(51) International Patent Classification 6:
C12N 15/12, C07K 14/705, 16/28, C12Q 1/68, G01N 33/68

(21) International Application Number: PCT/EP95/01968
(22) International Filing Date: 24 May 1995 (24.05.95)
(30) Priority Data:
9410664.8 27 May 1994 (27.05.94) GB
9502480.8 9 February 1995 (09.02.95) GB
(71) Applicant (for all designated States except US): GLAXO GROUP LIMITED [GB/GB]; Glaxo House, Berkeley Avenue, Greenford, Middlesex UB6 0NN (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): VALERA, Soledad [CH/CH]; Glaxo Institute for Molecular Biology, 14, chemin des Aulx, CH-1228 Plan-les-Ouates (CH). BUELL, Gary, Nutter [US/CH]; Glaxo Institute for Molecular Biology, 14, chemin des Aulx, CH-1228 Plan-les-Ouates (CH).

(74) Agents: DAWSON, Hugh, B. et al.; Glaxo Wellcome plc, Glaxo House, Berkeley Avenue, Greenford, Middlesex UB6 0NN (GB).

(51) International Patent Classification 6:
C12N 15/12, C07K 14/705, 16/28, C12Q 1/68, G01N 33/68

(21) International Application Number: PCT/EP95/01968
(22) International Filing Date: 24 May 1995 (24.05.95)
(30) Priority Data:
9410664.8 27 May 1994 (27.05.94) GB
9502480.8 9 February 1995 (09.02.95) GB
(71) Applicant (for all designated States except US): GLAXO GROUP LIMITED [GB/GB]; Glaxo House, Berkeley Avenue, Greenford, Middlesex UB6 0NN (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): VALERA, Soledad [CH/CH]; Glaxo Institute for Molecular Biology, 14, chemin des Aulx, CH-1228 Plan-les-Ouates (CH). BUELL, Gary, Nutter [US/CH]; Glaxo Institute for Molecular Biology, 14, chemin des Aulx, CH-1228 Plan-les-Ouates (CH).

(74) Agents: DAWSON, Hugh, B. et al.; Glaxo Wellcome plc, Glaxo House, Berkeley Avenue, Greenford, Middlesex UB6 0NN (GB).

(54) Title: P2X RECEPTORS (PURINOCEPTOR FAMILY)

(57) Abstract

The P2X receptor of ATP has been cloned and expressed by recombinant DNA technology, so the receptor can be prepared free from other ATP receptors. The P2X receptor enables antibodies to be prepared and is useful in screening compounds for use in a variety of diseases and conditions, including epilepsy, cognition, emesis, pain (especially migraine), asthma, peripheral vascular disease, hypertension, diseases of the immune system, irritable bowel syndrome and premature ejaculation.
FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>Austria</td>
<td>GB</td>
<td>United Kingdom</td>
<td>MR</td>
<td>Mauritania</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AU</td>
<td>Australia</td>
<td>GE</td>
<td>Georgia</td>
<td>MW</td>
<td>Malawi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>Barbados</td>
<td>GN</td>
<td>Guinea</td>
<td>NE</td>
<td>Niger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BE</td>
<td>Belgium</td>
<td>GR</td>
<td>Greece</td>
<td>NL</td>
<td>Netherlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BF</td>
<td>Burkina Faso</td>
<td>HU</td>
<td>Hungary</td>
<td>NO</td>
<td>Norway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BG</td>
<td>Bulgaria</td>
<td>IE</td>
<td>Ireland</td>
<td>NZ</td>
<td>New Zealand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BJ</td>
<td>Benin</td>
<td>IT</td>
<td>Italy</td>
<td>PL</td>
<td>Poland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>Brazil</td>
<td>JP</td>
<td>Japan</td>
<td>PT</td>
<td>Portugal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BY</td>
<td>Belarus</td>
<td>KE</td>
<td>Kenya</td>
<td>RO</td>
<td>Romania</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>Canada</td>
<td>KG</td>
<td>Kyrgyzstan</td>
<td>RU</td>
<td>Russian Federation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>Central African Republic</td>
<td>KP</td>
<td>Democratic People’s Republic of Korea</td>
<td>SD</td>
<td>Sudan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>Congo</td>
<td>KR</td>
<td>Republic of Korea</td>
<td>SE</td>
<td>Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>Switzerland</td>
<td>KZ</td>
<td>Kazakhstan</td>
<td>SI</td>
<td>Slovenia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td>Côte d’Ivoire</td>
<td>LIE</td>
<td>Liechtenstein</td>
<td>SK</td>
<td>Slovakia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM</td>
<td>Cameroon</td>
<td>LK</td>
<td>Sri Lanka</td>
<td>SN</td>
<td>Senegal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>China</td>
<td>LU</td>
<td>Luxembourg</td>
<td>TD</td>
<td>Chad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>Czechoslovakia</td>
<td>LV</td>
<td>Latvia</td>
<td>TG</td>
<td>Togo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ</td>
<td>Czech Republic</td>
<td>MC</td>
<td>Monaco</td>
<td>TJ</td>
<td>Tajikistan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>Germany</td>
<td>MD</td>
<td>Republic of Moldova</td>
<td>TT</td>
<td>Trinidad and Tobago</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DK</td>
<td>Denmark</td>
<td>MG</td>
<td>Madagascar</td>
<td>UA</td>
<td>Ukraine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>Spain</td>
<td>ML</td>
<td>Mali</td>
<td>US</td>
<td>United States of America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FI</td>
<td>Finland</td>
<td>MN</td>
<td>Mongolia</td>
<td>UZ</td>
<td>Uzbekistan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>France</td>
<td>MR</td>
<td>Mauritania</td>
<td>VN</td>
<td>Viet Nam</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>Gabon</td>
<td>GE</td>
<td>Georgia</td>
<td>MW</td>
<td>Malawi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB</td>
<td>United Kingdom</td>
<td>GR</td>
<td>Greece</td>
<td>NL</td>
<td>Netherlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>Georgia</td>
<td>HU</td>
<td>Hungary</td>
<td>NO</td>
<td>Norway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GN</td>
<td>Guinea</td>
<td>IE</td>
<td>Ireland</td>
<td>NZ</td>
<td>New Zealand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td>Greece</td>
<td>IT</td>
<td>Italy</td>
<td>PL</td>
<td>Poland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HK</td>
<td>Hong Kong</td>
<td>JP</td>
<td>Japan</td>
<td>PT</td>
<td>Portugal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HU</td>
<td>Hungary</td>
<td>KE</td>
<td>Kenya</td>
<td>RO</td>
<td>Romania</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IE</td>
<td>Ireland</td>
<td>KG</td>
<td>Kyrgyzstan</td>
<td>RU</td>
<td>Russian Federation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL</td>
<td>Israel</td>
<td>KP</td>
<td>Democratic People’s Republic of Korea</td>
<td>SD</td>
<td>Sudan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IN</td>
<td>India</td>
<td>KR</td>
<td>Republic of Korea</td>
<td>SE</td>
<td>Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IS</td>
<td>Iceland</td>
<td>KZ</td>
<td>Kazakhstan</td>
<td>SI</td>
<td>Slovenia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISL</td>
<td>Iceland</td>
<td>LI</td>
<td>Liechtenstein</td>
<td>SK</td>
<td>Slovakia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT</td>
<td>Italy</td>
<td>LK</td>
<td>Sri Lanka</td>
<td>SN</td>
<td>Senegal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>Japan</td>
<td>LU</td>
<td>Luxembourg</td>
<td>TD</td>
<td>Chad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KG</td>
<td>Kyrgyzstan</td>
<td>LV</td>
<td>Latvia</td>
<td>TG</td>
<td>Togo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KZ</td>
<td>Kazakhstan</td>
<td>MC</td>
<td>Monaco</td>
<td>TJ</td>
<td>Tajikistan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIE</td>
<td>Liechtenstein</td>
<td>MD</td>
<td>Republic of Moldova</td>
<td>TT</td>
<td>Trinidad and Tobago</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR</td>
<td>Liberia</td>
<td>MG</td>
<td>Madagascar</td>
<td>UA</td>
<td>Ukraine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS</td>
<td>Lesotho</td>
<td>ML</td>
<td>Mali</td>
<td>US</td>
<td>United States of America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LU</td>
<td>Luxembourg</td>
<td>MN</td>
<td>Mongolia</td>
<td>UZ</td>
<td>Uzbekistan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LV</td>
<td>Latvia</td>
<td>MR</td>
<td>Mauritania</td>
<td>VN</td>
<td>Viet Nam</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
P2X RECEPTORS (PURINOCEPTOR FAMILY)

