may be useful for isolating DNA encoding the human NMDA receptor. Other types of libraries include, but are not limited to, cDNA libraries derived from other tissues, cells or cell lines other than human hippocampal cells, and genomic DNA libraries.

Preparation of cDNA libraries can be performed by standard techniques well known in the art. Well known cDNA library construction techniques can be found, for example, in Maniatis, T., Fritsch, E.F., Sambrook, J., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Press, New York, 2nd edition, 1989).

It is also readily apparent to those skilled in the art that DNA encoding the human NMDA receptor may also be isolated from a suitable genomic DNA library.

Construction of genomic DNA libraries can be performed by standard techniques well known in the art. Well known genomic DNA library construction techniques can be found in Maniatis *et al.*, supra.

Using the preferred method, cDNA clones encoding the human NMDA receptor were isolated by cDNA library screening. ^{32}P -radiolabelled fragments of rodent NMDA receptor NMDA R1 and NMDA R2A subunit cDNAs served as probes for the isolation of human NMDA receptor cDNA

2148599

- 6 -

from a commercially available cDNA library derived from human hippocampal tissue.

For the NMDA R1 subunit, several positively hybridising clones were detected using the rodent NMDA R1 subunit cDNA probe. The longest of these cDNA clones, which was homologous to the published NMDA R1e isoform (Sugihara *et al.*, BBRC, 1992, 185, 826), lacked approximately 300 base pairs of the 5' end of the coding region. The missing sequence could be restored by conventional techniques, which essentially involved screening the same hippocampal cDNA library with an EcoRI-SmaI fragment encoding the last 300 nucleotides of the 5' end of the truncated NMDA R1e clone. A cDNA clone containing the missing sequence could be isolated and engineered by standard methods onto the truncated NMDA R1e cDNA at the internal SmaI site in order to generate a full-length NMDA R1e cDNA, as described in accompanying Example 1.

Analogous techniques applied to other positive clones isolated from the human hippocampal cDNA library have yielded cDNAs encoding isoforms of the human NMDA R1 receptor subunit corresponding to the published rat NMDA R1a and NMDA R1d isoforms (Sugihara *et al.*, *supra*).

For the NMDA R2A subunit, several hybridising clones were detected using the rodent NMDA R2A subunit cDNA probe. As with the situation encountered in connection with the R1 subunit as described above, none of these cDNAs obtained from the human hippocampal cDNA library encoded the complete human NMDA R2A deduced amino acid sequence. For this reason, a full-length cDNA was constructed from overlapping truncated cDNAs, as described in accompanying Example 2.

Figure 1 depicts the schematic structures of the R1a, R1d and R1e isoforms of the human NMDA R1 receptor subtype. Four putative transmembrane segments (TMI-IV) are represented by solid boxes. R1d has a deletion (deletion II) at the region corresponding to

0028035

- 7 -

nucleotide residues no. 2701-3056 of the R1a cDNA. This deletion results in the generation of a new 22 amino acid carboxyl-terminal (Ct) sequence that follows amino acid residue 900 of the R1a sequence. R1e has a further deletion (deletion I) and accordingly lacks the sequences of both deletions I and II, thus possessing a structure with the new 22 amino acid carboxyl-terminal sequence linked to the R1a sequence at position 863. This 22 amino acid carboxyl-terminal sequence, created by alternative splicing, is indicated in Figure 1 by hatched boxes.

The sequence for the full-length cDNA encoding the R1e isoform of the human NMDA receptor is shown in Figure 2A. The deduced amino acid sequence is shown alongside the cDNA sequence in Figure 2A. The positions of the four putative transmembrane segments (TM1-4) and of the predicted signal peptide (SP) are indicated in Figure 2A by solid lines.

Figure 2B depicts the alternative carboxy-terminal sequences of the cDNAs encoding the R1a, R1d and R1e isoforms of the human NMDA R1 receptor subtype. The nucleotide and deduced amino acid sequences are indicated in Figure 2B for each isoform. Indicated amino acid position 863 and nucleotide position 2589 relate to the corresponding positions in the sequence of the R1e isoform depicted in Figure 2A. The nucleotide and deduced amino acid sequences of deletion I (37 amino acids) and deletion II (38 amino acids) are indicated in Figure 2B. The R1d and R1e isoforms are generated from the R1a isoform by a combination of deletion I and deletions I+II respectively, followed by the addition of the alternative 22 amino acid carboxy-terminal sequence represented by the hatched boxes in Figure 1, the nucleotide and deduced amino acid sequences of which are indicated in Figure 2B. The untranslated domains are represented in lower case lettering.

2148599

The complete nucleotide and deduced amino acid sequences of the cDNA encoding the human NMDA R2A receptor subunit are shown in Figure 3. The cDNA is 4858 bases long. The open reading frame is between bases 47 and 4438. Bases 1-46 are 5' untranslated sequence, and bases 4439-4858 are 3' untranslated sequence. The deduced amino acid sequence is 1464 residues long and has 81 differences from the published rat NMDA R2A sequence (Monyer et al., Science, 1992, 256, 1217).

The cloned human NMDA receptor cDNA obtained through the methods described above may be recombinantly expressed by molecular cloning into an expression vector containing a suitable promoter and other appropriate transcription regulatory elements, and then transferring the expression vector into prokaryotic or eukaryotic host cells to produce a recombinant NMDA receptor. Techniques for such manipulations are fully described in Maniatis et al., supra, and are well known in the art.

DNA encoding the NMDA receptor cloned into an expression vector may then be transferred to a recombinant host cell for expression. Recombinant host cells may be prokaryotic or eukaryotic, including but not limited to bacteria, 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 derived cell lines). Cell lines derived from mammalian species which may be suitable and which are commercially available include, but are not limited to, L-M (ATCC CCL 1), HEK293 (ATCC CCL 1555), 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) and MRC-5 (ATCC CCL 171).

The expression vector may be introduced into host cells via any one of a number of techniques including but not limited to transformation,

- 9 -

transfection, protoplast fusion, and electroporation. The expression vector-containing cells are clonally propagated and individually analyzed to determine whether they produce NMDA receptor protein. Identification of
5 NMDA receptor expressing host cell clones may be done by several means, including but not limited to immunological reactivity with anti-NMDA receptor antibodies, and the presence of host cell-associated NMDA receptor activity.

Expression of NMDA receptor DNA may also be
10 performed using in vitro produced synthetic mRNA. Synthetic mRNA can be efficiently translated in various cell-free systems, including but not limited to wheat germ extracts and reticulocyte extracts, as well as efficiently translated in cell based systems, including
15 but not limited to microinjection into frog oocytes.

To determine the NMDA receptor cDNA sequence(s) that yield(s) optimal levels of binding activity and/or NMDA receptor protein, NMDA receptor cDNA molecules including but not limited to the following can be
20 constructed: the full-length open reading frame of the human NMDA receptor cDNA depicted in Figure 2A or Figure 3 and constructs containing portions of the cDNA encoding biologically active human NMDA receptor protein. All constructs can be designed to contain none, all, or
25 portions of the 5' and 3' untranslated region of human NMDA receptor cDNA. NMDA receptor activity and levels of protein expression can be determined following the introduction, both singly and in combination, of these constructs into appropriate host cells. Following
30 determination of the NMDA receptor cDNA cassette which yields optimal expression, this cDNA construct may be transferred to a variety of expression host cells, including but not limited to mammalian cells, baculovirus-infected insect cells, E. coli, and the yeast
35 S. cerevisiae.

Expression of the human NMDA receptor in a recombinant host cell affords NMDA receptor protein in

2148599

- 10 -

active form. Several purification procedures are available and suitable for use. Recombinant NMDA receptor may suitably be purified from cell lysates and extracts, or from conditioned culture medium, by various combinations of, or individual application of, salt fractionation, ion exchange chromatography, size exclusion chromatography, hydroxylapatite adsorption chromatography and hydrophobic interaction chromatography.

10 In addition, recombinant human NMDA receptor can be separated from other cellular proteins by use of an immuno-affinity column made with monoclonal or polyclonal antibodies specific for the NMDA receptor, or polypeptide fragments thereof. The preparation and
15 purification of monoclonal or polyclonal antibodies specific for the NMDA receptor or polypeptide fragments thereof can be accomplished by conventional techniques well known in the art. Typical procedures include those described, for example, by Maniatis et al., in Molecular Cloning, A Laboratory manual, Cold Harbor Press, New
20 York, 2nd edition, 1989, Chapter 18.

The present invention is concerned with the production of permanently transfected cells containing the NMDA receptor, which will be useful for screening for
25 drugs which act on this receptor. The NMDA receptor has previously been expressed in Xenopus oocytes (Ishii et al., J. Biol. Chem., 1993, 268, 2836) and in transiently transfected mammalian cells (Monyer et al., Science, 1992, 256, 1217). However, both of those systems involve
30 transient expression and are unsuitable for screening purposes.

We have now achieved the stable expression of the receptor.

Accordingly, the present invention provides a
35 stably co-transfected eukaryotic cell line capable of expressing an NMDA receptor, which receptor comprises at

- 11 -

least one R1 subunit isoform, or at least one R1 subunit isoform and one or two R2 subunits.

5 This has been achieved by co-transfecting cells with expression vectors, each harbouring cDNAs encoding for at least one R1 subunit isoform, or for at least one R1 subunit isoform and one or two R2 subunits respectively. In a further aspect, therefore, the present invention provides a process for the preparation of a eukaryotic cell line capable of expressing an NMDA
10 receptor, which comprises stably co-transfecting a eukaryotic host cell with one or more expression vectors, at least one such vector harbouring the cDNA sequence encoding for an NMDA R1 receptor subunit isoform (referred to in the art as the "key subunit"), and
15 optionally other such vectors harbouring the cDNA sequences encoding for one or two different NMDA R2 receptor subunits. The stable cell line which is established expresses an R1 or an R1+R2 NMDA receptor respectively. Each receptor thereby expressed,
20 comprising a unique combination of at least one R1 subunit isoform, or at least one R1 subunit isoform and one or two R2 subunits, will be referred to hereinafter as an NMDA receptor "subunit combination".
Pharmacological and electrophysiological data confirm
25 that the recombinant R1 or R1+R2 receptors expressed by the cells of the present invention have the properties expected of a native NMDA receptor.

As indicated above, expression of the NMDA receptor may be accomplished by a variety of different
30 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
35 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

2148599

- 12 -

incorporate at least one R1 subunit isoform, or at least one R1 subunit isoform and one or two R2 subunits, into the cell line in order to produce the required receptor, referred to as R1 or R1+R2, respectively. Within this
5 limitation, the choice of receptor subunit combination is made according to the type of activity or selectivity which is being screened for.

In order to employ this invention most effectively for screening purposes, it is preferable to
10 build up a library of cell lines, each with a different combination of subunits. Typically a library of 11 cell line types is convenient for this purpose. Preferred subunit combinations include: R1, R1+R2A, R1+R2B, R1+R2C, R1+R2D, R1+R2A+R2B, R1+R2A+R2C, R1+R2A+R2D, R1+R2B+R2C,
15 R1+R2B+R2D and R1+R2C+R2D. The nomenclature 'R1' signifies any one of the seven reported isoforms of the NMDA R1 subunit (Sugihara et al., BBRC, 1990, 185, 826): R1a, R1b, R1c, R1d, R1e, R1f and R1g. A particular subunit combination is R1a+R2A.

20 In a particular embodiment, the present invention provides a stably co-transfected eukaryotic cell line capable of expressing a human NMDA receptor comprising the R1a subunit isoform.

In a further embodiment, the present invention
25 provides a stably co-transfected eukaryotic cell line capable of expressing a human NMDA receptor comprising the R1a+R2A subunit combination.

As indicated above, the DNAs for the receptor subunits can be obtained from known sources, and are
30 generally obtained as specific nucleotide sequences harboured by a standard cloning vector such as those described, for example, by Maniatis et al., supra. Preferably the cDNA sequences are derived from the human gene. However, for screening purposes, cDNAs from other
35 species are also suitable, such as bovine or rat DNA. Known sources of NMDA receptor subunit cDNAs are as follows:

- 13 -

- R1 isoforms (rat): Sugihara et al., BBRC, 1992, 185,
826.
Durand et al., PNAS, 1992, 89, 9359.
5 Hollmann et al., Neuron, 1993, 10,
943.
- R2A, R2B, R2C, R2D (rat): Ishii et al., J. Biol.
Chem., 1993, 268, 2836.
- R2A, R2B, R2C (rat): Monyer et al., Science, 1992,
10 256, 1217.
- R1e (human): Planells-Cases, PNAS, 1993, 90, 5057.
- R1a (human): Karp et al., J. Biol. Chem., 1993, 268,
3728.
- R1a = ϵ^1 , R2A = ϵ^1 , R2B = ϵ^2 , R2C = ϵ^3 (mouse):
15 Kutsuwada et al., Nature, 1992, 358, 36.

In another aspect, the invention provides a recombinant expression vector comprising the nucleotide sequence of an NMDA receptor subunit together with additional sequences capable of directing the synthesis of the said NMDA 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; selectable markers; a limited number of useful restriction enzyme sites; a potential for high copy number; and strong promoters. A promoter is defined as a

- 14 -

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.

5 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
10 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
15 cosmids.

A variety of mammalian expression vectors may be used to express recombinant NMDA receptor in mammalian cells. Commercially available mammalian expression vectors which may be suitable for recombinant NMDA
20 receptor expression include, but are not limited to, pCDNAneo (Invitrogen), pCDNAI-Amp (Invitrogen), pCDM8 (Invitrogen), pMSGneo (Proc. Natl. Acad. Sci. USA, 1992, 89, 6378), pMC1neo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593), pBPV-1(8-2) (ATCC 37110), pδBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460), and gZD35 (ATCC 37565).
25

The recombinant expression vector in accordance with the invention may be prepared by inserting the
30 nucleotide sequence of the chosen NMDA 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
35 constructed by standard techniques from known expression vectors. The precursor vector suitably contains a selection marker, typically an antibiotic resistance

- 15 -

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 ion channel intrinsic to the NMDA 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 3 herein.

The recombinant expression vector of the present invention is then produced by cloning the NMDA 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 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 3, step (b) herein.

One suitable expression vector of the present invention is illustrated in Figure 4 of the accompanying drawings, in which R represents the nucleotide sequence of a chosen R1 or R2 subunit of the NMDA receptor, and the remainder of the expression vector depicted therein is derived from the precursor vector pMSGneo and constructed as described in accompanying Example 3, steps (a) and (b).

2148599

- 16 -

For each cell line of the present invention, one or more such vectors will be necessary. At least one such vector will contain the cDNA sequence encoding for an R1 subunit isoform. Vectors containing the cDNA sequences encoding for one or two different R2 subunits may also be utilised.

Cells are then co-transfected with the desired combination of one or more, typically one, two or 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 NMDA receptor subunit combinations, especially human NMDA receptor subunit combinations, derived from cultures of stably transfected eukaryotic cells. The invention also provides preparations of membranes containing subunit combinations of the NMDA receptor, especially human NMDA receptor subunit combinations, derived from cultures of stably transfected eukaryotic cells. In particular, the protein preparations and membrane preparations according to the invention will suitably contain the R1, R1+R2A, R1+R2B,

- 17 -

R1+R2C, R1+R2D, R1+R2A+R2B, R1+R2A+R2C, R1+R2A+R2D, R1+R2B+R2C, R1+R2B+R2D or R1+R2C+R2D subunit combinations of the human NMDA receptor, wherein the nomenclature 'R1' signifies any one of the several isoforms of the R1 subunit as described above. A preferred subunit combination is the R1a+R2A subunit combination. In an especially preferred embodiment, the invention provides cell membranes containing a human NMDA receptor consisting of the R1a+R2A subunit combination 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 NMDA receptor. The present invention accordingly provides the use of the cell line described above, of membrane preparations derived therefrom, and of the cloned human NMDA receptor as described herein, in screening for and designing medicaments which interact selectively with the NMDA receptor. Of particular interest in this context are molecules capable of interacting selectively with NMDA 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 NMDA receptor subtypes.

This invention provides an antisense oligonucleotide having a sequence capable of binding specifically with any sequences of an mRNA molecule which encodes the human NMDA receptor, so as to prevent translation of the mRNA molecule. The antisense oligonucleotide may have a sequence capable of binding specifically with any sequences of the cDNA molecule whose sequence is shown in, for example, Figure 2A or Figure 3. A particular example of an antisense

- 18 -

oligonucleotide is an antisense oligonucleotide comprising chemical analogues of nucleotides.

This invention provides a transgenic nonhuman mammal expressing DNA encoding the human NMDA receptor. This invention also provides a transgenic nonhuman mammal, so mutated as to be incapable of normal receptor activity, and not expressing the native NMDA receptor, which mammal is nonetheless capable of expressing DNA encoding the human NMDA receptor. This invention further provides a transgenic nonhuman mammal whose genome comprises antisense DNA complementary to DNA encoding the native NMDA receptor, so placed as to be transcribed into antisense mRNA which is complementary to mRNA encoding the receptor and which hybridizes to mRNA encoding the receptor thereby reducing its translation. The DNA may additionally comprise an inducible promoter or additionally comprise tissue specific regulatory elements, so that expression can be induced, or restricted to specific cell types. Examples of DNA are DNA or cDNA molecules having a coding sequence substantially the same as the coding sequence shown in, for example, Figure 2A or Figure 3. An example of a transgenic animal is a transgenic mouse. Examples of tissue specificity-determining regions are the metallothionein promoter (Low et al., Science, 1986, 231, 1002-1004) and the L7 promoter (Oberdick et al., Science, 1990, 248, 223-226).

The following non-limiting Examples illustrate the present invention.

- 19 -

EXAMPLE 1Isolation of cDNAs encoding three different isoforms of the human NMDA R1 receptor subunit

5

1. Methods

Oligonucleotide primers derived from the published rat NMDA R1 receptor sequence (Moriyoshi *et al.*, *Nature (London)*, 1991, 354, 31-37) [O₁: 5' GAC/CCC/AGG/CTC/AGA/ATT/CCC/TCA/GAC/AAG3' and O₂: 5' CAC/CAG/GAA/GAA/GTC/TGC/CAT/GTT/CTC/A3'] were used in the polymerase chain reaction (PCR) to isolate a 420bp cDNA encoding the putative large cytoplasmic loop of the human NMDA R1 receptor subtype. A human hippocampal cDNA library was screened at low stringency (55°C in 5 x SSPE, 5 x Denhardt's solution, 0.1% SDS, 100 µg/ml salmon sperm DNA) using this partial PCR cDNA as a [³²P]-labelled probe containing 0.5 x 10⁶ cpm/ml. Twenty positive clones encoding putative NMDA receptors were isolated. The longest clone, homologous to the published NMDA R1e, 0.6kb long, was lacking approximately 300bp of the 5' end of the coding region. This missing sequence was obtained by screening the same human hippocampal cDNA library with an EcoRI-SmaI fragment encoding the last 300 nucleotides of the 5' end of the truncated NMDA R1e clone. A cDNA clone containing the missing sequence was isolated and engineered onto the truncated NMDA R1e cDNA at the internal SmaI site in order to generate a full length NMDA R1e cDNA. This full length human NMDA R1e clone contains a 2.8kb coding sequence flanked by 0.05kb and 0.6kb of the 5' and the 3' untranslated regions respectively.