This invention relates to the P_{2X}-purinoceptor, its preparation and uses.

The P_{2X}-purinoceptor is a ligand-gated ion channel; that is, the receptor itself forms an ion channel which opens when extracellular adenosine 5'-triphosphate (ATP) binds to the receptor. There are five other classes of neurotransmitter receptors (nicotinic acetylcholine, glutamate, glycine, GABA\textsubscript{A} and 5-HT\textsubscript{3}); these form a structurally related superfamly of ligand-gated ion channels (Barnard, Trends Biochem. Sci. 17, 368-374, (1992)). The P_{2X}-receptor now identifies a new family of this type of receptor. The unique structure of this receptor, the widespread distribution of this receptor throughout the body, and the numerous physiological roles this receptor may play, make it an important protein that can be used to identify new, therapeutically effective, compounds for the treatment of a number of pathological states.

In 1929 the eminent physiologist Szent-Gyorgyi described powerful cardiovascular actions of extracellular purine nucleosides (e.g. adenosine) and nucleotides (e.g. ATP) (Drury & Szent-Gyorgyi, J. Physiol. 68 213-237 (1929)), but it was not until 1972 that pharmacological evidence was provided to suggest the existence of distinct receptors for extracellular ATP (ie. that recognise ATP but not adenosine) (Burnstock, Pharmacological Reviews 21 509-581 (1972)). The seminal and subsequent work on this area by Burnstock and colleagues was largely unaccepted throughout the 1970s and early 1980s until the development of a range of relatively selective ligands
and techniques for directly measuring ATP release overwhelmingly substantiated Burnstock's hypothesis (Barnard et al., Trends Pharmacol. Sci. 15 67-70 (1994)). In the past four or five years, unequivocal evidence for the role of ATP as a neurotransmitter has been provided for sympathetic control of blood flow to the intestine and smooth muscle tone (contractility) in genitourinary tissue such as vas deferens, bladder and ureter (Barnard et al. (loc. cit.) and Evans & Surpremant, Brit. J. Pharmacol. 106 242-249 (1992)). Substantial indirect evidence also exists for the role of ATP as a neurotransmitter in a number of distinct neurones in the spinal cord, autonomic ganglia and certain nuclei in the central nervous system (Bean, Trends Pharmacol. Sci. 15 67-70 (1992), Evans et al., Nature 357, 503-505 (1992) and Edwards et al., Nature 359 144-147 (1992)).

Purinoceptors are classified as P₁ (adenosine as ligand) and P₂ (ATP as ligand). The P₂ receptors are subclassified into two broad types - those that are 7-transmembrane receptors that couple to G-proteins (P₂y, P₂u, P₂τ, and perhaps P₂z) and those that form a directly gated ion channel (P₂x). Pharmacological and/or physiological evidence for subtypes of each of these types of receptors exists. The most recent nomenclature for these receptors is shown below.

<table>
<thead>
<tr>
<th>Type</th>
<th>P₂x</th>
<th>P₂y</th>
<th>P₂z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtype</td>
<td>P₂x₁, P₂x₂, P₂x₃</td>
<td>P₂y₁, P₂y₂, P₂y₃</td>
<td>P₂z₁, P₂z₂, P₂z₃</td>
</tr>
</tbody>
</table>

Various P₂ receptors have previously been cloned. P₂y₁ was cloned by the Barnard/Burnstock group (Webb et al., FEBS Lett. 324 219-225 (1993)) based on homology with
other 7-TM G-protein coupled receptors. This group used PCR technology and primers based on conserved domains of the second and sixth transmembrane regions to screen a mammalian brain cDNA library and, with final success, an embryonic chick whole-brain cDNA library.

P_{2Y2}/P_{2U} was cloned by the Julius laboratory (Lustig et al., Proc. Nat’l. Acad. Sci. USA 90 5113-5117 (1993)) by expression cloning in the oocyte from cDNA obtained from a NG108-15 neuroblastoma cell line.

P_{2Y1}/P_{2T} was also obtained by the Barnard/Burnstock group using the same probes and embryonic brain cDNA library used to obtain the P_{2Y1} receptor (Barnard et al., Trends Pharmacol. Sci. 15 67-70 (1994)).

However, as yet, cloning of the P_{2X} receptor has remained an elusive goal. The prior cloning exercises undertaken for the other P_{2} receptors do not provide an adequate lead to enable the P_{2X} receptor to be cloned. First, all the above purinoceptors are G-protein coupled 7-TM proteins. Their myriad functions (like those of all 7-TM receptors) occur through G-protein activation of one or more second messenger systems. There are over 200 currently identified proteins which belong to this 7-TM/G-protein coupled family. Agonists at these receptors activate cascades of intracellular transduction pathways, often involving several enzymes; the response of the cell is inherently slow (several seconds to minutes) and changes in excitability are subtle if they occur. In contrast, the P_{2X} receptor is a fundamentally different type of purinoceptor that incorporates an ion channel. Activation of P_{2X} receptors is rapid (milliseconds), has predominately local effects, and brings about immediate
depolarisation and excitation.

Secondly, the tissue distribution of the P2X receptor is distinctly different from other purinoceptors, and the physiological roles differ from other purinoceptors.

One of the principal established ways to clone a receptor is based on sequence relatedness of the nucleotides that encode the amino acids of the receptor protein; it depends on there being a fairly high level of homology between a known sequence and that of the unknown receptor. This method was used to clone the P2Y1 form (above). Several laboratories, including that of the applicants, invested significant effort in obtaining the P2X receptor using PCR techniques and primers based on conserved regions of various ligand-gated ion channels (ie. nicotinic ACh, GABA, glutamate, 5-HT3). This approach failed. With hindsight, this failure can be rationalised, as it can now, but only now, be seen that the structure of the P2X receptor bears no homology with any of these ligand-gated ion channels. For the same reason, approaches based on fragment hybridisation would not succeed.

However, by adopting a different approach, it has now been found possible to clone the P2X receptor, and it is on this achievement that the present invention is in part based.

According to a principal aspect of the present invention, there is provided a recombinant or isolated DNA molecule encoding a P2X receptor, wherein the receptor:

(a) has the amino sequence shown in Figure 1, Figure 2, Figure 3 or Figure 4; or
is substantially homologous to the sequence shown in Figure 1, Figure 2, Figure 3 or Figure 4;

or a fragment of such a DNA molecule, which fragment includes at least 15 nucleotides taken from nucleotides 1 to 813 shown in Figure 1, the full nucleotide sequences shown in Figures 2 and 3, or from nucleotides 1 to 1744 shown in Figure 4.

The sequence shown in Figure 1 is a cDNA sequence that encodes a rat vas deferens $P_{2X}$ receptor. This sequence is 1837 bases in length and encodes a protein of 399 amino acids. As was determined after the receptor was cloned, approximately one half of the protein-encoding sequence, from nucleotides 814 onwards, had been discovered previously but the function of the previously cloned sequence was not known except that it appeared to be implicated in apoptotic cell death (Owens et al., Mol. Cell. Biol. 11 4177-4188 (1991)). The Owens et al. sequence lacks a translation initiation site and could not be made into protein. (In Figure 1, the upstream portion of the reported sequence of Owens et al., namely PQLAHGCPYPHHR, which is not shared with the $P_{2X}$ receptor, is shown for comparative purposes and does not form part of the invention.)

Preferably the Figure 1 sequence fragments are taken from nucleotides 1-810. Often the Figure 4 sequence fragments are taken from nucleotides 1-777.

The sequence shown in Figure 2 is a cDNA sequence that encodes a rat superior cervical ganglion $P_{2X}$ receptor.

The sequence shown in Figure 3 is a cDNA sequence that encodes a rat dorsal root ganglion $P_{2X}$ receptor.
The sequence shown in Figure 4 is the cDNA sequence that encodes a human \( P_{2x} \) receptor. The cDNA was isolated from the human urinary bladder using a rat \( P_{2x} \) probe. It is 2643 bases long and encodes a 399 amino acid protein having an amino acid sequence which is highly homologous with the amino acid sequence of the rat \( P_{2x} \) receptor isolated from rat vas deferens and with the rat \( P_{2x} \) receptors isolated from a rat superior cervical ganglion and from a rat dorsal root ganglion. Recently we have become aware of an expressed sequence tag corresponding to residues 1745-1933 (Proc. Natl. Acad. Sci. USA 91,10645-10649 (Oct. 1994)).

Sequences which are substantially homologous to the Figure 1, Figure 2, Figure 3 or Figure 4 amino acid sequence include those which encode proteins having at least 40%, 50%, 60%, 70%, 80%, 90%, 95% or 99% homology in increasing order of preference. A protein having at least 99% homology with the amino acid sequence of Figure 1, Figure 2, Figure 3 or Figure 4 will have no more than four amino acid variations from such a sequence. Preferred substantially homologous sequences include \( P_{2x} \) sequences from other species. Thus for the rat \( P_{2x} \) receptor sequences a preferred substantially homologous sequence is a human \( P_{2x} \) sequence. One method of determining sequence homology is disclosed in WR Pearson and DJ Lipman, Proc Natl Acad Sci USA 85:2444-2448 (1988).

Fragments may of course be larger than 15 nucleotides. Fragments encoding substantially the whole of the \( P_{2x} \) rat receptors or human receptor may be expected to share the biological activity of the receptor, or at least some of its biological activities. Shorter fragments may be useful for encoding one or more selected domains of the receptor, or simply as probes for detecting or identifying other useful DNA sequences, including those encoding substantially homologous proteins. Fragments of
at least 20, 30 or 50 nucleotides may be more frequently of use than shorter ones.

DNA molecules of the invention are useful for a number of purposes. First, and not least, the P<sub>2x</sub> cDNA shown in Figure 1, in Figure 2, in Figure 3 and in Figure 4 enables the relevant proteins to be expressed in living cells. This would not be possible with fragments of the cDNA. However not only are fragments of DNA within the scope of the invention, for the various purposes mentioned above, but also genomic and other sequences of DNA (including synthetic DNA and "minigenes", which include at least one, but not all, of the introns naturally present in the gene) are included within its scope. cDNA sequences encoding the rat receptor proteins or human P<sub>2x</sub> receptor protein may be preferred in some circumstances because such sequences are smaller than either genomic or minigene DNA and therefore more amenable to cloning manipulations. The P<sub>2x</sub> receptor protein can be stably expressible in Chinese hamster ovary (CHO) cells, as will be described below.

Still on the subject of expression, while it would be possible to express genomic DNA in eukaryotic cells, it is much more difficult to manipulate the DNA for insertion into host cells due to the larger size that commonly results from introns. The size is particularly important for the expression of RNA; very long cRNAs -- the size of whole genes -- are difficult to make in sufficient quantity. On the other hand, expression from RNA is much preferred at least for the investigation of ion channel proteins, because the Xenopus oocyte is sufficiently large to be studied easily by electrophysiological methods.
Secondly, the cDNA sequences encode proteins that, in their predicted folding within the membrane, differ from other known proteins. This is advantageous because, based on historical precedent, this will lead to the discovery of a large family of related proteins and these may have functional roles unrelated to signalling mediated by ATP.

Thirdly, knowledge of the protein sequences encoded by rat and human \( P_{2X} \) cDNA allows the development of molecular models that predict the detailed disposition within the membrane. It further allows the correctness of such models to be determined by expression of mutagenised proteins. These two approaches are advantageous because they may permit the molecular design of complementary therapeutic agents that activate or block the receptor.

Fourthly, the \( P_{2X} \) cDNA sequences allow the distribution of the RNA that encodes this receptor, as well as the receptor protein itself, to be mapped in human tissues. RNA distribution can be determined by in situ hybridisation. Such hybridisation studies are disclosed in the present examples. Knowledge of a deduced amino acid sequence from cDNA allows synthetic peptides to be made that can be used to generate antibodies that selectively recognise a \( P_{2X} \) receptor. Thus a \( P_{2X} \) protein can be mapped by immunohistochemistry. This may suggest novel therapeutic applications for drugs that activate or block the \( P_{2X} \) receptor, that can not be predicted on the basis of less sensitive current methods for localising the receptor (radioactive ligand binding).

Fifthly, rat \( P_{2X} \) cDNA is advantageous because it can allow the isolation of a closely related cDNA from human tissue.
Sixthly, the isolation of the human Pₓ receptor cDNA clone will enable a human genomic clone to be obtained. It is probable that mutations of this gene will be discovered that lead to human genetic disease. The analysis of such mutations may lead to appropriate treatments of diseases or disorders caused by such mutations.