35

2. Results

The sequence of the human NMDA R1e cDNA (Figure 2A) is homologous (98.6% of homology) with the published rat NMDA R1e cDNA (Sugihara *et al.*, *BBRC*, 1992, 185, 826-

- 20 -

832). The amino acid sequence of the C-terminal domain of this NMDA R1e subunit differs from the original published rat NMDA R1a subunit (Moriyoshi *et al.*, *Nature* (London), 1991, 354, 31-37). Among the twenty positive clones isolated from a human hippocampal cDNA library, screened with the 420bp long PCR fragment, two other clones have revealed a sequence homologous to the NMDA R1e; they are the human version of the published rat NMDA R1a and NMDA R1d, respectively (Sugihara *et al.*, *BBRC*, 1992, 185, 826-832). Figure 1 shows the schematic structures of the three human NMDA R1a, R1d and R1e isoforms, on the basis of the structural comparison with the original NMDA R1 (now referred to as R1a). R1d has a deletion (deletion II, whose nucleotide and deduced amino acid sequences are shown in Figure 2B) at the region corresponding to nucleotide residues no. 2701-3056 of the R1a cDNA. This deletion results in the generation of a new 22 amino acid carboxyl-terminal sequence that follows amino acid residue 900 of the R1a sequence. The nucleotide and deduced amino acid sequences of this new 22 amino acid sequence are depicted in Figure 2B. R1e has a further deletion (deletion I, whose nucleotide and deduced amino acid sequences are shown in Figure 2B) and accordingly lacks the sequences of both deletions I and II, thus possessing a structure with the new 22 amino acid carboxyl-terminal sequence (cf. Figure 2B) linked to the R1a sequence at position 863.

EXAMPLE 2

30

Isolation of cDNAs encoding the human NMDA R2A subunit

Oligonucleotide primers derived from the published rat NMDA R2A cDNA sequence (Monyer *et al.*, *Science*, 1992, 256, 1217) were used in the polymerase chain reaction (PCR) to generate N-terminal and C-terminal cDNA probes. Primers for N-terminal probe

35

- 21 -

(encoding residues Leu16-Met74) were
5'TGCCGGGAATTCTGGTCTGGCGCGATCCGG3' and
5'GGGTCAAGCTTTTCATCAATAACGCCAC3'. Primers for the
C-terminal probe (encoding residues Leu898-Gln1209) were
5'TGCTAAGATTCTTCGGTCAGCTAAAAAC3' and
5'AATGTGAAGCTTTCTGCCGGTATCGATCAC3'. The sense
oligonucleotide (the first of each pair) had an EcoRI
restriction enzyme site incorporated into it, while the
antisense oligonucleotide had a HindIII restriction
enzyme site incorporated into it. PCR using rat brain
cDNA as template was performed using standard techniques
(e.g. Whiting et al., Proc. Natl. Acad. Sci. USA, 1990,
87, 9966-9970) to yield PCR products of 185 and 950 bases
in length (N-terminal cDNA probe and C-terminal cDNA
probe, respectively). These were subcloned into
pBluescript*Sk⁻ (Stratagene). A human hippocampal cDNA
library (Stratagene) was screened under moderate
stringency conditions (42°C in 5 x SSPE, 5 x Denhardt's
solution, 100 mg/ml salmon sperm DNA, 0.1% sodium dodecyl
sulphate, 30% formamide) using standard techniques (e.g.
Maniatis et al., in Molecular Cloning: A Laboratory
Manual, Cold Spring Harbor Press, New York, 2nd edition,
1989). A number of cDNA clones were isolated and
characterised. These cDNAs encoded the entire human NMDA
R2A receptor subunit, with the exception of a sequence
spanning bases approx. 300-900 of the coding region.
This missing sequence was obtained by PCR:
oligonucleotide primers derived from the 5' untranslated
region of the human NMDA R2A receptor subunit cDNA
sequence (5'CCCCACTGCATCCGGTACCTTCTCGGGCTACA3') and the
coding region of the human NMDA R2A receptor subunit cDNA
sequence, Figure 3, bases 873-902
(5'AGTCCCAGTCATATGAGACAGAAATGA3'), were used to generate
a 5' cDNA fragment using human hippocampal cDNA as
template. A full-length human NMDA R2A receptor subunit
cDNA was generated by assembling overlapping cDNAs using
standard molecular biology techniques, and the 5' cDNA

*Trade-mark

- 22 -

fragment generated by PCR and sequenced on both strands using Taq dideoxy terminator cycle sequencing (Applied Biosystems) and an Applied Biosystems 373A DNA Sequencer. The complete nucleotide and deduced amino acid sequence is shown in Figure 3.

5

EXAMPLE 3

Preparation of R1a transfected cells and R1a+R2A transfected cells

10

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.

15

20

25

b) Cloning of subunit cDNAs into pMSGneo

Human R1a and R2A cDNA were cloned as described in Examples 1 and 2 above respectively. Human R1a was subcloned from pBluescript Sk⁻ (Stratagene, San Diego, CA, USA) by digestion with SalI and EcoRI and ligation into the SalI and EcoRI sites of pMSGneo. Human R2A was subcloned from pBluescript Sk⁻ by restriction digestion with NheI and XhoI and ligation into the NheI and XhoI sites of pMSGneo.

30

35

- 23 -

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% CO₂, in Modified Eagles Medium containing penicillin, streptomycin and 10% fetal calf serum. The expression vector harbouring the NMDA receptor subunit DNAs for co-transfection was prepared by a standard protocol (Chen, C. and Okayama, H., BioTechniques, 6, 632-638, 1988). For co-transfection, Ltk⁻ cells were plated in dishes (approx. 2x10⁵ 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, United Kingdom)]. 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). In order to prevent cell death induced by the expression of the recombinant R1a+R2A receptor, 10mM MgCl₂ and 100 µM D-AP5 (D-2-amino-5-phosphonovalerate) (Tocris-Neuramin, Essex, United Kingdom) were added into the culture medium in addition to the dexamethasone. MgCl₂ and D-AP5 are substances known to act as antagonists at the NMDA receptor (Wong and Kemp, Ann. Rev. Pharmacol. Toxicol., 1991, 31, 401). 3-5 days later the cells were harvested, membranes prepared and used for radioligand binding (see Example 4, step (a) below), using the glycine site antagonist [³H]-

*Trade-mark

- 24 -

L-689,560 (custom synthesised by Amersham International plc, Amersham, United Kingdom) for the characterisation of the recombinant R1 receptor, and the channel blocker [³H]-dizocilpine (obtained from New England Nuclear, Du Pont Ltd., Stevenage, United Kingdom) for the

5 characterisation of the recombinant R1a+R2A receptor.

The recombinant R1a receptor clone expressing the highest amount of [³H]-L-689,560 binding was subcloned from a single cell by limiting dilution. The

10 resultant clonal population of these cells is referred to hereinbelow as population A.

The recombinant R1a+R2A receptor clone expressing the highest amount of [³H]-dizocilpine binding was subcloned from a single cell by limiting dilution.

15 The resultant clonal population of these cells is referred to hereinbelow as population B.

EXAMPLE 4

20 Characterization of R1a transfected cells and R1a+R2A transfected cells

a) Radioligand binding

The nature of the recombinant R1a NMDA

25 receptors prepared as described in Example 3 was addressed by characterization of the glycine site binding pharmacology, using the glycine site antagonist [³H]-L-689,560. For radioligand binding assays, cells which had been induced by culture in dexamethasone containing

30 medium for 3-5 days were scraped off into 5 mM Tris-Acetate, pH 7.0 at 4°C (buffer 1) and pelleted (20,000rpm, Sorvall RC5C centrifuge). The cell pellet was resuspended in buffer 1, homogenised using an Ultra-Turrax* homogeniser and then pelleted as above. This was

35 repeated once more, and the cells then resuspended in 50 mM Tris-Acetate, pH 7.0 at 4°C (buffer 2) at a protein concentration of 1 mg/ml. Radioligand binding was

*Trade-mark

- 25 -

performed in 0.5 ml final volume buffer 2, containing 1 nM of [^{3} H]-L-689,560 and 100 μ g of protein. After 2 hours incubation on ice the membranes were harvested onto filters using a Brandel cell harvester, washed with cold buffer 2, and bound radioactivity determined by scintillation counting. Non-specific binding was determined in a parallel incubation containing 10 mM glycine. The recombinant R1a receptors bound [^{3} H]-L-689,560 at levels of up to 3700 fmols/mg protein. No binding was seen to untransfected Ltk⁻ cells, confirming that the [^{3} H]-L-689,560 was binding to recombinant R1 NMDA receptors.

The nature of the recombinant R1a+R2A NMDA receptors prepared as described in Example 3 was addressed by demonstrating binding of the channel blocker [^{3} H]-dizocilpine. For radioligand binding assays, cells which had been induced by culture in dexamethasone, MgCl₂ and D-AP5 containing medium for 3-5 days were scraped off into 5mM Tris-Acetate, pH 7.0 at 4°C (buffer 1) and pelleted (20,000rpm, Sorvall RC5C centrifuge). The cell pellet was resuspended in buffer 1, homogenised using an Ultra-Turrax homogeniser and then pelleted as above. This was repeated once more, and the cells then resuspended in 5 mM Tris-Acetate, pH 7.4 at 25°C (buffer 3) at a protein concentration of 1 mg/ml. Radioligand binding was performed in 0.5 ml final volume buffer 3, containing 2 nM of [^{3} H]-dizocilpine and 100 μ g of protein. After 2 hours at 25°C in the presence of 30 μ M glycine and 30 μ M L-glutamate the membranes were washed with buffer 1 and bound radioactivity determined by scintillation counting. The recombinant R1a+R2A receptors bound [^{3} H]-dizocilpine at levels up to 200 fmols/mg protein. No binding was seen to untransfected Ltk⁻ cells, confirming that the [^{3} H]-dizocilpine was binding to recombinant R1a+R2A NMDA receptors. Non-specific binding was determined in a parallel incubation containing 100 μ M dizocilpine. The recombinant R1a+R2A

2148599

- 26 -

receptors bound also [^3H]-L-689,560 at levels of up to 500 fmols/mg protein, confirming the presence of a glycine binding site in population B cells.

5 b) Electrophysiology

10 The nature of the NMDA receptor expressed by population B 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 NMDA. Concentration response curves to glutamate in a saturating concentration of glycine (10 μM) gave a log EC₅₀ of 5.77, and a Hill coefficient of 1.1. The response to glutamate and glycine was antagonised by the competitive glutamate receptor antagonist CGS 19,755 (Research Biochemical International, MA, USA). All these electrophysiological data confirm that the recombinant R1a+R2A NMDA receptor expressed by population B cells has the properties expected of a bona fide NMDA receptor.

- 27 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5 (i) APPLICANT:

(A) NAME: Merck Sharp & Dohme Limited

(B) STREET: Hertford Road

(C) CITY: Hoddesdon

(D) STATE: Hertfordshire

10 (E) COUNTRY: England

(F) POSTAL CODE (ZIP): EN11 9BU

(ii) TITLE OF INVENTION: Stably transfected cell line
expressing NMDA receptors, and novel cloned NMDA receptor
15 subunit sequences

(iii) NUMBER OF SEQUENCES: 10

(iv) COMPUTER READABLE FORM:

20 (A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version
#1.25 (EPO)

25

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 2727 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- 28 -

(A) NAME/KEY: CDS

(B) LOCATION: 33..2621

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

5

GCCGCGCAGA GCCAGGCCCG CGGCCCCGAGC CC ATG AGC ACC ATG CGC CTG TTG 53

Met Ser Thr Met Arg Leu Leu

1

5

10

ACG CTC GCC CTG CTG TTC TCC TGC TCC CTC GCC CGT GCC GCG TGC GAC 101

Thr Leu Ala Leu Leu Phe Ser Cys Ser Leu Ala Arg Ala Ala Cys Asp

10

15

20

15

CCC AAG ATC GTC AAC ATT GGC GCG GTG CTG AGC ACG CGG AAG CAC GAG 149

Pro Lys Ile Val Asn Ile Gly Ala Val Leu Ser Thr Arg Lys His Glu

25

30

35

20

CAG ATG TTC CGC GAG GCC GTG AAC CAG GCC AAC AAG CGG CAC GGC TCC 197

Gln Met Phe Arg Glu Ala Val Asn Gln Ala Asn Lys Arg His Gly Ser

40

45

50

55

25

TGG AAG ATT CAG CTC AAT GCC ACC TCC GTC ACG CAC AAG CCC AAC GCC 245

Trp Lys Ile Gln Leu Asn Ala Thr Ser Val Thr His Lys Pro Asn Ala

60

65

70

ATC CAG ATG GCT CTG TCG GTG TGC GAG GAC CTC ATC TCC AGC CAG GTC 293

Ile Gln Met Ala Leu Ser Val Cys Glu Asp Leu Ile Ser Ser Gln Val

75

80

85

30

TAC GCC ATC CTA GTT AGC CAT CCA CCT ACC CCC AAC GAC CAC TTC ACT 341

Tyr Ala Ile Leu Val Ser His Pro Pro Thr Pro Asn Asp His Phe Thr

90

95

100

35

CCC ACC CCT GTC TCC TAC ACA GCC GGC TTC TAC CGC ATA CCC GTG CTG 389

Pro Thr Pro Val Ser Tyr Thr Ala Gly Phe Tyr Arg Ile Pro Val Leu

105

110

115

- 29 -

GGG CTG ACC ACC CGC ATG TCC ATC TAC TCG GAC AAG AGC ATC CAC CTG 437
 Gly Leu Thr Thr Arg Met Ser Ile Tyr Ser Asp Lys Ser Ile His Leu
 120 125 130 135

5 AGC TTC CTG CGC ACC GTG CCG CCC TAC TCC CAC CAG TCC AGC GTG TGG 485
 Ser Phe Leu Arg Thr Val Pro Pro Tyr Ser His Gln Ser Ser Val Trp
 140 145 150

10 TTT GAG ATG ATG CGT GTC TAC AGC TGG AAC CAC ATC ATC CTG CTG GTC 533
 Phe Glu Met Met Arg Val Tyr Ser Trp Asn His Ile Ile Leu Leu Val
 155 160 165

15 AGC GAC GAC CAC GAG GGC CGG GCG GCT CAG AAA CGC CTG GAG ACG CTG 581
 Ser Asp Asp His Glu Gly Arg Ala Ala Gln Lys Arg Leu Glu Thr Leu
 170 175 180

20 CTG GAG GAG CGT GAG TCC AAG GCA GAG AAG GTG CTG CAG TTT GAC CCA 629
 Leu Glu Glu Arg Glu Ser Lys Ala Glu Lys Val Leu Gln Phe Asp Pro
 185 190 195

GGG ACC AAG AAC GTG ACG GCC CTG CTG ATG GAG GCG AAA GAG CTG GAG 677
 Gly Thr Lys Asn Val Thr Ala Leu Leu Met Glu Ala Lys Glu Leu Glu
 200 205 210 215

25 GCC CGG GTC ATC ATC CTT TCT GCC AGC GAG GAC GAT GCT GCC ACT GTA 725
 Ala Arg Val Ile Ile Leu Ser Ala Ser Glu Asp Asp Ala Ala Thr Val
 220 225 230

30 TAC CGC GCA GCC GCG ATG CTG AAC ATG ACG GGC TCC GGG TAC GTG TGG 773
 Tyr Arg Ala Ala Ala Met Leu Asn Met Thr Gly Ser Gly Tyr Val Trp
 235 240 245

35 CTG GTC GGC GAG CGC GAG ATC TCG GGG AAC GCC CTG GCC TAC GCC CCA 821
 Leu Val Gly Glu Arg Glu Ile Ser Gly Asn Ala Leu Ala Tyr Ala Pro
 250 255 260

- 30 -

GAC GGC ATC CTC GGG CTG CAG CTC ATC AAC GGC AAG AAC GAG TCG GCC 869
Asp Gly Ile Leu Gly Leu Gln Leu Ile Asn Gly Lys Asn Glu Ser Ala
265 270 275

5 CAC ATC AGC GAC GCC GTG GGC GTG GTG GCC CAG GCC GTG CAC GAG CTC 917
His Ile Ser Asp Ala Val Gly Val Val Ala Gln Ala Val His Glu Leu
280 285 290 295

10 CTC GAG AAG GAG AAC ATC ACC GAC CCG CCG CGG GGC TGC GTG GGC AAC 965
Leu Glu Lys Glu Asn Ile Thr Asp Pro Pro Arg Gly Cys Val Gly Asn
300 305 310

15 ACC AAC ATC TGG AAG ACC GGG CCG CTC TTC AAG AGA GTG CTG ATG TCT 1013
Thr Asn Ile Trp Lys Thr Gly Pro Leu Phe Lys Arg Val Leu Met Ser
315 320 325

20 TCC AAG TAT GCG GAT GGG GTG ACT GGT CGC GTG GAG TTC AAT GAG GAT 1061
Ser Lys Tyr Ala Asp Gly Val Thr Gly Arg Val Glu Phe Asn Glu Asp
330 335 340

25 GGG GAC CGG AAG TTC GCC AAC TAC AGC ATC ATG AAC CTG CAG AAC CGC 1109
Gly Asp Arg Lys Phe Ala Asn Tyr Ser Ile Met Asn Leu Gln Asn Arg
345 350 355

30 AAG CTG GTG CAA GTG GGC ATC TAC AAT GGC ACC CAC GTC ATC CCT AAT 1157
Lys Leu Val Gln Val Gly Ile Tyr Asn Gly Thr His Val Ile Pro Asn
360 365 370 375

35 GAC AGG AAG ATC ATC TGG CCA GGC GGA GAG ACA GAG AAG CCT CGA GGC 1205
Asp Arg Lys Ile Ile Trp Pro Gly Gly Glu Thr Glu Lys Pro Arg Gly
380 385 390

40 TAC CAG ATG TCC ACC AGA CTG AAG ATT GTG ACG ATC CAC CAG GAG CCC 1253
Tyr Gln Met Ser Thr Arg Leu Lys Ile Val Thr Ile His Gln Glu Pro
395 400 405