In one aspect of the present invention rat vas deferens Pₓ receptor was cloned by a method which does not require prior inference about structure. Tissues were chosen that were believed to be rich in the RNA for the receptor of interest. A number of tissue sources were tried but they did not provide RNA that led to ATP responses in oocytes. Eventually, vas deferens was chosen. From extracted polyadenylated RNA, a cDNA library or bank that corresponds as far as possible to the DNAs in the tissue was constructed. It was not assured, either before work began or until it was satisfactorily completed, that a satisfactory cDNA library in which the rat Pₓ gene was represented could be constructed; nevertheless, this was achieved in plasmid pBKCMV.

An individual clone within the library that contains the rat vas deferens Pₓ cDNA of interest was detected by progressive fractionation of the library; at each step the fraction was tested to determine whether RNA made from it can direct the formation of the protein of interest. More specifically, RNA was transcribed in vitro from the cDNAs in the library (approximately 2 million) and the RNA ("cRNA") mixture was injected into immature Xenopus oocytes. cRNA is very susceptible to inadvertent enzymatic degradation, so all procedures were carried out under sterile conditions. The cDNA pools were made by the miniprep procedure and therefore
contained large amounts of E. coli RNA; this difficulty was overcome by precipitating any RNA before the cRNA was transcribed.

Detection of the protein can in principle be done by radioactive ligand binding or by a functional response. The activation of G proteins in the Xenopus oocyte and the subsequent cellular response was used to obtain the P_{2Y2}/P_{2U} receptor. In the present work, a decision was made to use the opening of the integral ion channel of the P_{2X} as the response. Individual oocytes were screened two days after injection to determine whether they had made P_{2X} receptor protein in their membrane. This was done by recording the current flowing across the oocyte membrane when ATP (30 μM) was applied to the outside of the oocyte; if the P_{2X} receptor has been produced, a small transient current would be expected. However, testing for expression of the receptor was not straightforward, as some batches of oocytes exhibit responses to ATP because they naturally express other kinds of ATP receptor. This difficulty was overcome as follows: when an oocyte responded to ATP with the expected current this was further tested by blockade with a P_{2X} receptor antagonist (suramin). The cDNA fraction that gave led to the positive response in such an oocyte was further divided, and each fraction was again tested. Such progressive fractionation led to isolation of a single clone. The insert in the plasmid was sequenced; the sequence is shown in Figure 1. This sequence was used to design PCR primers which were used in the cloning of cDNA encoding a P_{2X} receptor from a rat superior cervical ganglion (see Figure 2). A similar procedure was then used in the cloning of cDNA encoding a P_{2X} receptor from a rat dorsal root ganglion (see Figure 3).
DNA in accordance with the invention will usually be in recombinant or isolated form and may be in the form of a vector, such as a plasmid, phagemid, cosmid or virus, and in some embodiments contains elements to direct expression of the protein, for example in a heterologous host. Non-expressible vectors are useful as cloning vectors.

Although DNA in accordance with the invention may be prepared synthetically, it is preferred that it be prepared by recombinant DNA technology. Ultimately, both techniques depend on the linkage of successive nucleotides and/or the ligation of oligo- and/or poly-nucleotides.

The invention enables, for the first time, \( P_{2X} \) receptor to be prepared by recombinant DNA technology and hence free from protein with which it is naturally associated or contaminated (such as the \( P_{2\alpha} \) or, particularly, \( P_{2Y} \) receptor, or other ATP receptors or binding proteins), and this in itself forms another aspect of the invention. The protein will generally be associated with a lipid bilayer, such as a cell, organelle or artificial membrane. \( P_{2X} \) receptor prepared by expression of DNA in accordance with the first aspect may be glycosylated, but does not have to be. Generally speaking, receptor proteins and ion channels that are glycosylated will also function after carbohydrate removal or when expressed in cells that do not glycosylate the protein. However, there are often important quantitative differences in the function between the glycosylated and non-glycosylated protein. In the case of the rat vas deferens \( P_{2X} \) receptor, we believe that the native protein is glycosylated because it has a molecular weight of 62 kd
when purified from the rat vas deferens, as compared to the molecular weight of 45 kd for the cloned protein. Similar results were obtained for the human P<sub>2x</sub> receptor (see later).

There are also several asparagine residues in the extracellular domain that are likely sites of sugar attachment.

Knowledge of the amino acid sequence of a P<sub>2x</sub> receptor enables the protein or peptide fragments of it to be prepared by chemical synthesis, if required. However, preparation by expression from DNA, or at least translation from RNA, will usually be preferred.

Particularly useful peptide fragments within the scope of the invention include epitopes (which may contain at least 5, 6, 7, 10, 15 or 20 amino acid residues) of the P<sub>2x</sub> receptor which are immunologically non-cross reactive with the RP-2 polypeptide disclosed in Owens et al., loc. cit.

A P<sub>2x</sub> receptor, and fragments of it, can be used to prepare specific polyclonal and monoclonal antibodies, which themselves form part of the invention. Polyclonal and monoclonal antibodies may be prepared by methods well established in the art. Hybridoma and other cells expressing monoclonal antibodies are also within the invention.

RNA encoding a P<sub>2x</sub> receptor, transcribable from DNA in accordance with the invention and substantially free form other RNAs, also forms part of the invention, and may be useful for a number of purposes including hybridisation
studies, in vitro translation and translation in appropriate in vivo systems such as Xenopus oocytes.

The invention also relates to host cells transformed or transfected with a vector as described above. Host cells may be prokaryotic or eukaryotic and include mammalian cells (such as COS, CHO cells and human embryonic kidney cells (HEK 293 cells)), insect cells, yeasts (such as Saccharomyces cerevisiae) and bacteria (such as Escherichia coli). Host cells may only give transient expression of the receptor, as in the case of COS cells, but for preference the host cells are stably transfected with the vector. Host cells which appropriately glycosylate the receptor are preferred. A CHO cell line or any other cell line that stably expresses a P2X receptor can be used for electrophysiological, calcium-influx, calcium-imaging and ligand-binding studies. Host cells which do not express the receptor may still be useful as cloning hosts.

A P2X receptor prepared by recombinant DNA technology in accordance with the invention has a number of uses, either in situ in a membrane of the expression host or in in vitro systems. In particular, the receptor can be used as a screen for compounds useful in a variety of human (or other animal) diseases and conditions, as will now be briefly described. Such compounds include those present in combinatorial libraries, and extracts containing unknown compounds (e.g. plant extracts).

**Epilepsy** Epilepsy results from overexcitation of distinct neurones in specific regions of the brain, in particular in the hippocampus. Functional ATP P2X receptors are known to be present in some hippocampal
neurones. If the $P_{2X}$ receptors are expressed on inhibitory interneurons, then receptor agonists would be therapeutically useful. If the receptor is expressed on principal (pyramidal or granule) cells, then receptor antagonists will be useful. If will now be possible to determine which classes of neuron express the receptor.

Cognition Hippocampal neurones respond to ATP by activation of a $P_{2X}$ receptor; these areas are of primary importance to cognition. It is now possible to determine the cellular localisation of the $P_{2X}$ receptor with in the hippocampus; depending on this localisation, either agonists or antagonists might be effective to enhance memory.