- 31 -

TTC GTG TAC GTC AAG CCC ACG CTG AGT GAT GGG ACA TGC AAG GAG GAG 1301
 Phe Val Tyr Val Lys Pro Thr Leu Ser Asp Gly Thr Cys Lys Glu Glu
 410 415 420

5 TTC ACA GTC AAC GGC GAC CCA GTC AAG AAG GTG ATC TGC ACC GGG CCC 1349
 Phe Thr Val Asn Gly Asp Pro Val Lys Lys Val Ile Cys Thr Gly Pro
 425 430 435

10 AAC GAC ACG TCG CCG GGC AGC CCC CGC CAC ACG GTG CCT CAG TGT TGC 1397
 Asn Asp Thr Ser Pro Gly Ser Pro Arg His Thr Val Pro Gln Cys Cys
 440 445 450 455

15 TAC GGC TTT TGC ATC GAC CTG CTC ATC AAG CTG GCA CGG ACC ATG AAC 1445
 Tyr Gly Phe Cys Ile Asp Leu Leu Ile Lys Leu Ala Arg Thr Met Asn
 460 465 470

20 TTC ACC TAC GAG GTG CAC CTG GTG GCA GAT GGC AAG TTC GGC ACA CAG 1493
 Phe Thr Tyr Glu Val His Leu Val Ala Asp Gly Lys Phe Gly Thr Gln
 475 480 485

GAG CGG GTG AAC AAC AGC AAC AAG AAG GAG TGG AAT GGG ATG ATG GGC 1541
 Glu Arg Val Asn Asn Ser Asn Lys Lys Glu Trp Asn Gly Met Met Gly
 490 495 500

25 GAG CTG CTC AGC GGG CAG GCA GAC ATG ATC GTG GCG CCG CTA ACC ATA 1589
 Glu Leu Leu Ser Gly Gln Ala Asp Met Ile Val Ala Pro Leu Thr Ile
 505 510 515

30 AAC AAC GAG CGC GCG CAG TAC ATC GAG TTT TCC AAG CCC TTC AAG TAC 1637
 Asn Asn Glu Arg Ala Gln Tyr Ile Glu Phe Ser Lys Pro Phe Lys Tyr
 520 525 530 535

35 CAG GGC CTG ACT ATT CTG GTC AAG AAG GAG ATT CCC CGG AGC ACG CTG 1685
 Gln Gly Leu Thr Ile Leu Val Lys Lys Glu Ile Pro Arg Ser Thr Leu
 540 545 550

- 32 -

GAC TCG TTC ATG CAG CCG TTC CAG AGC ACA CTG TGG CTG CTG GTG GGG 1733

Asp Ser Phe Met Gln Pro Phe Gln Ser Thr Leu Trp Leu Leu Val Gly

555

560

565

5

CTG TCG GTG CAC GTG GTG GCC GTG ATG CTG TAC CTG CTG GAC CGC TTC 1781

Leu Ser Val His Val Val Ala Val Met Leu Tyr Leu Leu Asp Arg Phe

570

575

580

10

AGC CCC TTC GGC CGG TTC AAG GTG AAC AGC GAG GAG GAG GAG GAG GAC 1829

Ser Pro Phe Gly Arg Phe Lys Val Asn Ser Glu Glu Glu Glu Glu Asp

585

590

595

15

GCA CTG ACC CTG TCC TCG GCC ATG TGG TTC TCC TGG GGC GTC CTG CTC 1877

Ala Leu Thr Leu Ser Ser Ala Met Trp Phe Ser Trp Gly Val Leu Leu

600

605

610

615

20

AAC TCC GGC ATC GGG GAA GGC GCC CCC AGA AGC TTC TCA GCG CGC ATC 1925

Asn Ser Gly Ile Gly Glu Gly Ala Pro Arg Ser Phe Ser Ala Arg Ile

620

625

630

CTG GGC ATG GTG TGG GCC GGC TTT GCC ATG ATC ATC GTG GCC TCC TAC 1973

Leu Gly Met Val Trp Ala Gly Phe Ala Met Ile Ile Val Ala Ser Tyr

635

640

645

25

ACT GCC AAC TTG GCG GCC TTC CTG GTG CTG GAC CGG CCG GAG GAG CGC 2021

Thr Ala Asn Leu Ala Ala Phe Leu Val Leu Asp Arg Pro Glu Glu Arg

650

655

660

30

ATC ACG GGC ATC AAC GAC CCT CGG CTG AGG AAC CCC TCG GAC AAG TTT 2069

Ile Thr Gly Ile Asn Asp Pro Arg Leu Arg Asn Pro Ser Asp Lys Phe

665

670

675

35

ATC TAC GCC ACG GTG AAG CAG AGC TCC GTG GAT ATC TAC TTC CGG CGC 2117

Ile Tyr Ala Thr Val Lys Gln Ser Ser Val Asp Ile Tyr Phe Arg Arg

680

685

690

695

- 33 -

CAG GTG GAG CTG AGC ACC ATG TAC CGG CAT ATG GAG AAG CAC AAC TAC 2165
Gln Val Glu Leu Ser Thr Met Tyr Arg His Met Glu Lys His Asn Tyr

700

705

710

5 GAG AGT GCG GCG GAG GCC ATC CAG GCC GTG AGA GAC AAC AAG CTG CAT 2213
Glu Ser Ala Ala Glu Ala Ile Gln Ala Val Arg Asp Asn Lys Leu His

715

720

725

10 GCC TTC ATC TGG GAC TCG GCG GTG CTG GAG TTC GAG GCC TCG CAG AAG 2261
Ala Phe Ile Trp Asp Ser Ala Val Leu Glu Phe Glu Ala Ser Gln Lys

730

735

740

15 TGC GAC CTG GTG ACG ACT GGA GAG CTG TTT TTC CGC TCG GGC TTC GGC 2309
Cys Asp Leu Val Thr Thr Gly Glu Leu Phe Phe Arg Ser Gly Phe Gly

745

750

755

20 ATA GGC ATG CGC AAA GAC AGC CCC TGG AAG CAG AAC GTC TCC CTG TCC 2357
Ile Gly Met Arg Lys Asp Ser Pro Trp Lys Gln Asn Val Ser Leu Ser

760

765

770

775

ATC CTC AAG TCC CAC GAG AAT GGC TTC ATG GAA GAC CTG GAC AAG ACG 2405
Ile Leu Lys Ser His Glu Asn Gly Phe Met Glu Asp Leu Asp Lys Thr

780

785

790

25 TGG GTT CGG TAT CAG GAA TGT GAC TCG CGC AGC AAC GCC CCT GCA ACC 2453
Trp Val Arg Tyr Gln Glu Cys Asp Ser Arg Ser Asn Ala Pro Ala Thr

795

800

805

30 CTT ACT TTT GAG AAC ATG GCC GGG GTC TTC ATG CTG GTA GCT GGG GGC 2501
Leu Thr Phe Glu Asn Met Ala Gly Val Phe Met Leu Val Ala Gly Gly

810

815

820

ATC GTG GCC GGG ATC TTC CTG ATT TTC ATC GAG ATT GCC TAC AAG CGG 2549
Ile Val Ala Gly Ile Phe Leu Ile Phe Ile Glu Ile Ala Tyr Lys Arg

35

825

830

835

- 34 -

CAC AAG GAT GCT CGC CGG AAG CAG ATG CAG CTG GCC TTT GCC GCC GTT 2597
His Lys Asp Ala Arg Arg Lys Gln Met Gln Leu Ala Phe Ala Ala Val
840 845 850 855

5 AAC GTG TGG CGG AAG AAC CTG CAG 2621
Asn Val Trp Arg Lys Asn Leu Gln
860

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

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 885 amino acids

(B) TYPE: amino acid

15

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

20

Met Ser Thr Met Arg Leu Leu Thr Leu Ala Leu Leu Phe Ser Cys Ser

1 5 10 15

Leu Ala Arg Ala Ala Cys Asp Pro Lys Ile Val Asn Ile Gly Ala Val

25

20 25 30

Leu Ser Thr Arg Lys His Glu Gln Met Phe Arg Glu Ala Val Asn Gln

35 40 45

30

Ala Asn Lys Arg His Gly Ser Trp Lys Ile Gln Leu Asn Ala Thr Ser

50 55 60

Val Thr His Lys Pro Asn Ala Ile Gln Met Ala Leu Ser Val Cys Glu

65 70 75 80

35

Asp Leu Ile Ser Ser Gln Val Tyr Ala Ile Leu Val Ser His Pro Pro

85 90 95

- 35 -

Thr Pro Asn Asp His Phe Thr Pro Thr Pro Val Ser Tyr Thr Ala Gly
100 105 110

5 Phe Tyr Arg Ile Pro Val Leu Gly Leu Thr Thr Arg Met Ser Ile Tyr
115 120 125

Ser Asp Lys Ser Ile His Leu Ser Phe Leu Arg Thr Val Pro Pro Tyr
130 135 140

10 Ser His Gln Ser Ser Val Trp Phe Glu Met Met Arg Val Tyr Ser Trp
145 150 155 160

Asn His Ile Ile Leu Leu Val Ser Asp Asp His Glu Gly Arg Ala Ala
15 165 170 175

Gln Lys Arg Leu Glu Thr Leu Leu Glu Glu Arg Glu Ser Lys Ala Glu
180 185 190

20 Lys Val Leu Gln Phe Asp Pro Gly Thr Lys Asn Val Thr Ala Leu Leu
195 200 205

Met Glu Ala Lys Glu Leu Glu Ala Arg Val Ile Ile Leu Ser Ala Ser
210 215 220

25 Glu Asp Asp Ala Ala Thr Val Tyr Arg Ala Ala Ala Met Leu Asn Met
225 230 235 240

30 Thr Gly Ser Gly Tyr Val Trp Leu Val Gly Glu Arg Glu Ile Ser Gly
245 250 255

Asn Ala Leu Ala Tyr Ala Pro Asp Gly Ile Leu Gly Leu Gln Leu Ile
260 265 270

35 Asn Gly Lys Asn Glu Ser Ala His Ile Ser Asp Ala Val Gly Val Val
275 280 285

- 36 -

Ala Gln Ala Val His Glu Leu Leu Glu Lys Glu Asn Ile Thr Asp Pro

290

295

300

5

Pro Arg Gly Cys Val Gly Asn Thr Asn Ile Trp Lys Thr Gly Pro Leu

305

310

315

320

Phe Lys Arg Val Leu Met Ser Ser Lys Tyr Ala Asp Gly Val Thr Gly

325

330

335

10

Arg Val Glu Phe Asn Glu Asp Gly Asp Arg Lys Phe Ala Asn Tyr Ser

340

345

350

Ile Met Asn Leu Gln Asn Arg Lys Leu Val Gln Val Gly Ile Tyr Asn

15

355

360

365

Gly Thr His Val Ile Pro Asn Asp Arg Lys Ile Ile Trp Pro Gly Gly

370

375

380

20

Glu Thr Glu Lys Pro Arg Gly Tyr Gln Met Ser Thr Arg Leu Lys Ile

385

390

395

400

Val Thr Ile His Gln Glu Pro Phe Val Tyr Val Lys Pro Thr Leu Ser

405

410

415

25

Asp Gly Thr Cys Lys Glu Glu Phe Thr Val Asn Gly Asp Pro Val Lys

420

425

430

30

Lys Val Ile Cys Thr Gly Pro Asn Asp Thr Ser Pro Gly Ser Pro Arg

435

440

445

His Thr Val Pro Gln Cys Cys Tyr Gly Phe Cys Ile Asp Leu Leu Ile

450

455

460

35

Lys Leu Ala Arg Thr Met Asn Phe Thr Tyr Glu Val His Leu Val Ala

465

470

475

480

- 37 -

Asp Gly Lys Phe Gly Thr Gln Glu Arg Val Asn Asn Ser Asn Lys Lys
485 490 495

5 Glu Trp Asn Gly Met Met Gly Glu Leu Leu Ser Gly Gln Ala Asp Met
500 505 510

Ile Val Ala Pro Leu Thr Ile Asn Asn Glu Arg Ala Gln Tyr Ile Glu
515 520 525

10 Phe Ser Lys Pro Phe Lys Tyr Gln Gly Leu Thr Ile Leu Val Lys Lys
530 535 540

Glu Ile Pro Arg Ser Thr Leu Asp Ser Phe Met Gln Pro Phe Gln Ser
15 545 550 555 560

Thr Leu Trp Leu Leu Val Gly Leu Ser Val His Val Val Ala Val Met
565 570 575

20 Leu Tyr Leu Leu Asp Arg Phe Ser Pro Phe Gly Arg Phe Lys Val Asn
580 585 590

Ser Glu Glu Glu Glu Glu Asp Ala Leu Thr Leu Ser Ser Ala Met Trp
25 595 600 605

Phe Ser Trp Gly Val Leu Leu Asn Ser Gly Ile Gly Glu Gly Ala Pro
610 615 620

30 Arg Ser Phe Ser Ala Arg Ile Leu Gly Met Val Trp Ala Gly Phe Ala
625 630 635 640

Met Ile Ile Val Ala Ser Tyr Thr Ala Asn Leu Ala Ala Phe Leu Val
645 650 655

35 Leu Asp Arg Pro Glu Glu Arg Ile Thr Gly Ile Asn Asp Pro Arg Leu
660 665 670

- 38 -

Arg Asn Pro Ser Asp Lys Phe Ile Tyr Ala Thr Val Lys Gln Ser Ser
675 680 685

5 Val Asp Ile Tyr Phe Arg Arg Gln Val Glu Leu Ser Thr Met Tyr Arg
690 695 700

His Met Glu Lys His Asn Tyr Glu Ser Ala Ala Glu Ala Ile Gln Ala
705 710 715 720

10

Val Arg Asp Asn Lys Leu His Ala Phe Ile Trp Asp Ser Ala Val Leu
725 730 735

15

Glu Phe Glu Ala Ser Gln Lys Cys Asp Leu Val Thr Thr Gly Glu Leu
740 745 750

Phe Phe Arg Ser Gly Phe Gly Ile Gly Met Arg Lys Asp Ser Pro Trp
755 760 765

20

Lys Gln Asn Val Ser Leu Ser Ile Leu Lys Ser His Glu Asn Gly Phe
770 775 780

Met Glu Asp Leu Asp Lys Thr Trp Val Arg Tyr Gln Glu Cys Asp Ser
785 790 795 800

25

Arg Ser Asn Ala Pro Ala Thr Leu Thr Phe Glu Asn Met Ala Gly Val
805 810 815

30

Phe Met Leu Val Ala Gly Gly Ile Val Ala Gly Ile Phe Leu Ile Phe
820 825 830

Ile Glu Ile Ala Tyr Lys Arg His Lys Asp Ala Arg Arg Lys Gln Met
835 840 845

35

Gln Leu Ala Phe Ala Ala Val Asn Val Trp Arg Lys Asn Leu Gln
850 855 860

- 39 -

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 111 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- 15 (A) NAME/KEY: CDS
(B) LOCATION: 1..111

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

20 GAT AGA AAG AGT GGT AGA GCA GAG CCT GAC CCT AAA AAG AAA GCC ACA 48
Asp Arg Lys Ser Gly Arg Ala Glu Pro Asp Pro Lys Lys Lys Ala Thr
1 5 10 15
TTT AGG GCT ATC ACC TCC ACC CTG GCT TCC AGC TTC AAG AGG CGT AGG 96
25 Phe Arg Ala Ile Thr Ser Thr Leu Ala Ser Ser Phe Lys Arg Arg Arg
20 25 30
TCC TCC AAA GAC ACG 111
Ser Ser Lys Asp Thr
30 35

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 37 amino acids
(B) TYPE: amino acid

00220 - 40 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Asp Arg Lys Ser Gly Arg Ala Glu Pro Asp Pro Lys Lys Lys Ala Thr

1

5

10

15

10

Phe Arg Ala Ile Thr Ser Thr Leu Ala Ser Ser Phe Lys Arg Arg Arg

20

25

30

Ser Ser Lys Asp Thr

35

15

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

20

(A) LENGTH: 184 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

25

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..114

30

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

AGC ACC GGG GGT GGA CGC GGC GCT TTG CAA AAC CAA AAA GAC ACA GTG 48

Ser Thr Gly Gly Gly Arg Gly Ala Leu Gln Asn Gln Lys Asp Thr Val

35

1

5

10

15

- 41 -

CTG CCG CGA CGG CCT ATT GAG AGG GAG GAG GGC CAG CTG CAG CTG TGT 96
 Leu Pro Arg Arg Pro Ile Glu Arg Glu Glu Gly Gln Leu Gln Leu Cys
 20 25 30

5 TCC CGT CAT AGG GAG AGC TGAGACTCCC CGCCCGCCCT CCTCTGCCCC 144
 Ser Arg His Arg Glu Ser
 35

10 CCTCCCCCGC AGACAGACAG ACAGACGGAC GGGACAGCGG 184

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 38 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

20

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

Ser Thr Gly Gly Gly Arg Gly Ala Leu Gln Asn Gln Lys Asp Thr Val
 1 5 10 15
 25 Leu Pro Arg Arg Pro Ile Glu Arg Glu Glu Gly Gln Leu Gln Leu Cys
 20 25 30
 Ser Arg His Arg Glu Ser
 30 35

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 103 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

- 42 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5 (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

10

CAG TAC CAT CCC ACT GAT ATC ACG GGC CCG CTC AAC CTC TCA GAT CCC 48

Gln Tyr His Pro Thr Asp Ile Thr Gly Pro Leu Asn Leu Ser Asp Pro

1

5

10

15

15

TCG GTC AGC ACC GTG GTG TGAGGCC CCCGGAGCGC CCACCTGCCC AGTTTAGCCC 103

Ser Val Ser Thr Val Val

20

20

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

25

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

30

Gln Tyr His Pro Thr Asp Ile Thr Gly Pro Leu Asn Leu Ser Asp Pro

1

5

10

15

Ser Val Ser Thr Val

35

20

(2) INFORMATION FOR SEQ ID NO: 9:

- 43 -

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4858 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 47..4441