Emesis The acute trigger for emesis is rapid contraction of smooth muscle of the upper gastrointestinal tract. Activation of ATP $P_{2X}$ receptors present on smooth muscle of the GI tract, in particular the stomach and trachea, results in strong, rapid muscle contractions. $P_{2X}$-antagonists selective for visceral smooth muscle could be useful for emesis. Furthermore, $P_{2X}$ receptors are known to be expressed in the nucleus of the tractus solitarius (Ueno et al., J. Neurophysiol. 68 778-785 (1992)) and may be involved in transmission from primary visceral afferents; this could be blocked by selective $P_{2X}$ antagonists.

Pain First, $P_{2X}$ receptors are expressed in dorsal horn neurones of the spinal cord. Activation of these neurones by ATP causes fast depolarizing, excitatory responses (Jahr & Jessell, Nature 304 730-733 (1983)); if a component of the transmission from nociceptive fibres is mediated by ATP then this could be blocked by
a P2X antagonist. Secondly, ATP is one of the most noxious substance known when applied intradermally. This is because it activates directly the peripheral terminals of small diameter nociceptive fibres; it is known that the cell bodies in the dorsal root ganglion express P2X receptors. A P2X antagonist would be a peripherally active analgesic, and is likely to be effective in migraine.

Asthma Bronchial smooth muscles contract in response to activation of P2X receptors. This may occur in response to ATP released from sympathetic nerves, or from local immune cells. P2X antagonists may help to prevent stimulus-evoked spasms of bronchial smooth muscle and thereby reduce the frequency and/or severity of asthmatic attacks.

Peripheral vascular disease It is becoming clear that ATP and not noradrenaline is the primary vasoconstrictor neurotransmitter in small resistance arteries - those that comprise over 70% of total peripheral resistance. This has been shown for many vessels (Westfall et al., Ann. N.Y. Acad. Sci. 603 300-310 (1991)). A selective antagonist could be used for local collateral vasodilation.

Hypertension Hypertension that is associated with increased sympathetic tone could be treated with P2X receptor antagonists, because ATP is a major excitatory transmitter to many resistance vessels in several species including man (Westfall et al., loc. cit. and Martin et al., Br. J. Pharmacol. 102 645-650 (1991)).

Diseases of the immune system A molecule identical to part of the P2X receptor has been cloned from thymocytes that have been induced to die (Owens et al., loc. cit.).
The selective expression in these conditions implies that a molecule closely related to the P_2X receptor plays a role in the apoptosis that is an integral part of the selection of immunocompetent cells. The molecule described by Owens et al. (RP-2) was incomplete and could not have been translated into protein. The cloning of the P_2X receptor will now allow the isolation of full length RP-2 clones, their heterologous expression and the determination of their functional roles.

Irritable bowel syndrome ATP is an important transmitter to the smooth muscles of the intestinal tract, particularly in the colon. It is also a transmitter between neurons in the enteric nervous system, by activating P_2X receptors (Galligan, Gastroenterology, in press). Antagonists at P_2X receptors may therefore have utility in the management of this condition.

Premature ejaculation This could be prevented by preventing stimulus-evoked contraction of vas deferens smooth muscle. P_2X receptors are highly expressed in this tissue; antagonists at this site would prevent vas deferens contractility during sympathetic excitation.

Cystitis P_2X receptors may be implicated in increased bladder sensitivity in patients with cystitis. Thus antagonists of such P_2X receptors may be useful in treating cystitis.

Useful agonists and antagonists identified as described above also form an aspect of the invention.

The cloning of the hP_2X receptor is an important aspect of the present invention. hP_2X is the first human member of
a multigene family of ionotropic purinoceptors. Its strong similarity with P2X, isolated from rat vas deferens and with P2X isolated from rat superior cervical ganglion or from rat dorsal root ganglion, suggests that it is a human homolog of the rat proteins. The present inventors have found that differences between these two sequences are nearly all conservative substitutions of hydrophilic residues. Surprisingly, hP2X has only 41% identity with the other reported P2X receptor, that from rat PC12 cells (Brake et al, New structural motif for ligand-gated ion channels defined by an ionotropic ATP receptor Nature 371: 519-523 (1994)). The PC12 derived receptor was proposed to have a similar membrane topography and shares the conserved spacing of cysteine residues, indicated for the two smooth muscle sequences in Figure 5.

The computed molecular weight of the hP2X polypeptide (45 kd) agrees with that of the in vitro translation product when made in absence of pancreatic microsomal membranes. A larger product, 60 kd, produced in presence of microsomes suggests glycosylation and supports the idea of a central extracellular domain. The predicted hP2X protein thus has the general features of other cloned members of this family (Valera et al, A new class of ligand-gated ion channel defined by P2X receptor for extracellular ATP Nature 371: 516-519 (1994); Brake - supra): a large, cysteine-rich extracellular central domain flanked by two transmembrane spans and short internal N- and C-termini.

The distribution of the hP2X mRNA was examined by northern blot analysis. Hybridisation of a principal 2.6 kb species was seen in all RNA samples tested, with the exception of brain. A smaller, 1.8 kb band, observed in
spleen, and lung mRNAs could be due to a shorter 3' untranslated portion of the mRNA, as occurs for P2X mRNA from the rat vas deferens. The hybridisation observed in thymus, lung, spleen and liver RNA may reflect the content of smooth muscle in those organs. However, hP2X is likely to have roles in other cell types, as demonstrated by its presence in adrenal gland, and the hemopoietic cell line HL60. The strong induction of hP2X mRNA by HL60 differentiation may reflect a parallel observation in rat in which the smooth muscle form of P2X mRNA can be induced in immature thymocytes by dexamethasone (RP2 mRNA; Owens et al, Identification of mRNAs associated with programmed cell death in immature thymocytes J J Molec Cell Biol 11: 4177-4188 (1991)).

The present invention has enabled the first comprehensive pharmacological characterization of a cloned P2X-purinoceptor to be made. The time course of the responses to ATP and the sensitivity to α,β,-methylen ATP are similar to those reported for the native hP2X in urinary bladder (Inoue & Brading, Human, pig and guinea-pig bladder smooth muscle cells generate similar inward currents in response to purinoceptor activation Br J Pharmacol 103: 1840-1841 (1991)). Thus the functional properties of some native P2X purinoceptors can be obtained by the expression of a single molecular species. The agonist induced current recorded from oocytes expressing the hP2X clone gives a direct measure of the activation of P2X-purinoceptors in a system with low levels of endogenous ectonucleotidase activity. The agonist profile 2MeSATP>ATP>α,β,-meATP for hP2X is similar to that of the cloned rat vas deferens P2X-purinoceptor. The high potency of α,β,-meATP in whole tissue studies (α,β,-meATP >> 2MeSATP>ATP) probably reflects, its
resistance to ectonucleotidases.

The concentration-effect curves for ATP, 2MeSATP and 2-chloro-ATP were superimposable, indicating that these particular substitutions at the 2' position on the adenine ring do not affect agonist binding to the P_{2X}-purinoceptor. The agonist activity of AP_{5}A is likely to be because diadenosine phosphates (AP_{5}A, and AP_{6}A) released from the platelets can act as vasoactive agents through activation of P_{2X}-purinoceptors.

Preferred features of each aspect of the invention are as for each other aspect, mutatis mutandis.

The invention will now be illustrated by the following examples. The examples refer to the accompanying drawings, in which:

FIGURE 1 shows DNA and amino acid sequences of the rat vas deferens P_{2X} receptor as determined in Example 2. (SEQ ID NO 4).

FIGURE 2 shows DNA and amino acid sequences of a rat superior cervical ganglion P_{2X} receptor, as determined in Example 11. (SEQ ID NO 5).

FIGURE 3 shows DNA and amino acid sequences of a rat dorsal root ganglion P_{2X} receptor, as determined in Example 12. (SEQ ID NO 6).

FIGURE 4 shows DNA and amino acid sequences of a
human P<sub>2X</sub> receptor as determined in Example 6. (SEQ ID NO 7)

FIGURE 5 shows the alignment of the predicted amino acid sequence of hP<sub>2X</sub> with the rat vas deferens P<sub>2X</sub>, and in vitro translation of hP<sub>2X</sub> protein.

TM1 and TM2 filled boxes indicate the hydrophobic regions and boxed amino acids indicate the differences between the two sequences.

o indicates conserved cysteine residues.
* Indicates potential sites of N-glycosylation.

FIGURE 6 shows an SDS-PAGE analysis of <sup>35</sup>S-methionine labelled hP<sub>2X</sub> protein. Lanes 1 and 2 show in vitro coupled transcription/translation of pBKCMV-hP<sub>2X</sub> cDNA in the absence and presence of microsomal membranes, respectively.

FIGURES 7 AND 8 show Northern analyses of the hP<sub>2X</sub> cDNA, wherein:

A) FIGURE 7 shows Northern blot with 8 µg of total RNA from differentiated HL60 cells.

0 indicates HL60 cells without treatment; PMA2 and PMA3 indicate respectively cells treated 2 days, and 3 days with PMA; DMSO indicates cells treated 6 days with DMSO; dcAMP indicates cells treated 5 days with dibutryl
cAMP;
UB indicates 100 ng of polyA⁺ RNA from human urinary bladder; and

B) FIGURE 8 shows distribution of hP₂ₓ in human tissues. Lanes contained 1 μg polyA⁺ RNA except for the urinary bladder which contained 0.2 μg of polyA⁺ RNA.

FIGURES 9, 10 and 11 show the response of oocytes expressing hP₂ₓ to purinoceptor agonists, wherein:

A) FIGURE 9 shows traces which show inward currents evoked by ATP, 2 me SATP and α,β, me ATP (0.1, 1, and 100 μM). Records for each agonist are from separate oocytes;

B) FIGURE 10 shows concentration response relationships of full P₂ₓ-purinoceptor agonists. Data are expressed relative to the peak response to 100 μM ATP; and

C) FIGURE 11 shows concentration response of partial P₂ₓ-purinoceptor agonists. Data are fitted with a Hill slope of 1 (n = 4-8).

FIGURES 12 and 13 show the effects of P₂-purinoceptor antagonists of hP₂ₓ mediated responses, wherein;

A) FIGURE 12 shows concentration response curves for ATP in the presence of the P₂-purinoceptor
agonist suramin (1, 10 and 100 μM) (n = 4 for each point); and

B) FIGURE 13 shows concentration dependence of suramin, DIDS, PPADS and P5P in inhibiting the response to 10 μM ATP (n = 4 for each point).

FIGURE 14 shows the results of the functional characterisation of rat superior ganglion P2X receptors (as encoded by clone 3, described in Example 10). These experiments provided electrical recordings from transfected HEX293 cells.

Top left: Superimposed currents evoked by ATP (30 μM) during the time are indicated by the bar. Holding potential was changed from -70 to 20 mV.

Top right: Peak current as a function of membrane potential.

Bottom left: Superimposed currents evoked by ATP, from 1 to 300 μM.

Bottom right: Concentration-response curves for ATP and αβmethylene-ATP (points are mean ± s.e. mean for 5 - 8 experiments).

FIGURE 15 shows the inhibition of currents caused by various substances acting on the clone 3 form of the P2X receptor (as described in Example 11), compared with PC12 and human bladder forms in HEK293 cells.

Top: inhibition by suramin.
Middle: inhibition by PPADS.
Bottom: inhibition by pyridoxal 5-phosphate.
(i) RAT VAS DEFERENS P2X RECEPTOR

EXAMPLE 1 Cloning of the Rat vas deferens P2X Receptor
Total RNA was isolated by the guanidinium isothiocyanate method (Sambrook et al., "Molecular Cloning: A Laboratory Manual" Cold Spring Harbor Laboratory Press, second edition (1989)) from vas deferens of 4 weeks old Sprague-Dawley male rats, and the poly A+ RNA was subsequently purified by oligo(dT)-cellulose. First strand cDNA primed with the sequence 5'-GAGAGAGAGAGCGGCCGCTTTTTTTTTTTTTTTTTT-3' (SEQ ID NO 1) was synthesised with SUPERSCRIPT™ (BRL, Gaithersburg, MD, USA). After conversion of the cDNA to double stranded (Gubler & Hoffman, Gene 25 263-269 (1983)) EcoRI linkers were ligated to the cDNA, and the product was digested with NotI. The EcoRI-NotI cDNA of 1.3 to 9 kb was isolated by gel electrophoresis, and a unidirectional library was constructed by ligation of the cDNA to pBKCMV (Stratagene, San Diego, CA, USA) digested with the same enzymes. The library was electroporated into E. coli DH10B cells and divided in 24 pools of 8 x 10⁴ clones. The plasmid DNA from the pools was prepared by minialkaline lysis followed by LiCl precipitation (Sambrook et al., loc. cit). NotI-linearised cDNA was transcribed in vitro with T3 RNA polymerase in the presence of the cap analogue m7GpppG (Sambrook et al., loc. cit). The in vitro transcribed RNA (cRNA) was concentrated to 4 mg/ml.
EXAMPLE 2 Sequencing of the Rat vas deferens P<sub>2X</sub> Receptor cDNA

The cDNA insert was sequenced the exonuclease method (Henikoff Meth. Enzymol. 155 156-164 (1987)). The sequence is shown in Figure 1.

EXAMPLE 3 Functional characterisation of the Rat vas deferens P<sub>2X</sub> Receptor cDNA in Oocytes

50 nl (200 ng) of RNA was injected into defolliculated Xenopus oocytes. After incubation for 2-6 days at 18°C, the oocytes were assayed for ATP-evoked currents by a two-electrode voltage clamp (GENECCLAMP); one electrode is to hold the voltage constant (at -100 mV), and the other is to measure the currents. A cDNA pool which showed ATP induced currents was subdivided to obtain a single clone (P<sub>2X</sub>). Electrophysiological measurements were done at -100 mV, in a perfusion medium containing 96 mM NaCl, 2 mM KCl, 1.8 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 5 mM Hepes pH 7.6, and 5 mM sodium pyruvate. For dose-response curves and suramin inhibition, oocytes were injected with 100 ng P<sub>2X</sub> cRNA, and all recordings were performed at -60 mV, with Ba<sup>2+</sup> substituted for external Ca<sup>2+</sup> to prevent activation of endogenous Ca<sup>2+</sup>-activated Cl<sup>-</sup> currents. Microelectrodes (0.5-2 MΩ) were filled with 3M KCl.