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CTGCATCCTC GACCTTCTCG GGCTACAGGG ACCGTCAGTG GCGACT ATG GGC AGA 55

Met Gly Arg

1

GTG GGC TAT TGG ACC CTG CTG GTG CTG CCG GCC CTT CTG GTC TGG CGC 103

Val Gly Tyr Trp Thr Leu Leu Val Leu Pro Ala Leu Leu Val Trp Arg

5

10

15

GGT CCG GCG CCG AGC GCG GCG GCG GAG AAG GGT CCC CCC GCG CTA AAT 151

Gly Pro Ala Pro Ser Ala Ala Ala Glu Lys Gly Pro Pro Ala Leu Asn

20

25

30

35

ATT GCG GTG ATG CTG GGT CAC AGC CAC GAC GTG ACA GAG CGC GAA CTT 199

Ile Ala Val Met Leu Gly His Ser His Asp Val Thr Glu Arg Glu Leu

40

45

50

CGA ACA CTG TGG GGC CCC GAG CAG GCG GGC GGG CTG CCC CTG GAC GTG 247

Arg Thr Leu Trp Gly Pro Glu Gln Ala Gly Gly Leu Pro Leu Asp Val

55

60

65

2148599

- 44 -

AAC GTG GTA GCT CTG CTG ATG AAC CGC ACC GAC CCC AAG AGC CTC ATC 295
 Asn Val Val Ala Leu Leu Met Asn Arg Thr Asp Pro Lys Ser Leu Ile

70

75

80

5

ACG CAC GTG TGC GAC CTC ATG TCC GGG GCA CGC ATC CAC GGC CTC GTG 343
 Thr His Val Cys Asp Leu Met Ser Gly Ala Arg Ile His Gly Leu Val

85

90

95

10

TTT GGG GAC GAC ACG GAC CAG GAG GTC GTA GCC CAG ATG CTG GAT TTT 391
 Phe Gly Asp Asp Thr Asp Gln Glu Val Val Ala Gln Met Leu Asp Phe
 100 105 110 115

ATC TCC TCC CAC ACC TTC GTC CCC ATC TTG GGC ATT CAT GGG GGC GCA 439
 Ile Ser Ser His Thr Phe Val Pro Ile Leu Gly Ile His Gly Gly Ala

15

120

125

130

TCT ATG ATC ATG GCT GAC AAG GAT CCG ACG TCT ACC TTC TTC CAG TTT 487
 Ser Met Ile Met Ala Asp Lys Asp Pro Thr Ser Thr Phe Phe Gln Phe

20

135

140

145

GGA GCG TCC ATC CAG CAG CAA GCC ACG GTC ATG CTG AAG ATC ATG CAG 535
 Gly Ala Ser Ile Gln Gln Gln Ala Thr Val Met Leu Lys Ile Met Gln

150

155

160

25

GAT TAT GAC TGG CAT GTC TTC TCC CTG GTG ACC ACT ATC TTC CCT GGC 583
 Asp Tyr Asp Trp His Val Phe Ser Leu Val Thr Thr Ile Phe Pro Gly

165

170

175

30

TAC AGG GAA TTC ATC AGC TTC GTC AAG ACC ACA GTG GAC AAC AGC TTT 631
 Tyr Arg Glu Phe Ile Ser Phe Val Lys Thr Thr Val Asp Asn Ser Phe
 180 185 190 195

GTG GGC TGG GAC ATG CAG AAT GTG ATC ACA CTG GAC ACT TCC TTT GAG 679
 Val Gly Trp Asp Met Gln Asn Val Ile Thr Leu Asp Thr Ser Phe Glu

35

200

205

210

- 45 -

GAT GCA AAG ACA CAA GTC CAG CTG AAG AAG ATC CAC TCT TCT GTC ATC 727
Asp Ala Lys Thr Gln Val Gln Leu Lys Lys Ile His Ser Ser Val Ile
215 220 225

5 TTG CTC TAC TGT TCC AAA GAC GAG GCT GTT CTC ATT CTG AGT GAG GCC 775
Leu Leu Tyr Cys Ser Lys Asp Glu Ala Val Leu Ile Leu Ser Glu Ala
230 235 240

10 CGC TCC CTT GGC GTC ACC GGG TAT GAT TTC TTC TGG ATT GTC CCC AGC 823
Arg Ser Leu Gly Val Thr Gly Tyr Asp Phe Phe Trp Ile Val Pro Ser
245 250 255

15 TTG GTC TCT GGG AAC ACG GAG CTC ATC CCA AAA GAG TTT CCA TCG GGA 871
Leu Val Ser Gly Asn Thr Glu Leu Ile Pro Lys Glu Phe Pro Ser Gly
260 265 270 275

20 CTC ATT TCT GTC TCC TAC GAT GAC TGG GAC TAC AGC CTG GAG GCG AGA 919
Leu Ile Ser Val Ser Tyr Asp Asp Trp Asp Tyr Ser Leu Glu Ala Arg
280 285 290

25 GTG AGG GAC GGC ATT GGC ATC CTA ACC ACC GCT GCA TCT TCT ATG CTG 967
Val Arg Asp Gly Ile Gly Ile Leu Thr Thr Ala Ala Ser Ser Met Leu
295 300 305

30 GAG AAG TTC TCC TAC ATC CCC GAG GCC AAG GCC AGC TGC TAC GGG CAG 1015
Glu Lys Phe Ser Tyr Ile Pro Glu Ala Lys Ala Ser Cys Tyr Gly Gln
310 315 320

35 ATG GAG AGG CCA GAG GTC CCG ATG CAC ACC TTG CAC CCA TTT ATG GTC 1063
Met Glu Arg Pro Glu Val Pro Met His Thr Leu His Pro Phe Met Val
325 330 335

AAT GTT ACA TGG GAT GGC AAA GAC TTA TCC TTC ACT GAG GAA GGC TAC 1111
Asn Val Thr Trp Asp Gly Lys Asp Leu Ser Phe Thr Glu Glu Gly Tyr
340 345 350 355

000000

- 46 -

CAG GTG CAC CCC AGG CTG GTG GTG ATT GTG CTG AAC AAA GAC CGG GAA 1159
Gln Val His Pro Arg Leu Val Val Ile Val Leu Asn Lys Asp Arg Glu
360 365 370

5 TGG GAA AAG GTG GCC AAG TGG AAG AAC AAT ACG CTG AGC CTG AGC CAC 1207
Trp Glu Lys Val Ala Lys Trp Lys Asn Asn Thr Leu Ser Leu Ser His
375 380 385

10 GCC GTG TGG CCC AGG TAC AAG TCC TTC TCC GAC TGT GAG CCG GAT GAC 1255
Ala Val Trp Pro Arg Tyr Lys Ser Phe Ser Asp Cys Glu Pro Asp Asp
390 395 400

15 AAC CAT CTC AGC ATC GTC ACC CTG GAG GAG GCC CCA TTC GTC ATC GTG 1303
Asn His Leu Ser Ile Val Thr Leu Glu Glu Ala Pro Phe Val Ile Val
405 410 415

20 GAA GAC ATA GAC CCC CTG ACC GAG ACG TGT GTG AGG AAC ACC GTG CCA 1351
Glu Asp Ile Asp Pro Leu Thr Glu Thr Cys Val Arg Asn Thr Val Pro
420 425 430 435

TGT CGG AAG TTC GTC AAA ATC AAC AAT TCA ACC AAT GAG GGG ATG AAT 1399
Cys Arg Lys Phe Val Lys Ile Asn Asn Ser Thr Asn Glu Gly Met Asn
440 445 450

25 GTG AAG AAA TGC TGC AAG GGG TTC TGC ATT GAT ATT CTG AAG AAG CTT 1447
Val Lys Lys Cys Cys Lys Gly Phe Cys Ile Asp Ile Leu Lys Lys Leu
455 460 465

30 TCC AGA ACT GTG AAG TTT ACT TAC GAC CTC TAT CTG GTG ACC AAT GGG 1495
Ser Arg Thr Val Lys Phe Thr Tyr Asp Leu Tyr Leu Val Thr Asn Gly
470 475 480

35 AAG CAT GGC AAG AAA GTT AAC AAT GTG TGG AAT GGA ATG ATC GGT GAA 1543
Lys His Gly Lys Lys Val Asn Asn Val Trp Asn Gly Met Ile Gly Glu
485 490 495

- 47 -

GTG GTC TAT CAA CGG GCA GTC ATG GCA GTT GGC TCG CTC ACC ATC AAT 1591
 Val Val Tyr Gln Arg Ala Val Met Ala Val Gly Ser Leu Thr Ile Asn
 500 505 510 515

5 GAG GAA CGT TCT GAA GTG GTG GAC TTC TCT GTG CCC TTT GTG GAA ACG 1639
 Glu Glu Arg Ser Glu Val Val Asp Phe Ser Val Pro Phe Val Glu Thr
 520 525 530

10 GGA ATC AGT GTC ATG GTT TCA AGA AGT AAT GGC ACC GTC TCA CCT TCT 1687
 Gly Ile Ser Val Met Val Ser Arg Ser Asn Gly Thr Val Ser Pro Ser
 535 540 545

15 GCT TTT CTA GAA CCA TTC AGC GCC TCT GTC TGG GTG ATG ATG TTT GTG 1735
 Ala Phe Leu Glu Pro Phe Ser Ala Ser Val Trp Val Met Met Phe Val
 550 555 560

20 ATG CTG CTC ATT GTT TCT GCC ATA GCT GTT TTT GTC TTT GAA TAC TTC 1783
 Met Leu Leu Ile Val Ser Ala Ile Ala Val Phe Val Phe Glu Tyr Phe
 565 570 575

AGC CCT GTT GGA TAC AAC AGA AAC TTA GCC AAA GGG AAA GCA CCC CAT 1831
 Ser Pro Val Gly Tyr Asn Arg Asn Leu Ala Lys Gly Lys Ala Pro His
 580 585 590 595

25 GGG CCT TCT TTT ACA ATT GGA AAA GCT ATA TGG CTT CTT TGG GGC CTG 1879
 Gly Pro Ser Phe Thr Ile Gly Lys Ala Ile Trp Leu Leu Trp Gly Leu
 600 605 610

30 GTG TTC AAT AAC TCC GTG CCT GTC CAG AAT CCT AAA GGG ACC ACC AGC 1927
 Val Phe Asn Asn Ser Val Pro Val Gln Asn Pro Lys Gly Thr Thr Ser
 615 620 625

35 AAG ATC ATG GTA TCT GTA TGG GCC TTC TTC GCT GTC ATA TTC CTG GCT 1975
 Lys Ile Met Val Ser Val Trp Ala Phe Phe Ala Val Ile Phe Leu Ala
 630 635 640

2148599

- 48 -

AGC TAC ACA GCC AAT CTG GCT GCC TTC ATG ATC CAA GAG GAA TTT GTG 2023
 Ser Tyr Thr Ala Asn Leu Ala Ala Phe Met Ile Gln Glu Glu Phe Val
 645 650 655

5 GAC CAA GTG ACC GGC CTC AGT GAC AAA AAG TTT CAG AGA CCT CAT GAC 2071
 Asp Gln Val Thr Gly Leu Ser Asp Lys Lys Phe Gln Arg Pro His Asp
 660 665 670 675

10 TAT TCC CCA CCT TTT CGA TTT GGG ACA GTG CCT AAT GGA AGC ACG GAG 2119
 Tyr Ser Pro Pro Phe Arg Phe Gly Thr Val Pro Asn Gly Ser Thr Glu
 680 685 690

15 AGA AAC ATT CGG AAT AAC TAT CCC TAC ATG CAT CAG TAC ATG ACC AAA 2167
 Arg Asn Ile Arg Asn Asn Tyr Pro Tyr Met His Gln Tyr Met Thr Lys
 695 700 705

20 TTT AAT CAG AAA GGA GTA GAG GAC GCC TTG GTC AGC CTG AAA ACG GGG 2215
 Phe Asn Gln Lys Gly Val Glu Asp Ala Leu Val Ser Leu Lys Thr Gly
 710 715 720

AAG CTG GAC GCT TTC ATC TAC GAT GCC GCA GTC TTG AAT TAC AAG GCT 2263
 Lys Leu Asp Ala Phe Ile Tyr Asp Ala Ala Val Leu Asn Tyr Lys Ala
 725 730 735

25 GGG AGG GAT GAA GGC TGC AAA CTG GTG ACC ATC GGG AGT GGG TAC ATC 2311
 Gly Arg Asp Glu Gly Cys Lys Leu Val Thr Ile Gly Ser Gly Tyr Ile
 740 745 750 755

30 TTT GCC ACC ACC AGT TAT GGA ATT GCC CTT CAG AAA GGC TCT CCT TGG 2359
 Phe Ala Thr Thr Ser Tyr Gly Ile Ala Leu Gln Lys Gly Ser Pro Trp
 760 765 770

AAG AGG CAG ATC GAC CTG GCC TTG CTT CAG TTT GTG GGT GAT GGT GAG 2407
 Lys Arg Gln Ile Asp Leu Ala Leu Leu Gln Phe Val Gly Asp Gly Glu
 775 780 785

- 49 -

ATG GAG GAG CTG GAG ACC CTG TGG CTC ACT GGG ATC GTC CAC AAC GAG 2455

Met Glu Glu Leu Glu Thr Leu Trp Leu Thr Gly Ile Val His Asn Glu

790

795

800

5

AAG AAC GAG GTG ATG AGC AAC CAG CTG GAC ATT GAC AAC ATG GCG GGC 2503

Lys Asn Glu Val Met Ser Asn Gln Leu Asp Ile Asp Asn Met Ala Gly

805

810

815

GTA TTC TAC ATG CTG GCT GCC GCC ATG GCC CTT AGC CTC ATC ACC TTC 2551

10

Val Phe Tyr Met Leu Ala Ala Ala Met Ala Leu Ser Leu Ile Thr Phe

820

825

830

835

ATC TGG GAG CAC CTC TTC TAC TGG AAG CTG CGC TTC TGT TTC ACG GGC 2599

Ile Trp Glu His Leu Phe Tyr Trp Lys Leu Arg Phe Cys Phe Thr Gly

15

840

845

850

GTG TGC TCC GAC CGG CCT GGG TTG CTC TTC TCC ATC AGC AGG GGC ATC 2647

Val Cys Ser Asp Arg Pro Gly Leu Leu Phe Ser Ile Ser Arg Gly Ile

855

860

865

20

TAC AGC TGC ATT CAT GGA GTG CAC ATT GAA GAA AAG AAG AAG TCT CCA 2695

Tyr Ser Cys Ile His Gly Val His Ile Glu Glu Lys Lys Lys Ser Pro

870

875

880

25

GAC TTC AAT CTG ACG GGA TCC CAG AGC AAC ATG TTA AAA CTC CTC CGG 2743

Asp Phe Asn Leu Thr Gly Ser Gln Ser Asn Met Leu Lys Leu Leu Arg

885

890

895

TCA GCC AAA AAC ATT TCC AGC ATG TCC AAC ATG AAC TCC TCA AGA ATG 2791

30

Ser Ala Lys Asn Ile Ser Ser Met Ser Asn Met Asn Ser Ser Arg Met

900

905

910

915

GAC TCA CCC AAA AGA GCT GCT GAC TTC ATC CAA AGA GGT TCC CTC ATC 2839

Asp Ser Pro Lys Arg Ala Ala Asp Phe Ile Gln Arg Gly Ser Leu Ile

35

920

925

930

- 50 -

ATG GAC ATG GTT TCA GAT AAG GGG AAT TTG ATG TAC TCA GAC AAC AGG 2887

Met Asp Met Val Ser Asp Lys Gly Asn Leu Met Tyr Ser Asp Asn Arg

935

940

945

5

TCC TTT CAG GGG AAA GAG AGC ATT TTT GGA GAC AAC ATG AAC GAA CTC 2935

Ser Phe Gln Gly Lys Glu Ser Ile Phe Gly Asp Asn Met Asn Glu Leu

950

955

960

10

CAA ACA TTT GTG GCC AAC CGG CAG AAG GAT AAC CTC AAT AAC TAT GTA 2983

Gln Thr Phe Val Ala Asn Arg Gln Lys Asp Asn Leu Asn Asn Tyr Val

965

970

975

15

TTC CAG GGA CAA CAT CCT CTT ACT CTC AAT GAG TCC AAC CCT AAC ACG 3031

Phe Gln Gly Gln His Pro Leu Thr Leu Asn Glu Ser Asn Pro Asn Thr

980

985

990

995

20

GTG GAG GTG GCC GTG AGC ACA GAA TCC AAA GCG AAC TCT AGA CCC CGG 3079

Val Glu Val Ala Val Ser Thr Glu Ser Lys Ala Asn Ser Arg Pro Arg

1000

1005

1010

CAG CTG TGG AAG AAA TCC GTG GAT TCC ATA CGC CAG GAT TCA CTA TCC 3127

Gln Leu Trp Lys Lys Ser Val Asp Ser Ile Arg Gln Asp Ser Leu Ser

1015

1020

1025

25

CAG AAT CCA GTC TCC CAG AGG GAT GAG GCA ACA GCA GAG AAT AGG ACC 3175

Gln Asn Pro Val Ser Gln Arg Asp Glu Ala Thr Ala Glu Asn Arg Thr

1030

1035

1040

30

CAC TCC CTA AAG AGC CCT AGG TAT CTT CCA GAA GAG ATG GCC CAC TCT 3223

His Ser Leu Lys Ser Pro Arg Tyr Leu Pro Glu Glu Met Ala His Ser

1045

1050

1055

35

GAC ATT TCA GAA ACG TCA AAT CGG GCC ACG TGC CAC AGG GAA CCT GAC 3271

Asp Ile Ser Glu Thr Ser Asn Arg Ala Thr Cys His Arg Glu Pro Asp

1060

1065

1070

1075

- 51 -

AAC AGT AAG AAC CAC AAA ACC AAG GAC AAC TTT AAA AGG TCA GTG GCC 3319
 Asn Ser Lys Asn His Lys Thr Lys Asp Asn Phe Lys Arg Ser Val Ala

1080

1085

1090

5

TCC AAA TAC CCC AAG GAC TGT AGT GAG GTC GAG CGC ACC TAC CTG AAA 3367
 Ser Lys Tyr Pro Lys Asp Cys Ser Glu Val Glu Arg Thr Tyr Leu Lys

1095

1100

1105

10

ACC AAA TCA AGC TCC CCT AGA GAC AAG ATC TAC ACT ATA GAT GGT GAG 3415
 Thr Lys Ser Ser Ser Pro Arg Asp Lys Ile Tyr Thr Ile Asp Gly Glu

1110

1115

1120

15

AAG GAG CCT GGT TTC CAC TTA GAT CCA CCC CAG TTT GTT GAA AAT GTG 3463
 Lys Glu Pro Gly Phe His Leu Asp Pro Pro Gln Phe Val Glu Asn Val

1125

1130

1135

20

ACC CTG CCC GAG AAC GTG GAC TTC CCG GAC CCC TAC CAG GAT CCC AGT 3511
 Thr Leu Pro Glu Asn Val Asp Phe Pro Asp Pro Tyr Gln Asp Pro Ser

1140

1145

1150

1155

GAA AAC TTC CGC AAG GGG GAC TCC ACG CTG CCA ATG AAC CGG AAC CCC 3559
 Glu Asn Phe Arg Lys Gly Asp Ser Thr Leu Pro Met Asn Arg Asn Pro

1160

1165

1170

25

TTG CAT AAT GAA GAG GGG CTT TCC AAC AAC GAC CAG TAT AAA CTC TAC 3607
 Leu His Asn Glu Glu Gly Leu Ser Asn Asn Asp Gln Tyr Lys Leu Tyr

1175

1180

1185

30

TCC AAG CAC TTC ACC TTG AAA GAC AAG GGT TCC CCG CAC AGT GAG ACC 3655
 Ser Lys His Phe Thr Leu Lys Asp Lys Gly Ser Pro His Ser Glu Thr

1190

1195

1200

35

AGC GAG CGA TAC CGG CAG AAC TCC ACG CAC TGC AGA AGC TGC CTT TCC 3703
 Ser Glu Arg Tyr Arg Gln Asn Ser Thr His Cys Arg Ser Cys Leu Ser

1205

1210

1215

- 52 -

AAC ATG CCC ACC TAT TCA GGC CAC TTC ACC ATG AGG TCC CCC TTC AAG 3751

Asn Met Pro Thr Tyr Ser Gly His Phe Thr Met Arg Ser Pro Phe Lys

1220

1225

1230

1235

5

TGC GAT GCC TGC CTG CGG ATG GGG AAC CTC TAT GAC ATC GAT GAA GAC 3799

Cys Asp Ala Cys Leu Arg Met Gly Asn Leu Tyr Asp Ile Asp Glu Asp

1240

1245

1250

10

CAG ATG CTT CAG GAG ACA GGT AAC CCA GCC ACC GGG GAG GAG GTC TAC 3847

Gln Met Leu Gln Glu Thr Gly Asn Pro Ala Thr Gly Glu Glu Val Tyr

1255

1260

1265

CAG CAG GAC TGG GCA CAG AAC AAT GCC CTT CAA TTA CAA AAG AAC AAG 3895

Gln Gln Asp Trp Ala Gln Asn Asn Ala Leu Gln Leu Gln Lys Asn Lys

15

1270

1275

1280

CTA AGG ATT AGC CGT CAG CAT TCC TAC GAT AAC ATT GTC GAC AAA CCT 3943

Leu Arg Ile Ser Arg Gln His Ser Tyr Asp Asn Ile Val Asp Lys Pro

1285

1290

1295

20

AGG GAG CTA GAC CTT AGC AGG CCC TCC CGG AGC ATA AGC CTC AAG GAC 3991

Arg Glu Leu Asp Leu Ser Arg Pro Ser Arg Ser Ile Ser Leu Lys Asp

1300

1305

1310

1315

25

AGG GAA CGG CTT CTG GAG GGA AAT TTT TAC GGC AGC CTG TTT AGT GTC 4039

Arg Glu Arg Leu Leu Glu Gly Asn Phe Tyr Gly Ser Leu Phe Ser Val

1320

1325

1330

CCC TCA AGC AAA CTC TCG GGG AAA AAA AGC TCC CTT TTC CCC CAA GGT 4087

Pro Ser Ser Lys Leu Ser Gly Lys Lys Ser Ser Leu Phe Pro Gln Gly

30

1335

1340

1345

CTG GAG GAC AGC AAG AGG AGC AAG TCT CTC TTG CCA GAC CAC ACC TCC 4135

Leu Glu Asp Ser Lys Arg Ser Lys Ser Leu Leu Pro Asp His Thr Ser

35

1350

1355

1360

- 53 -

GAT AAC CCT TTC CTC CAC TCC CAC AGG GAT GAC CAA CGC TTG GGT ATT 4183
Asp Asn Pro Phe Leu His Ser His Arg Asp Asp Gln Arg Leu Gly Ile
1365 1370 1375

5 GGG AGA TGC CCC TCG GAC CCT TAC AAA CAC TCG TTG CCA TCC CAG GCG 4231
Gly Arg Cys Pro Ser Asp Pro Tyr Lys His Ser Leu Pro Ser Gln Ala
1380 1385 1390 1395

10 GTG AAT GAC AGC TAT CTT CGG TCG TCC TTG AGG TCA ACG GCA TCG TAC 4279
Val Asn Asp Ser Tyr Leu Arg Ser Ser Leu Arg Ser Thr Ala Ser Tyr
1400 1405 1410

15 TGT TCC AGG GAC AGT CGG GGC CAC AAT GAT GTG TAT ATT TCG GAG CAT 4327
Cys Ser Arg Asp Ser Arg Gly His Asn Asp Val Tyr Ile Ser Glu His
1415 1420 1425

GTT ATG CCT TAT GCT GCA AAT AAG AAT AAT ATG TAC TCT ACC CCC AGG 4375
Val Met Pro Tyr Ala Ala Asn Lys Asn Asn Met Tyr Ser Thr Pro Arg
1430 1435 1440

20 GTT TTA AAT TCC TGC AGC AAT AGA CGC GTG TAC AAG AAA ATG CCT AGT 4423
Val Leu Asn Ser Cys Ser Asn Arg Arg Val Tyr Lys Lys Met Pro Ser
1445 1450 1455

25 ATC GAA TCT GAT GTT TAAAAATCTT CCATTAATGT TTTATCTATA GGGAAATATA 4478
Ile Glu Ser Asp Val
1460 146

30 CGTAATGCCC AATGTTCTGG AGGGTAAATG TTGAATGTCC AATAGTGCCC TGCTAAGAGG 4538

AAGAAGATGT AGGGAGGTAT TTTGTTGTTG TTGTTGTTGG CTCTTTTGCA CACGGCTTCA 4598

35 TGCCATAATC TTCCACTCAA GGAATCTTGT GAGGTGTGTG CTGAGCATGG CAGACACCAG 4658

- 54 -

ATAGGTGAGT CCTTAACCAA AAATAACTAA CTACATAAGG GCAAGTCTCC GGGACATGCC

4718

TACTGGGTAT GTTGCCAATA ATGATGCATT GGATGCCAAT GGTGATGTTA TGATTTCCTA

5

4778

TATTCCAAAT TCCATTAAGG TCAGCCCACC ATGTAATTTT CTCATCAGAA ATGCCTAATG

4838

10

GTTTCTCTAA TACAGAATAA

4858

(2) INFORMATION FOR SEQ ID NO: 10:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1464 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Gly Arg Val Gly Tyr Trp Thr Leu Leu Val Leu Pro Ala Leu Leu

25

1 5 10 15

Val Trp Arg Gly Pro Ala Pro Ser Ala Ala Ala Glu Lys Gly Pro Pro

20 25 30

30

Ala Leu Asn Ile Ala Val Met Leu Gly His Ser His Asp Val Thr Glu

35 40 45

Arg Glu Leu Arg Thr Leu Trp Gly Pro Glu Gln Ala Gly Gly Leu Pro

50 55 60

35

Leu Asp Val Asn Val Val Ala Leu Leu Met Asn Arg Thr Asp Pro Lys

65 70 75 80

- 55 -

Ser Leu Ile Thr His Val Cys Asp Leu Met Ser Gly Ala Arg Ile His

85

90

95

5

Gly Leu Val Phe Gly Asp Asp Thr Asp Gln Glu Val Val Ala Gln Met

100

105

110

Leu Asp Phe Ile Ser Ser His Thr Phe Val Pro Ile Leu Gly Ile His

115

120

125

10

Gly Gly Ala Ser Met Ile Met Ala Asp Lys Asp Pro Thr Ser Thr Phe

130

135

140

Phe Gln Phe Gly Ala Ser Ile Gln Gln Gln Ala Thr Val Met Leu Lys

15

145

150

155

160

Ile Met Gln Asp Tyr Asp Trp His Val Phe Ser Leu Val Thr Thr Ile

165

170

175

20

Phe Pro Gly Tyr Arg Glu Phe Ile Ser Phe Val Lys Thr Thr Val Asp

180

185

190

Asn Ser Phe Val Gly Trp Asp Met Gln Asn Val Ile Thr Leu Asp Thr

25

195

200

205

Ser Phe Glu Asp Ala Lys Thr Gln Val Gln Leu Lys Lys Ile His Ser

210

215

220

30

Ser Val Ile Leu Leu Tyr Cys Ser Lys Asp Glu Ala Val Leu Ile Leu

225

230

235

240

Ser Glu Ala Arg Ser Leu Gly Val Thr Gly Tyr Asp Phe Phe Trp Ile

245

250

255

35

Val Pro Ser Leu Val Ser Gly Asn Thr Glu Leu Ile Pro Lys Glu Phe

260

265

270

- 56 -

Pro Ser Gly Leu Ile Ser Val Ser Tyr Asp Asp Trp Asp Tyr Ser Leu

275

280

285

5

Glu Ala Arg Val Arg Asp Gly Ile Gly Ile Leu Thr Thr Ala Ala Ser

290

295

300

Ser Met Leu Glu Lys Phe Ser Tyr Ile Pro Glu Ala Lys Ala Ser Cys

305

310

315

320

10

Tyr Gly Gln Met Glu Arg Pro Glu Val Pro Met His Thr Leu His Pro

325

330

335

15

Phe Met Val Asn Val Thr Trp Asp Gly Lys Asp Leu Ser Phe Thr Glu

340

345

350

Glu Gly Tyr Gln Val His Pro Arg Leu Val Val Ile Val Leu Asn Lys

355

360

365

20

Asp Arg Glu Trp Glu Lys Val Ala Lys Trp Lys Asn Asn Thr Leu Ser

370

375

380

25

Leu Ser His Ala Val Trp Pro Arg Tyr Lys Ser Phe Ser Asp Cys Glu

385

390

395

400

Pro Asp Asp Asn His Leu Ser Ile Val Thr Leu Glu Glu Ala Pro Phe

405

410

415

30

Val Ile Val Glu Asp Ile Asp Pro Leu Thr Glu Thr Cys Val Arg Asn

420

425

430

Thr Val Pro Cys Arg Lys Phe Val Lys Ile Asn Asn Ser Thr Asn Glu

35

435

440

445

- 57 -

Gly Met Asn Val Lys Lys Cys Cys Lys Gly Phe Cys Ile Asp Ile Leu

450

455

460

Lys Lys Leu Ser Arg Thr Val Lys Phe Thr Tyr Asp Leu Tyr Leu Val

5

465

470

475

480

Thr Asn Gly Lys His Gly Lys Lys Val Asn Asn Val Trp Asn Gly Met

485

490

495

10

Ile Gly Glu Val Val Tyr Gln Arg Ala Val Met Ala Val Gly Ser Leu

500

505

510

Thr Ile Asn Glu Glu Arg Ser Glu Val Val Asp Phe Ser Val Pro Phe

515

520

525

15

Val Glu Thr Gly Ile Ser Val Met Val Ser Arg Ser Asn Gly Thr Val

530

535

540

Ser Pro Ser Ala Phe Leu Glu Pro Phe Ser Ala Ser Val Trp Val Met

20

545

550

555

560

Met Phe Val Met Leu Leu Ile Val Ser Ala Ile Ala Val Phe Val Phe

565

570

575

25

Glu Tyr Phe Ser Pro Val Gly Tyr Asn Arg Asn Leu Ala Lys Gly Lys

580

585

590

Ala Pro His Gly Pro Ser Phe Thr Ile Gly Lys Ala Ile Trp Leu Leu

595

600

605

30

Trp Gly Leu Val Phe Asn Asn Ser Val Pro Val Gln Asn Pro Lys Gly

610

615

620

Thr Thr Ser Lys Ile Met Val Ser Val Trp Ala Phe Phe Ala Val Ile

35

625

630

635

640

- 58 -

Phe Leu Ala Ser Tyr Thr Ala Asn Leu Ala Ala Phe Met Ile Gln Glu
645 650 655

5 Glu Phe Val Asp Gln Val Thr Gly Leu Ser Asp Lys Lys Phe Gln Arg
660 665 670

Pro His Asp Tyr Ser Pro Pro Phe Arg Phe Gly Thr Val Pro Asn Gly
675 680 685

10 Ser Thr Glu Arg Asn Ile Arg Asn Asn Tyr Pro Tyr Met His Gln Tyr
690 695 700

Met Thr Lys Phe Asn Gln Lys Gly Val Glu Asp Ala Leu Val Ser Leu
705 710 715 720

15 Lys Thr Gly Lys Leu Asp Ala Phe Ile Tyr Asp Ala Ala Val Leu Asn
725 730 735

20 Tyr Lys Ala Gly Arg Asp Glu Gly Cys Lys Leu Val Thr Ile Gly Ser
740 745 750

Gly Tyr Ile Phe Ala Thr Thr Ser Tyr Gly Ile Ala Leu Gln Lys Gly
755 760 765

25 Ser Pro Trp Lys Arg Gln Ile Asp Leu Ala Leu Leu Gln Phe Val Gly
770 775 780

30 Asp Gly Glu Met Glu Glu Leu Glu Thr Leu Trp Leu Thr Gly Ile Val
785 790 795 800

His Asn Glu Lys Asn Glu Val Met Ser Asn Gln Leu Asp Ile Asp Asn
805 810 815

35 Met Ala Gly Val Phe Tyr Met Leu Ala Ala Ala Met Ala Leu Ser Leu
820 825 830

- 59 -

Ile Thr Phe Ile Trp Glu His Leu Phe Tyr Trp Lys Leu Arg Phe Cys
835 840 845

5 Phe Thr Gly Val Cys Ser Asp Arg Pro Gly Leu Leu Phe Ser Ile Ser
850 855 860

Arg Gly Ile Tyr Ser Cys Ile His Gly Val His Ile Glu Glu Lys Lys
865 870 875 880

10 Lys Ser Pro Asp Phe Asn Leu Thr Gly Ser Gln Ser Asn Met Leu Lys
885 890 895

Leu Leu Arg Ser Ala Lys Asn Ile Ser Ser Met Ser Asn Met Asn Ser
900 905 910

15 Ser Arg Met Asp Ser Pro Lys Arg Ala Ala Asp Phe Ile Gln Arg Gly
915 920 925

20 Ser Leu Ile Met Asp Met Val Ser Asp Lys Gly Asn Leu Met Tyr Ser
930 935 940

Asp Asn Arg Ser Phe Gln Gly Lys Glu Ser Ile Phe Gly Asp Asn Met
945 950 955 960

25 Asn Glu Leu Gln Thr Phe Val Ala Asn Arg Gln Lys Asp Asn Leu Asn
965 970 975

Asn Tyr Val Phe Gln Gly Gln His Pro Leu Thr Leu Asn Glu Ser Asn
980 985 990

30 Pro Asn Thr Val Glu Val Ala Val Ser Thr Glu Ser Lys Ala Asn Ser
995 1000 1005

35 Arg Pro Arg Gln Leu Trp Lys Lys Ser Val Asp Ser Ile Arg Gln Asp
1010 1015 1020

- 60 -

Ser Leu Ser Gln Asn Pro Val Ser Gln Arg Asp Glu Ala Thr Ala Glu
1025 1030 1035 1040

5 Asn Arg Thr His Ser Leu Lys Ser Pro Arg Tyr Leu Pro Glu Glu Met
1045 1050 1055

Ala His Ser Asp Ile Ser Glu Thr Ser Asn Arg Ala Thr Cys His Arg
1060 1065 1070

10 Glu Pro Asp Asn Ser Lys Asn His Lys Thr Lys Asp Asn Phe Lys Arg
1075 1080 1085

15 Ser Val Ala Ser Lys Tyr Pro Lys Asp Cys Ser Glu Val Glu Arg Thr
1090 1095 1100

Tyr Leu Lys Thr Lys Ser Ser Ser Pro Arg Asp Lys Ile Tyr Thr Ile
1105 1110 1115 1120

20 Asp Gly Glu Lys Glu Pro Gly Phe His Leu Asp Pro Pro Gln Phe Val
1125 1130 1135

Glu Asn Val Thr Leu Pro Glu Asn Val Asp Phe Pro Asp Pro Tyr Gln
1140 1145 1150

25 Asp Pro Ser Glu Asn Phe Arg Lys Gly Asp Ser Thr Leu Pro Met Asn
1155 1160 1165

Arg Asn Pro Leu His Asn Glu Glu Gly Leu Ser Asn Asn Asp Gln Tyr
1170 1175 1180

30 Lys Leu Tyr Ser Lys His Phe Thr Leu Lys Asp Lys Gly Ser Pro His
1185 1190 1195 1200

35 Ser Glu Thr Ser Glu Arg Tyr Arg Gln Asn Ser Thr His Cys Arg Ser
1205 1210 1215

- 61 -

Cys Leu Ser Asn Met Pro Thr Tyr Ser Gly His Phe Thr Met Arg Ser

1220

1225

1230

Pro Phe Lys Cys Asp Ala Cys Leu Arg Met Gly Asn Leu Tyr Asp Ile

5

1235

1240

1245

Asp Glu Asp Gln Met Leu Gln Glu Thr Gly Asn Pro Ala Thr Gly Glu

1250

1255

1260

10

Glu Val Tyr Gln Gln Asp Trp Ala Gln Asn Asn Ala Leu Gln Leu Gln

1265

1270

1275

1280

Lys Asn Lys Leu Arg Ile Ser Arg Gln His Ser Tyr Asp Asn Ile Val

1285

1290

1295

15

Asp Lys Pro Arg Glu Leu Asp Leu Ser Arg Pro Ser Arg Ser Ile Ser

1300

1305

1310

Leu Lys Asp Arg Glu Arg Leu Leu Glu Gly Asn Phe Tyr Gly Ser Leu

20

1315

1320

1325

Phe Ser Val Pro Ser Ser Lys Leu Ser Gly Lys Lys Ser Ser Leu Phe

1330

1335

1340

25

Pro Gln Gly Leu Glu Asp Ser Lys Arg Ser Lys Ser Leu Leu Pro Asp

1345

1350

1355

1360

His Thr Ser Asp Asn Pro Phe Leu His Ser His Arg Asp Asp Gln Arg

1365

1370

1375

30

Leu Gly Ile Gly Arg Cys Pro Ser Asp Pro Tyr Lys His Ser Leu Pro

1380

1385

1390

Ser Gln Ala Val Asn Asp Ser Tyr Leu Arg Ser Ser Leu Arg Ser Thr

35

1395

1400

1405

- 62 -

Ala Ser Tyr Cys Ser Arg Asp Ser Arg Gly His Asn Asp Val Tyr Ile

1410

1415

1420

Ser Glu His Val Met Pro Tyr Ala Ala Asn Lys Asn Asn Met Tyr Ser

5

1425

1430

1435

1440

Thr Pro Arg Val Leu Asn Ser Cys Ser Asn Arg Arg Val Tyr Lys Lys

1445

1450

1455

10

Met Pro Ser Ile Glu Ser Asp Val

1460

-63-

CLAIMS:

1. A stably co-transfected eukaryotic cell line capable of expressing an NMDA receptor, which receptor comprises at least one R1 subunit isoform, or at least one R1 subunit isoform and one or two R2 subunits.