EXAMPLE 4 Functional characterisation of the Rat vas deferens P<sub>2X</sub> Receptor cDNA in HEK 293 Cells

HEK 293 cells were transfected by the lipofectin method (Felgner et al., Proc. Nat’l. Acad. Sci. USA 84 7413-7417 (1987)) with P<sub>2X</sub>-plasmid. DNA concentration used was 1 mg/2 ml medium placed into a 35 mm petri dish containing four 11 mm diameter coverslips on which HEK cells were placed at 10,000 cells per coverslip. Cells were exposed to lipofectin/DNA for 6 h and recordings made 16 - 36 h
later; 40 - 60% of cells from which recordings were made exhibited $P_{2x}$ responses. Currents were recorded from HEK 293 cells using whole-cell recording methods and the AXOPATCH™ 200 amplifier (Axon Instruments); patch pipettes (5 MΩ) contained (mM) Cs or K aspartate 140, NaCl 5, EGTA 11, HEPES 5. The external solution was (mM) NaCl 150, KCl 2, CaCl$_2$ 2, MgCl$_2$ 1, HEPES 5 and glucose 11; the pH and osmolarity of both solutions were maintained at 7.3 and 305 mosmol/l respectively. All recordings performed at room temperature. Data acquisition and analysis were performed using PCLAMP™ and AXOGRAPH™ software (Axon Instruments). Solutions for experiments examining calcium permeability of ATP currents in HEK cells contained (mM): internal solution NaCl 150, HEPES 5, CaCl$_2$ 0.5 and EGTA 5 (free calcium concentration about 5 nM); external sodium solution NaCl 150, glucose 11, histidine 5, CaCl$_2$ 2; external calcium solution CaCl$_2$ 115, glucose 11 and histidine 5. The pH and osmolarity of the solutions were 7.4 and 295 mosmol/l respectively. For single channel measurements, a GENECLAMP™ 500 amplifier and outside-out recording methods were used (Adelman et al., Neuron 9 209-216 (1992)). Wax-coated patch pipettes (5 - 10 MΩ) contained (mM) K-gluconate 115, HEPES 5, BAPTA 5 and MgCl$_2$ 0.5, external solution was 96 mM NaCl, 2 mM KCl, 1.8 mM CaCl$_2$, 1 mM MgCl$_2$, 5 mM Hepes pH 7.6, and 5 mM sodium pyruvate. ATP was applied by U-tube typically for 1 s; data was sampled at 5 kHz in 2 s segments beginning 300 ms prior to onset of agonist (ATP) application and filtered at 1 kHz.

**EXAMPLE 5 Transfection of the Rat vas deferens $P_{2x}$ Receptor cDNA into CHO and HEK293 Cells**

CHO cells were stably transfected by a method used for other ion channels (Claudio, Meth. Enzymol. 207 391-408
Transfection was confirmed by a) electrophysiological recording and b) radioligand binding. ATP and other agonists (up to 30 μM) caused rapidly desensitising inward currents in 14 of 14 CHO cells stably transfected, and had no effect in 45 of 45 non-transfected cells. [3H]α,β-methyleneATP binding was more than 600 cpm per million transfected cells with less than 80 cpm nonspecific binding.

Stable transfection of HEK293 cells was also achieved. This was confirmed by electrophysiological recording.
(ii) HUMAN P$_{2X}$ RECEPTOR

The materials and methods used in the human P$_{2X}$ receptor examples are set out below:

**In Vitro translation**  
In vitro coupled transcription/translation were performed using Promega's TNT Coupled reticulocyte lysate Systems with or without 2 µl of canine pancreatic microsomal membranes (Promega). µg Circular pBKCMV-hP$_{2X}$ (0.5 µg) was transcribed with the T3 RNA polymerase as described in the system manual in a 25 µl reaction for 2 h are 30°C. Synthesized proteins (5 µl) were analysed by SDS-PAGE and autoradiography.

**Differentiation of HL60 cells**  
HL60 cells (human promyelocytes ATCC CCL240) were passaged twice weekly in RPMI-1640 supplemented with 25 mM HEPES, 2 mM Glutamax II, and 10% heat-inactivated fetal calf serum (GIBCO BRL). For each experiment 33 x 10$^6$ cells were resuspended at 2.5 x 10$^5$ cells/ml in medium containing either phorbol mystate acetate (100 nM), 1.1% DMSO, or dibutyryl cAMP (200 µM) (SIGMA) for the indicated times.

**Northern blot analysis**  
PolyA$^+$ RNAs were obtained from Clontech Laboratories Inc. (Palo Alto) except for the urinary bladder and HL60 mRNA which were prepared as described (Valera et al (1994) - supra). Samples were quantified by measuring the O.D. at 260 nm, and by staining the membrane with methylene blue. The RNA were fractionated on a 1% agarose - 6% formaldehyde gel and electroblotted to a non-charged nylon membrane (BDH). Prehybridisation at 68°C was performed for 6 hours in hybridisation buffer (50% formamide, 5X SSC, 2% blocking buffer (Boehringer Mannheim), 0.1% laurolysarscosine,
0.02% SDS). Hybridisation was overnight at 68°C in fresh hybridisation buffer with a digoxigenin-UTP labelled riboprobe (100 ng/ml) corresponding to the entire hP2X sequence. The membrane was washed at 68°C; twice in 2X SSC + 0.1% SDS, and twice in 0.1X SSC + 0.1% SDS. Chemiluminescent detection of hybridisation was carried at room temperature as follows: the membrane was rinsed 5 min in buffer B1 (0.1 M maleic acid, 0.15 M NaCl, pH 7.5), saturated for 1 hour in 1% blocking buffer (B2), incubated 30 min with anti-digoxigenin-antibody alkaline phosphatase conjugated (750 u/ml, Boehringer Mannheim) diluted 1:15000 in B2, washed in B1 + 0.3% tween 20 (1X 5 min, 1X 15 min, 1X 1 h), equilibrated for 5 min in buffer B3 (0.1 M Tris HCl pH 9.5, 0.1 M NaCl, 50 mM MgCl2), incubated 45-60 sec in lumigen PPD (Boehringer Mannheim) diluted 1:100 in B3. The humid membrane was sealed in a plastic bag, incubated 15 min at 37°C, and exposed 15 to 20 min to Hyperfilm-ECL (Amersham).