2. A cell line as claimed in Claim 1, wherein the receptor expressed is a human NMDA receptor.

3. A stably co-transfected eukaryotic cell line capable of expressing a human NMDA receptor comprising the R1a subunit isoform.

4. A stably co-transfected eukaryotic cell line capable of expressing a human NMDA receptor comprising the R1a+R2A subunit combination.

5. A process for the preparation of a eukaryotic cell line capable of expressing a human NMDA receptor, which comprises stably co-transfecting a mouse Ltk⁻ host cell with one or more expression vectors, at least one such vector harbouring the cDNA sequence encoding for a human NMDA R1 receptor subunit isoform, and optionally other such vectors harbouring the cDNA sequences encoding for one or two different human NMDA R2 receptor subunits.

6. An isolated DNA molecule encoding the R1a subunit isoform of the human NMDA R1 receptor subtype comprising the sequence depicted in Figure 2B herein, SEQ ID NO. 1 + 3 + 5.

7. An isolated DNA molecule encoding the R1d subunit isoform of the human NMDA R1 receptor subtype comprising the sequence depicted in Figure 2B herein, SEQ ID NO. 1 + 3 + 7.

-64-

8. An isolated DNA molecule encoding the R1e subunit isoform of the human NMDA R1 receptor subtype comprising the sequence depicted in Figure 2A herein, SEQ ID NO. 1 + 7.

5

9. An isolated DNA molecule encoding the R2A subunit of the human NMDA receptor comprising the sequence depicted in Figure 3 herein, SEQ ID NO. 9.

10

10. A recombinant expression vector comprising the nucleotide sequence of a human NMDA receptor subunit as claimed in any one of Claims 6 to 9 together with additional sequences capable of directing the synthesis of the said human NMDA receptor subunit in cultures of stably co-transfected eukaryotic cells.

15

11. A protein preparation of human NMDA receptor subunit combinations derived from a culture of the stably co-transfected eukaryotic cells as claimed in Claim 2.

20

12. A membrane preparation containing subunit combinations of the human NMDA receptor derived from a culture of the stably co-transfected eukaryotic cells as claimed in Claim 2.

25

13. A preparation as claimed in Claim 11 or 12 containing a human NMDA receptor comprising the human R1a subunit isoform isolated from stably co-transfected mouse Ltk⁻ fibroblast cells.

30

14. A preparation as claimed in Claim 11 or 12 containing a human NMDA receptor comprising the R1a+R2A subunit combination isolated from stably co-transfected mouse Ltk⁻ fibroblast cells.

35

-65-

15. The use of the cell line as claimed in any one of Claims 2 to 4, and protein and membrane preparations derived therefrom, in screening for and designing medicaments which act upon the human NMDA receptor.

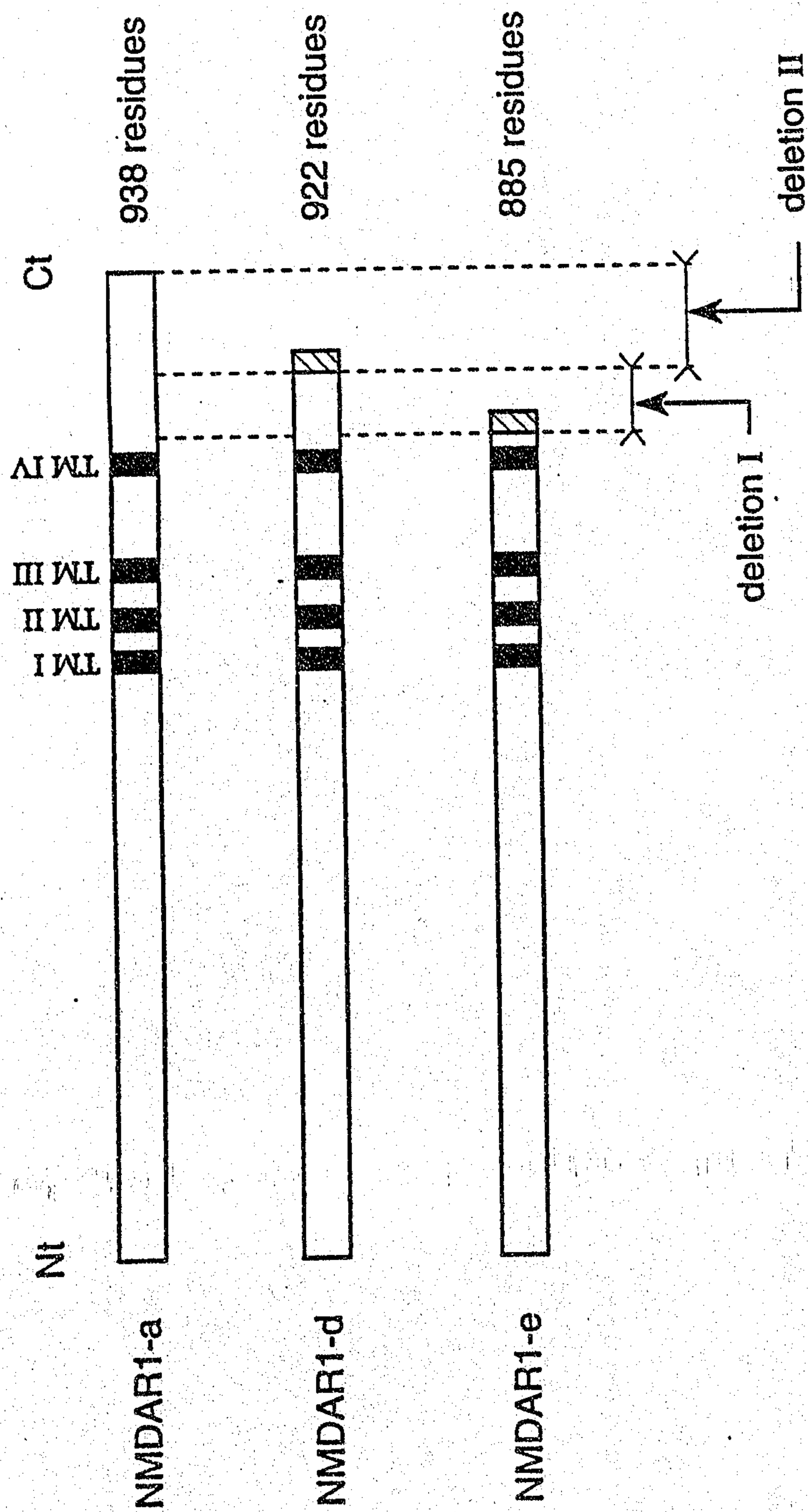


Figure 1

2/13

Figure 2A

gccgcgcagagccaggccccgcggccccgagccc																		-31
SP																		18
Met	Ser	Thr	Met	Arg	Leu	Leu	Thr	Leu	Ala	Leu	Leu	Phe	Ser	Cys	Ser	Leu	Ala	54
ATG	AGC	ACC	ATG	CGC	CTG	TTG	ACG	CTC	GCC	CTG	CTG	TTC	TCC	TGC	TCC	CTC	GCC	
																		36
Arg	Ala	Ala	Cys	Asp	Pro	Lys	Ile	Val	Asn	Ile	Gly	Ala	Val	Leu	Ser	Thr	Arg	108
CGT	GCC	GCG	TGC	GAC	CCC	AAG	ATC	GTC	AAC	ATT	GGC	GCG	GTG	CTG	AGC	ACG	CGG	
																		54
Lys	His	Glu	Gln	Met	Phe	Arg	Glu	Ala	Val	Asn	Gln	Ala	Asn	Lys	Arg	His	Gly	162
AAG	CAC	GAG	CAG	ATG	TTC	CGC	GAG	GCC	GTG	AAC	CAG	GCC	AAC	AAG	CGG	CAC	GGC	
																		72
Ser	Trp	Lys	Ile	Gln	Leu	Asn	Ala	Thr	Ser	Val	Thr	His	Lys	Pro	Asn	Ala	Ile	216
TCC	TGG	AAG	ATT	CAG	CTC	AAT	GCC	ACC	TCC	GTC	ACG	CAC	AAG	CCC	AAC	GCC	ATC	
																		90
Gln	Met	Ala	Leu	Ser	Val	Cys	Glu	Asp	Leu	Ile	Ser	Ser	Gln	Val	Tyr	Ala	Ile	270
CAG	ATG	GCT	CTG	TCG	GTG	TGC	GAG	GAC	CTC	ATC	TCC	AGC	CAG	GTG	TAC	GCC	ATC	
																		108
Leu	Val	Ser	His	Pro	Pro	Thr	Pro	Asn	Asp	His	Phe	Thr	Pro	Thr	Pro	Val	Ser	324
CTA	GTT	AGC	CAT	CCA	CCT	ACC	CCC	AAC	GAC	CAC	TTC	ACT	CCC	ACC	CCT	GTG	TCC	
																		126
Tyr	Thr	Ala	Gly	Phe	Tyr	Arg	Ile	Pro	Val	Leu	Gly	Leu	Thr	Thr	Arg	Met	Ser	378
TAC	ACA	GCC	GGC	TTC	TAC	CGC	ATA	CCC	GTG	CTG	GGG	CTG	ACC	ACC	CGC	ATG	TCC	
																		144
Ile	Tyr	Ser	Asp	Lys	Ser	Ile	His	Leu	Ser	Phe	Leu	Arg	Thr	Val	Pro	Pro	Tyr	432
ATC	TAC	TCG	GAC	AAG	AGC	ATC	CAC	CTG	AGC	TTC	CTG	CGC	ACC	GTG	CCG	CCC	TAC	
																		162
Ser	His	Gln	Ser	Ser	Val	Trp	Phe	Glu	Met	Met	Arg	Val	Tyr	Ser	Trp	Asn	His	486
TCC	CAC	CAG	TCC	AGC	GTG	TGG	TTT	GAG	ATG	ATG	CGT	GTC	TAC	AGC	TGG	AAC	CAC	
																		180
Ile	Ile	Leu	Leu	Val	Ser	Asp	Asp	His	Glu	Gly	Arg	Ala	Ala	Gln	Lys	Arg	Leu	540
ATC	ATC	CTG	CTG	GTC	AGC	GAC	GAC	CAC	GAG	GGC	CGG	GCG	GCT	CAG	AAA	CGC	CTG	
																		198
Glu	Thr	Leu	Leu	Glu	Glu	Arg	Glu	Ser	Lys	Ala	Glu	Lys	Val	Leu	Gln	Phe	Asp	594
GAG	ACG	CTG	CTG	GAG	GAG	CGT	GAG	TCC	AAG	GCA	GAG	AAG	GTG	CTG	CAG	TTT	GAC	
																		216
Pro	Gly	Thr	Lys	Asn	Val	Thr	Ala	Leu	Leu	Met	Glu	Ala	Lys	Glu	Leu	Glu	Ala	648
CCA	GGG	ACC	AAG	AAC	GTG	ACG	GCC	CTG	CTG	ATG	GAG	GCG	AAA	GAG	CTG	GAG	GCC	
																		234
Arg	Val	Ile	Ile	Leu	Ser	Ala	Ser	Glu	Asp	Asp	Ala	Ala	Thr	Val	Tyr	Arg	Ala	702
CGG	GTC	ATC	ATC	CTT	TCT	GCC	AGC	GAG	GAC	GAT	GCT	GCC	ACT	GTA	TAC	CGC	GCA	
																		252
Ala	Ala	Met	Leu	Asn	Met	Thr	Gly	Ser	Gly	Tyr	Val	Trp	Leu	Val	Gly	Glu	Arg	756
GCC	GCG	ATG	CTG	AAC	ATG	ACG	GGC	TCC	GGG	TAC	GTG	TGG	CTG	GTC	GGC	GAG	CGC	
																		270
Glu	Ile	Ser	Gly	Asn	Ala	Leu	Ala	Tyr	Ala	Pro	Asp	Gly	Ile	Leu	Gly	Leu	Gln	810
GAG	ATC	TCG	GGG	AAC	GCC	CTG	GCC	TAC	GCC	CCA	GAC	GGC	ATC	CTC	GGG	CTG	CAG	
																		288
Leu	Ile	Asn	Gly	Lys	Asn	Glu	Ser	Ala	His	Ile	Ser	Asp	Ala	Val	Gly	Val	Val	864
CTC	ATC	AAC	GGC	AAG	AAC	GAG	TCG	GCC	CAC	ATC	AGC	GAC	GCC	GTG	GGC	GTG	GTG	

3/13

Figure 2A (continued)

Ala	Gln	Ala	Val	His	Glu	Leu	Leu	Glu	Lys	Glu	Asn	Ile	Thr	Asp	Pro	Pro	Arg	306	
GCC	CAG	GCC	GTG	CAC	GAG	CTC	CTC	GAG	AAG	GAG	AAC	ATC	ACC	GAC	CCG	CCG	CGG	918	
Gly	Cys	Val	Gly	Asn	Thr	Asn	Ile	Trp	Lys	Thr	Gly	Pro	Leu	Phe	Lys	Arg	Val	324	
GGC	TGC	GTG	GGC	AAC	ACC	AAC	ATC	TGG	AAG	ACC	GGG	CCG	CTC	TTC	AAG	AGA	GTG	972	
Leu	Met	Ser	Ser	Lys	Tyr	Ala	Asp	Gly	Val	Thr	Gly	Arg	Val	Glu	Phe	Asn	Glu	342	
CTG	ATG	TCT	TCC	AAG	TAT	GCG	GAT	GGG	GTG	ACT	GGT	CGC	GTG	GAG	TTC	AAT	GAG	1026	
Asp	Gly	Asp	Arg	Lys	Phe	Ala	Asn	Tyr	Ser	Ile	Met	Asn	Leu	Gln	Asn	Arg	Lys	360	
GAT	GGG	GAC	CGG	AAG	TTC	GCC	AAC	TAC	AGC	ATC	ATG	AAC	CTG	CAG	AAC	CGC	AAG	1080	
Leu	Val	Gln	Val	Gly	Ile	Tyr	Asn	Gly	Thr	His	Val	Ile	Pro	Asn	Asp	Arg	Lys	378	
CTG	GTG	CAA	GTG	GGC	ATC	TAC	AAT	GGC	ACC	CAC	GTC	ATC	CCT	AAT	GAC	AGG	AAG	1134	
Ile	Ile	Trp	Pro	Gly	Gly	Glu	Thr	Glu	Lys	Pro	Arg	Gly	Tyr	Gln	Met	Ser	Thr	396	
ATC	ATC	TGG	CCA	GGC	GGA	GAG	ACA	GAG	AAG	CCT	CGA	GGG	TAC	CAG	ATG	TCC	ACC	1188	
Arg	Leu	Lys	Ile	Val	Thr	Ile	His	Gln	Glu	Pro	Phe	Val	Tyr	Val	Lys	Pro	Thr	414	
AGA	CTG	AAG	ATT	GTG	ACG	ATC	CAC	CAG	GAG	CCC	TTC	GTG	TAC	GTC	AAG	CCC	ACG	1242	
Leu	Ser	Asp	Gly	Thr	Cys	Lys	Glu	Glu	Phe	Thr	Val	Asn	Gly	Asp	Pro	Val	Lys	432	
CTG	AGT	GAT	GGG	ACA	TGC	AAG	GAG	GAG	TTC	ACA	GTC	AAC	GGC	GAC	CCA	GTC	AAG	1296	
Lys	Val	Ile	Cys	Thr	Gly	Pro	Asn	Asp	Thr	Ser	Pro	Gly	Ser	Pro	Arg	His	Thr	450	
AAG	GTG	ATC	TGC	ACC	GGG	CCC	AAC	GAC	ACG	TGC	CCG	GGC	AGC	CCC	CGC	CAC	ACG	1350	
Val	Pro	Gln	Cys	Cys	Tyr	Gly	Phe	Cys	Ile	Asp	Leu	Leu	Ile	Lys	Leu	Ala	Arg	468	
GTG	CCT	CAG	TGT	TGC	TAC	GGC	TTT	TGC	ATC	GAC	CTG	CTC	ATC	AAG	CTG	GCA	CGG	1404	
Thr	Met	Asn	Phe	Thr	Tyr	Glu	Val	His	Leu	Val	Ala	Asp	Gly	Lys	Phe	Gly	Thr	486	
ACC	ATG	AAC	TTC	ACC	TAC	GAG	GTG	CAC	CTG	GTG	GCA	GAT	GGC	AAG	TTC	GGC	ACA	1458	
Gln	Glu	Arg	Val	Asn	Asn	Ser	Asn	Lys	Lys	Glu	Trp	Asn	Gly	Met	Met	Gly	Glu	504	
CAG	GAG	CGG	GTG	AAC	AAC	AGC	AAC	AAG	AAG	GAG	TGG	AAT	GGG	ATG	ATG	GGC	GAG	1512	
Leu	Leu	Ser	Gly	Gln	Ala	Asp	Met	Ile	Val	Ala	Pro	Leu	Thr	Ile	Asn	Asn	Glu	522	
CTG	CTC	AGC	GGG	CAG	GCA	GAC	ATG	ATC	GTG	GCG	CCG	CTA	ACC	ATA	AAC	AAC	GAG	1566	
Arg	Ala	Gln	Tyr	Ile	Glu	Phe	Ser	Lys	Pro	Phe	Lys	Tyr	Gln	Gly	Leu	Thr	Ile	540	
CGC	GCG	CAG	TAC	ATC	GAG	TTT	TCC	AAG	CCC	TTC	AAG	TAC	CAG	GGC	CTG	ACT	ATT	1620	
Leu	Val	Lys	Lys	Glu	Ile	Pro	Arg	Ser	Thr	Leu	Asp	Ser	Phe	Met	Gln	Pro	Phe	558	
CTG	GTC	AAG	AAG	GAG	ATT	CCC	CGG	AGC	ACG	CTG	GAC	TCG	TTC	ATG	CAG	CCG	TTC	1674	

2148599

4/13

Figure 2A (continued)

Gln	Ser	Thr	Leu	Trp	Leu	Leu	Val	Gly	Leu	Ser	Val	His	Val	Val	Ala	Val	Met	576	
CAG	AGC	ACA	CTG	TGG	CTG	CTG	GTG	GGG	CTG	TCG	GTG	CAC	GTG	GTG	GCC	GTG	ATG	1728	
Leu	Tyr	Leu	Leu	Asp	Arg	Phe	Ser	Pro	Phe	Gly	Arg	Phe	Lys	Val	Asn	Ser	Glu	594	
CTG	TAC	CTG	CTG	GAC	CGC	TTC	AGC	CCC	TTC	GGC	CGG	TTC	AAG	GTG	AAC	AGC	GAG	1782	
Glu	Glu	Glu	Glu	Asp	Ala	Leu	Thr	Leu	Ser	Ser	Ala	Met	Trp	Phe	Ser	Trp	Gly	612	
GAG	GAG	GAG	GAG	GAC	GCA	CTG	ACC	CTG	TCC	TCG	GCC	ATG	TGG	TTC	TCC	TGG	GGC	1836	
Val	Leu	Leu	Asn	Ser	Gly	Ile	Gly	Glu	Gly	Ala	Pro	Arg	Ser	Phe	Ser	Ala	Arg	630	
GTC	CTG	CTC	AAC	TCC	GGC	ATC	GGG	GAA	GGC	GCC	CCC	AGA	AGC	TTC	TCA	GCG	CGC	1890	
Ile	Leu	Gly	Met	Val	Trp	Ala	Gly	Phe	Ala	Met	Ile	Ile	Val	Ala	Ser	Tyr	Thr	648	
ATC	CTG	GGC	ATG	GTG	TGG	GCC	GGC	TTT	GCC	ATG	ATC	ATC	GTG	GCC	TCC	TAC	ACT	1944	
Ala	Asn	Leu	Ala	Ala	Phe	Leu	Val	Leu	Asp	Arg	Pro	Glu	Glu	Arg	Ile	Thr	Gly	666	
GCC	AAC	TTG	GCG	GCC	TTC	CTG	GTG	CTG	GAC	CGG	CCG	GAG	GAG	CGC	ATC	ACG	GGC	1998	
Ile	Asn	Asp	Pro	Arg	Leu	Arg	Asn	Pro	Ser	Asp	Lys	Phe	Ile	Tyr	Ala	Thr	Val	684	
ATC	AAC	GAC	CCT	CGG	CTG	AGG	AAC	CCC	TCG	GAC	AAG	TTT	ATC	TAC	GCC	ACG	GTG	2052	
Lys	Gln	Ser	Ser	Val	Asp	Ile	Tyr	Phe	Arg	Arg	Gln	Val	Glu	Leu	Ser	Thr	Met	702	
AAG	CAG	AGC	TCC	GTG	GAT	ATC	TAC	TTC	CGG	CGC	CAG	GTG	GAG	CTG	AGC	ACC	ATG	2106	
Tyr	Arg	His	Met	Glu	Lys	His	Asn	Tyr	Glu	Ser	Ala	Ala	Glu	Ala	Ile	Gln	Ala	720	
TAC	CGG	CAT	ATG	GAG	AAG	CAC	AAC	TAC	GAG	AGT	GCG	GCG	GAG	GCC	ATC	CAG	GCC	2160	
Val	Arg	Asp	Asn	Lys	Leu	His	Ala	Phe	Ile	Trp	Asp	Ser	Ala	Val	Leu	Glu	Phe	738	
GTG	AGA	GAC	AAC	AAG	CTG	CAT	GCC	TTC	ATC	TGG	GAC	TCG	GCG	GTG	CTG	GAG	TTC	2214	
Glu	Ala	Ser	Gln	Lys	Cys	Asp	Leu	Val	Thr	Thr	Gly	Glu	Leu	Phe	Phe	Arg	Ser	756	
GAG	GCC	TCG	CAG	AAG	TGC	GAC	CTG	GTG	ACG	ACT	GGA	GAG	CTG	TTT	TTC	CGC	TCG	2268	
Gly	Phe	Gly	Ile	Gly	Met	Arg	Lys	Asp	Ser	Pro	Trp	Lys	Gln	Asn	Val	Ser	Leu	774	
GGC	TTC	GGC	ATA	GGC	ATG	CGC	AAA	GAC	AGC	CCC	TGG	AAG	CAG	AAC	GTC	TCC	CTG	2322	
Ser	Ile	Leu	Lys	Ser	His	Glu	Asn	Gly	Phe	Met	Glu	Asp	Leu	Asp	Lys	Thr	Trp	792	
TCC	ATC	CTC	AAG	TCC	CAC	GAG	AAT	GGC	TTC	ATG	GAA	GAC	CTG	GAC	AAG	ACG	TGG	2376	
Val	Arg	Tyr	Gln	Glu	Cys	Asp	Ser	Arg	Ser	Asn	Ala	Pro	Ala	Thr	Leu	Thr	Phe	810	
GTT	CGG	TAT	CAG	GAA	TGT	GAC	TCG	CGC	AGC	AAC	GCC	CCT	GCA	ACC	CTT	ACT	TTT	2430	
Glu	Asn	Met	Ala	Gly	Val	Phe	Met	Leu	Val	Ala	Gly	Gly	Ile	Val	Ala	Gly	Ile	828	
GAG	AAC	ATG	GCC	GGG	GTC	TTC	ATG	CTG	GTA	GCT	GGG	GGC	ATC	GTG	GCC	GGG	ATC	2484	
Phe	Leu	Ile	Phe	Ile	Glu	Ile	Ala	Tyr	Lys	Arg	His	Lys	Asp	Ala	Arg	Arg	Lys	846	
TTC	CTG	ATT	TTC	ATC	GAG	ATT	GCC	TAC	AAG	CGG	CAC	AAG	GAT	GCT	CGC	CGG	AAG	2538	
Gln	Met	Gln	Leu	Ala	Phe	Ala	Ala	Val	Asn	Val	Trp	Arg	Lys	Asn	Leu	Gln	Gln	864	
CAG	ATG	CAG	CTG	GCC	TTT	GCC	GCC	GTT	AAC	GTG	TGG	CGG	AAG	AAC	CTG	CAG	CAG	2592	
Tyr	His	Pro	Thr	Asp	Ile	Thr	Gly	Pro	Leu	Asn	Leu	Ser	Asp	Pro	Ser	Val	Ser	882	
TAC	CAT	CCC	ACT	GAT	ATC	ACG	GGC	CCG	CTC	AAC	CTC	TCA	GAT	CCC	TCG	GTC	AGC	2646	
Thr	Val	Val																885	
ACC	GTG	GTG	tga	ggccccccggagcgcgccacctgcccagtttagccccggc															

SUBSTITUTE SHEET

5/13
Figure 2B

NMDA R1a:
863
Asn Leu Gln
AAC CTG CAG
2589

864 900
Asp Arg Lys.....Lys Asp Thr
GAT AGA AAG.....AAA Asp ACG
2592 2700
deletion 1

901 938
Ser Thr Gly.....Arg Glu Ser
AGC ACC GGG.....AGG GAG AGCtga gactc
2703
deletion 2

NMDA R1d:
863
Asn Leu Gln
AAC CTG CAG
2589

864 900
Asp Arg Lys.....Lys Asp Thr
GAT AGA AAG.....AAA GAC ACG
2592 2700
deletion 1

901 922
Gln Tyr His.....Thr Val Val
CAG TAC CAT.....ACC GTG GTGtga ggcccc
2703
alternate sequence

NMDA R1e:
863
Asn Leu Gln
AAC CTG CAG
2589

864 885
Gln Tyr His.....Thr Val Val
CAG TAC CAT.....ACC GTG GTGtga ggcc
2592
alternate sequence

deletion 1: 864
Asp Arg Lys Ser Gly Arg Ala Glu Pro Asp Pro Lys Lys Lys Ala Thr
GAT AGA AAG AGT GGT AGA GCA GAG CCT GAC CCT AAA AAG AAA GCC ACA
2592

Phe Arg Ala Ile Thr Ser Thr Leu Ala Ser Ser Phe Lys Arg Arg Arg
TTT AGG GCT ATC ACC TCC ACC CTG GCT TCC AGC TTC AAG AGG CGT AGG

900
Ser Ser Lys Asp Thr
TCC TCC AAA GAC ACG
2700

deletion 2: 901
Ser Thr Gly Gly Gly Arg Gly Ala Leu Gln Asn Gln Lys Asp Thr Val
AGC ACC GGG GGT GGA CGC GGC GCT TTG CAA AAC CAA AAA GAC ACA GTG
2703

Leu Pro Arg Arg Pro Ile Glu Arg Glu Glu Gly Gln Leu Gln Leu Cys
CTG CCG CGA CGG CCT ATT GAG AGG GAG GAG GGC CAG CTG CAG CTG TGT

938
Ser Arg His Arg Glu Ser
TCC CGT CAT AGG GAG AGC tga gactccccgcccgccctctctgccccctcccc
gcagacagacagacagacggacgggacagogg

alternate sequence : 864
Gln Tyr His Pro Thr Asp Ile Thr Gly Pro Leu Asn Leu Ser Asp Pro
CAG TAC CAT CCC ACT GAT ATC ACG GGC CCG CTC AAC CTC TCA GAT CCC
2592

885
Ser Val Ser Thr Val Val
TCG GTC AGC ACC GTG GTG tga ggcccccgagcgccacctgcccagtttagccc

6/13

Figure 3

10			20			30			40			49			58		
CTGCATCCTC GACCTTCTCG GGCTACAGGG ACCGTCAGTG GCGACT												ATG	GGC	AGA	GTG	GGC	
												MET	Gly	Arg	Val	Gly	
67			76			85			94			103			112		
TAT	TGG	ACC	CTG	CTG	GTG	CTG	CCG	GCC	CTT	CTG	GTC	TGG	CGC	GGT	CCG	GCG	CCG
Tyr	Trp	Thr	Leu	Leu	Val	Leu	Pro	Ala	Leu	Leu	Val	Trp	Arg	Gly	Pro	Ala	Pro
121			130			139			148			157			166		
AGC	GCG	GCG	GCG	GAG	AAG	GGT	CCC	CCC	GCG	CTA	AAT	ATT	GCG	GTG	ATG	CTG	GGT
Ser	Ala	Ala	Ala	Glu	Lys	Gly	Pro	Pro	Ala	Leu	Asn	Ile	Ala	Val	MET	Leu	Gly
175			184			193			202			211			220		
CAC	AGC	CAC	GAC	GTG	ACA	GAG	CGC	GAA	CTT	CGA	ACA	CTG	TGG	GGC	CCC	GAG	CAG
His	Ser	His	Asp	Val	Thr	Glu	Arg	Glu	Leu	Arg	Thr	Leu	Trp	Gly	Pro	Glu	Gln
229			238			247			256			265			274		
GCG	GGC	GGG	CTG	CCC	CTG	GAC	GTG	AAC	GTG	GTA	GCT	CTG	CTG	ATG	AAC	CGC	ACC
Ala	Gly	Gly	Leu	Pro	Leu	Asp	Val	Asn	Val	Val	Ala	Leu	Leu	MET	Asn	Arg	Thr
283			292			301			310			319			328		
GAC	CCC	AAG	AGC	CTC	ATC	ACG	CAC	GTG	TGC	GAC	CTC	ATG	TCC	GGG	GCA	CGC	ATC
Asp	Pro	Lys	Ser	Leu	Ile	Thr	His	Val	Cys	Asp	Leu	MET	Ser	Gly	Ala	Arg	Ile
337			346			355			364			373			382		
CAC	GGC	CTC	GTG	TTT	GGG	GAC	GAC	ACG	GAC	CAG	GAG	GTC	GTA	GCC	CAG	ATG	CTG
His	Gly	Leu	Val	Phe	Gly	Asp	Asp	Thr	Asp	Gln	Glu	Val	Val	Ala	Gln	MET	Leu
391			400			409			418			427			436		
GAT	TTT	ATC	TCC	TCC	CAC	ACC	TTC	GTC	CCC	ATC	TTG	GGC	ATT	CAT	GGG	GGC	GCA
Asp	Phe	Ile	Ser	Ser	His	Thr	Phe	Val	Pro	Ile	Leu	Gly	Ile	His	Gly	Gly	Ala
445			454			463			472			481			490		
TCT	ATG	ATC	ATG	GCT	GAC	AAG	GAT	CCG	ACG	TCT	ACC	TTC	TTC	CAG	TTT	GGA	GCG
Ser	MET	Ile	MET	Ala	Asp	Lys	Asp	Pro	Thr	Ser	Thr	Phe	Phe	Gln	Phe	Gly	Ala
499			508			517			526			535			544		
TCC	ATC	CAG	CAG	CAA	GCC	ACG	GTC	ATG	CTG	AAG	ATC	ATG	CAG	GAT	TAT	GAC	TGG
Ser	Ile	Gln	Gln	Gln	Ala	Thr	Val	MET	Leu	Lys	Ile	MET	Gln	Asp	Tyr	Asp	Trp
553			562			571			580			589			598		
CAT	GTC	TTC	TCC	CTG	GTG	ACC	ACT	ATC	TTC	CCT	GGC	TAC	AGG	GAA	TTC	ATC	AGC
His	Val	Phe	Ser	Leu	Val	Thr	Thr	Ile	Phe	Pro	Gly	Tyr	Arg	Glu	Phe	Ile	Ser
607			616			625			634			643			652		
TTC	GTC	AAG	ACC	ACA	GTG	GAC	AAC	AGC	TTT	GTG	GGC	TGG	GAC	ATG	CAG	AAT	GTG
Phe	Val	Lys	Thr	Thr	Val	Asp	Asn	Ser	Phe	Val	Gly	Trp	Asp	MET	Gln	Asn	Val
661			670			679			688			697			706		
ATC	ACA	CTG	GAC	ACT	TCC	TTT	GAG	GAT	GCA	AAG	ACA	CAA	GTC	CAG	CTG	AAG	AAG
Ile	Thr	Leu	Asp	Thr	Ser	Phe	Glu	Asp	Ala	Lys	Thr	Gln	Val	Gln	Leu	Lys	Lys

7/13

Figure 3 (continued)

715			724			733			742			751			760		
ATC Ile	CAC His	TCT Ser	TCT Ser	GTC Val	ATC Ile	TTG Leu	CTC Leu	TAC Tyr	TGT Cys	TCC Ser	AAA Lys	GAC Asp	GAG Glu	GCT Ala	GTT Val	CTC Leu	ATT Ile
769			778			787			796			805			814		
CTG Leu	AGT Ser	GAG Glu	GCC Ala	CGC Arg	TCC Ser	CTT Leu	GGC Gly	GTC Val	ACC Thr	GGG Gly	TAT Tyr	GAT Asp	TTC Phe	TTC Phe	TGG Trp	ATT Ile	GTC Val
823			832			841			850			859			868		
CCC Pro	AGC Ser	TTG Leu	GTC Val	TCT Ser	GGG Gly	AAC Asn	ACG Thr	GAG Glu	CTC Leu	ATC Ile	CCA Pro	AAA Lys	GAG Glu	TTT Phe	CCA Pro	TCG Ser	GGA Gly
877			886			895			904			913			922		
CTC Leu	ATT Ile	TCT Ser	GTC Val	TCC Ser	TAC Tyr	GAT Asp	GAC Asp	TGG Trp	GAC Asp	TAC Tyr	AGC Ser	CTG Leu	GAG Glu	GCG Ala	AGA Arg	GTG Val	AGG Arg
931			940			949			958			967			976		
GAC Asp	GGC Gly	ATT Ile	GGC Gly	ATC Ile	CTA Leu	ACC Thr	ACC Thr	GCT Ala	GCA Ala	TCT Ser	TCT Ser	ATG MET	CTG Leu	GAG Glu	AAG Lys	TTC Phe	TCC Ser
985			994			1003			1012			1021			1030		
TAC Tyr	ATC Ile	CCC Pro	GAG Glu	GCC Ala	AAG Lys	GCC Ala	AGC Ser	TGC Cys	TAC Tyr	GGG Gly	CAG Gln	ATG MET	GAG Glu	AGG Arg	CCA Pro	GAG Glu	GTC Val
1039			1048			1057			1066			1075			1084		
CCG Pro	ATG MET	CAC His	ACC Thr	TTG Leu	CAC His	CCA Pro	TTT Phe	ATG MET	GTC Val	AAT Asn	GTT Val	ACA Thr	TGG Trp	GAT Asp	GGC Gly	AAA Lys	GAC Asp
1093			1102			1111			1120			1129			1138		
TTA Leu	TCC Ser	TTC Phe	ACT Thr	GAG Glu	GAA Glu	GGC Gly	TAC Tyr	CAG Gln	GTG Val	CAC His	CCC Pro	AGG Arg	CTG Leu	GTG Val	GTG Val	ATT Ile	GTG Val
1147			1156			1165			1174			1183			1192		
CTG Leu	AAC Asn	AAA Lys	GAC Asp	CGG Arg	GAA Glu	TGG Trp	GAA Glu	AAG Lys	GTG Val	GCC Ala	AAG Lys	TGG Trp	AAG Lys	AAC Asn	AAT Asn	ACG Thr	CTG Leu
1201			1210			1219			1228			1237			1246		
AGC Ser	CTG Leu	AGC Ser	CAC His	GCC Ala	GTG Val	TGG Trp	CCC Pro	AGG Arg	TAC Tyr	AAG Lys	TCC Ser	TTC Phe	TCC Ser	GAC Asp	TGT Cys	GAG Glu	CCG Pro
1255			1264			1273			1282			1291			1300		
GAT Asp	GAC Asp	AAC Asn	CAT His	CTC Leu	AGC Ser	ATC Ile	GTC Val	ACC Thr	CTG Leu	GAG Glu	GAG Glu	GCC Ala	CCA Pro	TTC Phe	GTC Val	ATC Ile	GTG Val
1309			1318			1327			1336			1345			1354		
GAA Glu	GAC Asp	ATA Ile	GAC Asp	CCC Pro	CTG Leu	ACC Thr	GAG Glu	ACG Thr	TGT Cys	GTG Val	AGG Arg	AAC Asn	ACC Thr	GTG Val	CCA Pro	TGT Cys	CGG Arg
1363			1372			1381			1390			1399			1408		
AAG Lys	TTC Phe	GTC Val	AAA Lys	ATC Ile	AAC Asn	AAT Asn	TCA Ser	ACC Thr	AAT Asn	GAG Glu	GGG Gly	ATG MET	AAT Asn	GTG Val	AAG Lys	AAA Lys	TGC Cys

2148599

8/13

Figure 3 (continued)

1417	1426	1435	1444	1453	1462
TGC AAG GGG TTC TGC ATT GAT ATT CTG AAG AAG CTT TCC AGA ACT GTG AAG TTT					
Cys Lys Gly Phe Cys Ile Asp Ile Leu Lys Lys Leu Ser Arg Thr Val Lys Phe					
1471	1480	1489	1498	1507	1516
ACT TAC GAC CTC TAT CTG GTG ACC AAT GGG AAG CAT GGC AAG AAA GTT AAC AAT					
Thr Tyr Asp Leu Tyr Leu Val Thr Asn Gly Lys His Gly Lys Lys Val Asn Asn					
1525	1534	1543	1552	1561	1570
GTG TGG AAT GGA ATG ATC GGT GAA GTG GTC TAT CAA CGG GCA GTC ATG GCA GTT					
Val Trp Asn Gly MET Ile Gly Glu Val Val Tyr Gln Arg Ala Val MET Ala Val					
1579	1588	1597	1606	1615	1624
GGC TCG CTC ACC ATC AAT GAG GAA CGT TCT GAA GTG GTG GAC TTC TCT GTG CCC					
Gly Ser Leu Thr Ile Asn Glu Glu Arg Ser Glu Val Val Asp Phe Ser Val Pro					
1633	1642	1651	1660	1669	1678
TTT GTG GAA ACG GGA ATC AGT GTC ATG GTT TCA AGA AGT AAT GGC ACC GTC TCA					
Phe Val Glu Thr Gly Ile Ser Val MET Val Ser Arg Ser Asn Gly Thr Val Ser					
1687	1696	1705	1714	1723	1732
CCT TCT GCT TTT CTA GAA CCA TTC AGC GCC TCT GTC TGG GTG ATG ATG TTT GTG					
Pro Ser Ala Phe Leu Glu Pro Phe Ser Ala Ser Val Trp Val MET MET Phe Val					
1741	1750	1759	1768	1777	1786
ATG CTG CTC ATT GTT TCT GCC ATA GCT GTT TTT GTC TTT GAA TAC TTC AGC CCT					
MET Leu Leu Ile Val Ser Ala Ile Ala Val Phe Val Phe Glu Tyr Phe Ser Pro					
1795	1804	1813	1822	1831	1840
GTT GGA TAC AAC AGA AAC TTA GCC AAA GGG AAA GCA CCC CAT GGG CCT TCT TTT					
Val Gly Tyr Asn Arg Asn Leu Ala Lys Gly Lys Ala Pro His Gly Pro Ser Phe					
1849	1858	1867	1876	1885	1894
ACA ATT GGA AAA GCT ATA TGG CTT CTT TGG GGC CTG GTG TTC AAT AAC TCC GTG					
Thr Ile Gly Lys Ala Ile Trp Leu Leu Trp Gly Leu Val Phe Asn Asn Ser Val					
1903	1912	1921	1930	1939	1948
CCT GTC CAG AAT CCT AAA GGG ACC ACC AGC AAG ATC ATG GTA TCT GTA TGG GCC					
Pro Val Gln Asn Pro Lys Gly Thr Thr Ser Lys Ile MET Val Ser Val Trp Ala					
1957	1966	1975	1984	1993	2002
TTC TTC GCT GTC ATA TTC CTG GCT AGC TAC ACA GCC AAT CTG GCT GCC TTC ATG					
Phe Phe Ala Val Ile Phe Leu Ala Ser Tyr Thr Ala Asn Leu Ala Ala Phe MET					
2011	2020	2029	2038	2047	2056
ATC CAA GAG GAA TTT GTG GAC CAA GTG ACC GGC CTC AGT GAC AAA AAG TTT CAG					
Ile Gln Glu Glu Phe Val Asp Gln Val Thr Gly Leu Ser Asp Lys Lys Phe Gln					
2065	2074	2083	2092	2101	2110
AGA CCT CAT GAC TAT TCC CCA CCT TTT CGA TTT GGG ACA GTG CCT AAT GGA AGC					
Arg Pro His Asp Tyr Ser Pro Pro Phe Arg Phe Gly Thr Val Pro Asn Gly Ser					

2148599

9/13

Figure 3 (continued)

2119	2128	2137	2146	2155	2164
ACG GAG AGA AAC ATT CGG AAT AAC TAT CCC TAC ATG CAT CAG TAC ATG ACC AAA Thr Glu Arg Asn Ile Arg Asn Asn Tyr Pro Tyr MET His Gln Tyr MET Thr Lys					
2173	2182	2191	2200	2209	2218
TTT AAT CAG AAA GGA GTA GAG GAC GCC TTG GTC AGC CTG AAA ACG GGG AAG CTG Phe Asn Gln Lys Gly Val Glu Asp Ala Leu Val Ser Leu Lys Thr Gly Lys Leu					
2227	2236	2245	2254	2263	2272
GAC GCT TTC ATC TAC GAT GCC GCA GTC TTG AAT TAC AAG GCT GGG AGG GAT GAA Asp Ala Phe Ile Tyr Asp Ala Ala Val Leu Asn Tyr Lys Ala Gly Arg Asp Glu					
2281	2290	2299	2308	2317	2326
GGC TGC AAA CTG GTG ACC ATC GGG AGT GGG TAC ATC TTT GCC ACC ACC AGT TAT Gly Cys Lys Leu Val Thr Ile Gly Ser Gly Tyr Ile Phe Ala Thr Thr Ser Tyr					
2335	2344	2353	2362	2371	2380
GGA ATT GCC CTT CAG AAA GGC TCT CCT TGG AAG AGG CAG ATC GAC CTG GCC TTG Gly Ile Ala Leu Gln Lys Gly Ser Pro Trp Lys Arg Gln Ile Asp Leu Ala Leu					
2389	2398	2407	2416	2425	2434
CTT CAG TTT GTG GGT GAT GGT GAG ATG GAG GAG CTG GAG ACC CTG TGG CTC ACT Leu Gln Phe Val Gly Asp Gly Glu MET Glu Glu Leu Glu Thr Leu Trp Leu Thr					
2443	2452	2461	2470	2479	2488
GGG ATC GTC CAC AAC GAG AAG AAC GAG GTG ATG AGC AAC CAG CTG GAC ATT GAC Gly Ile Val His Asn Glu Lys Asn Glu Val MET Ser Asn Gln Leu Asp Ile Asp					
2497	2506	2515	2524	2533	2542
AAC ATG GCG GGC GTA TTC TAC ATG CTG GCT GCC GCC ATG GCC CTT AGC CTC ATC Asn MET Ala Gly Val Phe Tyr MET Leu Ala Ala Ala MET Ala Leu Ser Leu Ile					
2551	2560	2569	2578	2587	2596
ACC TTC ATC TGG GAG CAC CTC TTC TAC TGG AAG CTG CGC TTC TGT TTC ACG GGC Thr Phe Ile Trp Glu His Leu Phe Tyr Trp Lys Leu Arg Phe Cys Phe Thr Gly					
2605	2614	2623	2632	2641	2650
GTG TGC TCC GAC CGG CCT GGG TTG CTC TTC TCC ATC AGC AGG GGC ATC TAC AGC Val Cys Ser Asp Arg Pro Gly Leu Leu Phe Ser Ile Ser Arg Gly Ile Tyr Ser					
2659	2668	2677	2686	2695	2704
TGC ATT CAT GGA GTG CAC ATT GAA GAA AAG AAG AAG TCT CCA GAC TTC AAT CTG Cys Ile His Gly Val His Ile Glu Glu Lys Lys Lys Ser Pro Asp Phe Asn Leu					
2713	2722	2731	2740	2749	2758
ACG GGA TCC CAG AGC AAC ATG TTA AAA CTC CTC CGG TCA GCC AAA AAC ATT TCC Thr Gly Ser Gln Ser Asn MET Leu Lys Leu Leu Arg Ser Ala Lys Asn Ile Ser					
2767	2776	2785	2794	2803	2812
AGC ATG TCC AAC ATG AAC TCC TCA AGA ATG GAC TCA CCC AAA AGA GCT GCT GAC Ser MET Ser Asn MET Asn Ser Ser Arg MET Asp Ser Pro Lys Arg Ala Ala Asp					

2148599

10/13

Figure 3 (continued)

2821	2830	2839	2848	2857	2866
TTC ATC CAA AGA GGT TCC CTC ATC ATG GAC ATG GTT TCA GAT AAG GGG AAT TTG					
Phe Ile Gln Arg Gly Ser Leu Ile MET Asp MET Val Ser Asp Lys Gly Asn Leu					
2875	2884	2893	2902	2911	2920
ATG TAC TCA GAC AAC AGG TCC TTT CAG GGG AAA GAG AGC ATT TTT GGA GAC AAC					
MET Tyr Ser Asp Asn Arg Ser Phe Gln Gly Lys Glu Ser Ile Phe Gly Asp Asn					
2929	2938	2947	2956	2965	2974
ATG AAC GAA CTC CAA ACA TTT GTG GCC AAC CGG CAG AAG GAT AAC CTC AAT AAC					
MET Asn Glu Leu Gln Thr Phe Val Ala Asn Arg Gln Lys Asp Asn Leu Asn Asn					
2983	2992	3001	3010	3019	3028
TAT GTA TTC CAG GGA CAA CAT CCT CTT ACT CTC AAT GAG TCC AAC CCT AAC ACG					
Tyr Val Phe Gln Gly Gln His Pro Leu Thr Leu Asn Glu Ser Asn Pro Asn Thr					
3037	3046	3055	3064	3073	3082
GTG GAG GTG GCC GTG AGC ACA GAA TCC AAA GCG AAC TCT AGA CCC CGG CAG CTG					
Val Glu Val Ala Val Ser Thr Glu Ser Lys Ala Asn Ser Arg Pro Arg Gln Leu					
3091	3100	3109	3118	3127	3136
TGG AAG AAA TCC GTG GAT TCC ATA CGC CAG GAT TCA CTA TCC CAG AAT CCA GTC					
Trp Lys Lys Ser Val Asp Ser Ile Arg Gln Asp Ser Leu Ser Gln Asn Pro Val					
3145	3154	3163	3172	3181	3190
TCC CAG AGG GAT GAG GCA ACA GCA GAG AAT AGG ACC CAC TCC CTA AAG AGC CCT					
Ser Gln Arg Asp Glu Ala Thr Ala Glu Asn Arg Thr His Ser Leu Lys Ser Pro					
3199	3208	3217	3226	3235	3244
AGG TAT CTT CCA GAA GAG ATG GCC CAC TCT GAC ATT TCA GAA ACG TCA AAT CGG					
Arg Tyr Leu Pro Glu Glu MET Ala His Ser Asp Ile Ser Glu Thr Ser Asn Arg					
3253	3262	3271	3280	3289	3298
GCC ACG TGC CAC AGG GAA CCT GAC AAC AGT AAG AAC CAC AAA ACC AAG GAC AAC					
Ala Thr Cys His Arg Glu Pro Asp Asn Ser Lys Asn His Lys Thr Lys Asp Asn					
3307	3316	3325	3334	3343	3352
TTT AAA AGG TCA GTG GCC TCC AAA TAC CCC AAG GAC TGT AGT GAG GTC GAG CGC					
Phe Lys Arg Ser Val Ala Ser Lys Tyr Pro Lys Asp Cys Ser Glu Val Glu Arg					
3361	3370	3379	3388	3397	3406
ACC TAC CTG AAA ACC AAA TCA AGC TCC CCT AGA GAC AAG ATC TAC ACT ATA GAT					
Thr Tyr Leu Lys Thr Lys Ser Ser Ser Pro Arg Asp Lys Ile Tyr Thr Ile Asp					
3415	3424	3433	3442	3451	3460
GGT GAG AAG GAG CCT GGT TTC CAC TTA GAT CCA CCC CAG TTT GTT GAA AAT GTG					
Gly Glu Lys Glu Pro Gly Phe His Leu Asp Pro Pro Gln Phe Val Glu Asn Val					
3469	3478	3487	3496	3505	3514
ACC CTG CCC GAG AAC GTG GAC TTC CCG GAC CCC TAC CAG GAT CCC AGT GAA AAC					
Thr Leu Pro Glu Asn Val Asp Phe Pro Asp Pro Tyr Gln Asp Pro Ser Glu Asn					

SUBSTITUTE SHEET

2148599

11/13

Figure 3 (continued)

3523	3532	3541	3550	3559	3568
TTC CGC AAG GGG GAC TCC ACG CTG CCA ATG AAC CGG AAC CCC TTG CAT AAT GAA Phe Arg Lys Gly Asp Ser Thr Leu Pro MET Asn Arg Asn Pro Leu His Asn Glu					
3577	3586	3595	3604	3613	3622
GAG GGG CTT TCC AAC AAC GAC CAG TAT AAA CTC TAC TCC AAG CAC TTC ACC TTG Glu Gly Leu Ser Asn Asn Asp Gln Tyr Lys Leu Tyr Ser Lys His Phe Thr Leu					
3631	3640	3649	3658	3667	3676
AAA GAC AAG GGT TCC CCG CAC AGT GAG ACC AGC GAG CGA TAC CGG CAG AAC TCC Lys Asp Lys Gly Ser Pro His Ser Glu Thr Ser Glu Arg Tyr Arg Gln Asn Ser					
3685	3694	3703	3712	3721	3730
ACG CAC TGC AGA AGC TGC CTT TCC AAC ATG CCC ACC TAT TCA GGC CAC TTC ACC Thr His Cys Arg Ser Cys Leu Ser Asn MET Pro Thr Tyr Ser Gly His Phe Thr					
3739	3748	3757	3766	3775	3784
ATG AGG TCC CCC TTC AAG TGC GAT GCC TGC CTG CGG ATG GGG AAC CTC TAT GAC MET Arg Ser Pro Phe Lys Cys Asp Ala Cys Leu Arg MET Gly Asn Leu Tyr Asp					
3793	3802	3811	3820	3829	3838
ATC GAT GAA GAC CAG ATG CTT CAG GAG ACA GGT AAC CCA GCC ACC GGG GAG GAG Ile Asp Glu Asp Gln MET Leu Gln Glu Thr Gly Asn Pro Ala Thr Gly Glu Glu					
3847	3856	3865	3874	3883	3892
GTC TAC CAG CAG GAC TGG GCA CAG AAC AAT GCC CTT CAA TTA CAA AAG AAC AAG Val Tyr Gln Gln Asp Trp Ala Gln Asn Asn Ala Leu Gln Leu Gln Lys Asn Lys					
3901	3910	3919	3928	3937	3946
CTA AGG ATT AGC CGT CAG CAT TCC TAC GAT AAC ATT GTC GAC AAA CCT AGG GAG Leu Arg Ile Ser Arg Gln His Ser Tyr Asp Asn Ile Val Asp Lys Pro Arg Glu					
3955	3964	3973	3982	3991	4000
CTA GAC CTT AGC AGG CCC TCC CGG AGC ATA AGC CTC AAG GAC AGG GAA CGG CTT Leu Asp Leu Ser Arg Pro Ser Arg Ser Ile Ser Leu Lys Asp Arg Glu Arg Leu					
4009	4018	4027	4036	4045	4054
CTG GAG GGA AAT TTT TAC GGC AGC CTG TTT AGT GTC CCC TCA AGC AAA CTC TCG Leu Glu Gly Asn Phe Tyr Gly Ser Leu Phe Ser Val Pro Ser Ser Lys Leu Ser					
4063	4072	4081	4090	4099	4108
GGG AAA AAA AGC TCC CTT TTC CCC CAA GGT CTG GAG GAC AGC AAG AGG AGC AAG Gly Lys Lys Ser Ser Leu Phe Pro Gln Gly Leu Glu Asp Ser Lys Arg Ser Lys					
4117	4126	4135	4144	4153	4162
TCT CTC TTG CCA GAC CAC ACC TCC GAT AAC CCT TTC CTC CAC TCC CAC AGG GAT Ser Leu Leu Pro Asp His Thr Ser Asp Asn Pro Phe Leu His Ser His Arg Asp					
4171	4180	4189	4198	4207	4216
GAC CAA CGC TTG GGT ATT GGG AGA TGC CCC TCG GAC CCT TAC AAA CAC TCG TTG Asp Gln Arg Leu Gly Ile Gly Arg Cys Pro Ser Asp Pro Tyr Lys His Ser Leu					

2148599

12/13

Figure 3 (continued)

4225	4234	4243	4252	4261	4270	
CCA TCC CAG GCG GTG AAT GAC AGC TAT CTT CGG TCG TCC TTG AGG TCA ACG GCA Pro Ser Gln Ala Val Asn Asp Ser Tyr Leu Arg Ser Ser Leu Arg Ser Thr Ala						
4279	4288	4297	4306	4315	4324	
TCG TAC TGT TCC AGG GAC AGT CGG GGC CAC AAT GAT GTG TAT ATT TCG GAG CAT Ser Tyr Cys Ser Arg Asp Ser Arg Gly His Asn Asp Val Tyr Ile Ser Glu His						
4333	4342	4351	4360	4369	4378	
GTT ATG CCT TAT GCT GCA AAT AAG AAT AAT ATG TAC TCT ACC CCC AGG GTT TTA Val MET Pro Tyr Ala Ala Asn Lys Asn Asn MET Tyr Ser Thr Pro Arg Val Leu						
4387	4396	4405	4414	4423	4432	
AAT TCC TGC AGC AAT AGA CGC GTG TAC AAG AAA ATG CCT AGT ATC GAA TCT GAT Asn Ser Cys Ser Asn Arg Arg Val Tyr Lys Lys MET Pro Ser Ile Glu Ser Asp						
4441	4451	4461	4471	4481	4491	4501
GTT TAA ^{>} AAATCTTCCA TTAATGTTTT ATCTATAGGG AAATATACGT AATGGCCAAT GTTCTGGAGG Val .						
4511	4521	4531	4541	4551	4561	4571
GTAAATGTTG AATGTCCAAT AGTGCCCTGC TAAGAGGAAG AAGATGTAGG GAGGTATTTT GTTGTGTTG						
4581	4591	4601	4611	4621	4631	4641
TTGTTGGCTC TTTTGCACAC GGCTTCATGC CATAATCTTC CACTCAAGGA ATCTTGTGAG GTGTGTGCTG						
4651	4661	4671	4681	4691	4701	4711
AGCATGGCAG ACACCAGATA GGTGAGTCCT TAACCAAAAA TAACTAATA CATAAGGGCA AGTCTCCGGG						
4721	4731	4741	4751	4761	4771	4781
ACATGCCTAC TGGGTATGTT GCCAATAATG ATGCATTGGA TGCCAATGGT GATGTTATGA TTTCCTATAT						
4791	4801	4811	4821	4831	4841	4851
TCCAAATTCC ATTAAGGTCA GCCCACCATG TAATTTTCTC ATCAGAAATG CCTAATGGTT TCTCTAATAC						

AGAATAA

2148599

13/13