P2X expression into oocytes Human urinary bladder P2X cDNA, subcloned into the PBKCMV expression vector, was linearized with NotI, and transcribed in vitro with T3 polymerase in the presence of cap analogue m7G(5')ppp(5')G. Defolliculated Xenopus oocytes (Bertrand et al., Electrophysiology of neuronal nicotinic acetylcholine receptors expressed in Xenopus oocytes following nuclear injection of genes or cDNAs Meth Neurosci 4: 174-193 (1991)) were injected with 50 ng of human P2X in vitro transcribed RNA, and incubated at 18°C for 2-6 days in the ND96 solution (mM): NaCl96, KCl2, MgCl2 1, CaCl2 2, sodium pyruvate 5, HEPES 5, ph 7.6 - 7.5, penicillin (10 U/ml), and streptomycin (10 µg/ml).
**Electrophysiology** Oocytes were placed in a 1 ml chamber and superfused at 2 - 3 ml/min with ND96 solution with 0.1 mM BaCl₂ replacing the 2 mM CaCl₂ to prevent activation of endogenous calcium-activated chloride currents (Barish, A transient calcium-dependent chloride current in the immature Xenopus oocytes *J Physiol* **342**: 309-325 (1983)). Currents were measured using a two-electrode voltage-clamp amplifier (Geneclamp Axon Instruments) at a holding potential of -60 mV. Microelectrodes were filled with 3 M KCl (0.5 - 2 MΩ). Data were collected using PClamp software (Axon Instruments). ATP and other purinoceptor agonists were applied by a U-tube perfusion system (Fenwick et al, A patch clamp study of bovine chromaffin cells and their sensitivity to acetylcholine *J Physiol* **331**: 577-597 (1982)) placed close (200 - 500 μm) to the oocyte. Initial studies showed that reproducible responses (<10% variation in peak amplitude) could be obtained when ATP (at concentrations up to 1 mM) was applied to hP₂x injected oocytes for 5 s every 10 mins. Concentration response relationships to ATP and its analogs were determined by measuring the peak amplitude of responses to a 5 s application of agonist applied at 10 min intervals. Responses to agonists were normalized in each oocyte to the peak response evoked by 100 μM ATP; 100 μM ATP was usually applied at the beginning and at the end of an experiment to determine if there was any rundown of the response. No inward current was recorded in uninjected oocytes in response to application of purinoceptor agonists at the maximal concentration used (n = 3 for each agonist). Antagonists were applied both in the superfusate and together with ATP in the U-tube solution. Antagonists were superfused for 5 - 10 min prior to the application of ATP.
Data analysis  Concentration response curves for purinoceptor agonists were fitted with a Hill slope of 1. Equi-effective concentrations i.e. concentration of agonist, giving 50% of the response to 100 μM ATP, (EEC_{50}) were determined from individual concentration response curves. For antagonists the concentration required to give 50% inhibition (IC50) of the response to 10 μM ATP (approximately 90% of peak response to ATP) were determined. Data are presented throughout as mean ± SEM for a given number of oocytes.

Drugs  Adenosine, adenosine 5’-monophosphate sodium salt (AMP), adenosine 5’-diphosphate sodium salt (ADP), adenosine 5’-triphosphate magnesium salt (ATP), adenosine 5’-O-(-3-thiophosphate) tetralithium salt (ATP-γ-S), uridine 5’-triphosphate sodium salt (UTP), α,β-methylene ATP lithium salt (α,β,-meATP), β,γ-methylene-D-ATP sodium salt (D-β,γ-meATP), 2’-3’-O-(4-benzoylbenzol)ATP tetraethylammonium salt, (BzATP), 4,4’-diisothiocyanatostilbene 2,2’-disulphonic acid, disodium salt (DIDS) were obtained from Sigma. 2-MethylthioATP tetra sodium salt (2MeSATP), 2-chloro-ATP tetra sodium salt, and β-γ-methylene-1-ATP (1-β-γ-meATP) were obtained from RBL. Pyridoxal 5-phosphate monohydrate (Aldrich), p1, p5-di[adenosine-5’]pentaphosphate trilithium salt (AP5A) (Boehringer Mannheim), pyridoxal phosphate 6-azophenyl 2’,4’-disulphonic acid (PPADS, gift of G. Lambrecht, University of Frankfurt) and suramin (Bayer) were tested. Drugs were prepared from frozen aliquots of stock solutions and diluted to give the required final concentration.
**EXAMPLE 6 Sequence and characteristics of hP$_{2X}$ from urinary bladder**

**Isolation of human P$_{2X}$ cDNA** Human urinary bladder tissue was obtained from a cystectomy for a bladder tumor. The patient showed no symptoms of bladder instability or urodynamic abnormalities. Only those portions, surrounding the tumor, which appeared macroscopically normal (Palea et al - supra) were used. Total RNA was isolated by guanidinium isothiocyanate and poly A$^+$ RNA was purified as described (Valera et al (1994) - supra). Preparation of a cDNA library in λgt10, random primer labelling of a rat smooth muscle P$_{2X}$ probe (Valera et al (1994) - supra), low stringency hybridisation screening and lambda phage DNA isolation were all done by standard protocols (Sambrook et al, Molecular Cloning, A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, New York (1989)). Several independent phage isolates were examined and the cDNA insert from one was chosen for subcloning into Eco RI-Not I digested pBKCMV. This 2677 bp hP$_{2X}$ cDNA was sequenced as described (Valera et al (1994) - supra).

The 2677 bp cDNA, hP$_{2X}$, contained a single long open reading frame which corresponds to a protein of 399 amino acids (Figure 4). This amino acid sequence is highly homologous with that of the P$_{2X}$ receptor, isolated from rat vas deferens (89% identity). There are two regions of hydrophobicity near either end of the protein which are sufficiently long to traverse the membrane but there is no hydrophobic N-terminal leader sequence. All five potential sites for glycosylation and all ten cysteine residues in the central section of the protein are conserved. *In vitro* translation of hP$_{2X}$ RNA in the
presence of microsomes produced a 60 kD product, whereas translation in the absence of microsomes produced the 45 kD peptide (Figure 6). 45 kD is the computed molecular weight, suggesting that the additional 15 kD results from glycosylation.

Some human urinary bladder P_2X cDNA was used to transf ect HEK293 cells. Stable transfection was confirmed by electrophysiological recording.

**EXAMPLE 7 Distribution of human urinary bladder P_2X mRNA**

The distribution of the human urinary bladder P_2X mRNA was examined by northern analysis. A single 2.6 kb mRNA species was observed in bladder, placenta, liver and adrenal gland (Figure 8). In thymus, spleen, and lung samples, the 2.6 kb band plus additional higher molecular weight RNAs of 3.6 and 4.2 kb were seen. A smaller additional RNA species of 1.8 kb was observed in spleen and lung. No hybridisation was detected with brain mRNA.

**EXAMPLE 8 Induction of hP_2X mRNA in HL60 cells**

A portion of the 3'-untranslated region had been previously deposited in the database (HSGS01701) as an expressed sequence tag for the differentiation of the human promyelocytic cell line, HL60 (Okubo unpublished). We examined the induction of hP_2X mRNA in HL60 cells by Northern blot analysis (Figure 7). HL60 cells can be differentiated into distinct lineages, depending on the inductant (Koeffler, Induction of Differentiation of Human Acute Myelogenous Leukemia Cells: Therapeutic Implications Blood 62: 709-721 (1983)). Induction of macrophage-like characteristics with phorbol diesters or
granulocytic differentiation with DMSO or dibutryl cAMP, each produced an increase in P2X mRNA (Figure 7, lane 6). HL60 RNA (lane 1-5) showed hybridisation of two bands (1.8 and 2.6 kb) and both of these were inducible. This contrasts with the bladder, where Northern analysis showed only a single RNA species (2.6 kb) (Figure 7, lane 6).

**EXAMPLE 9 Pharmacological characterization of hP2X**

Application of ATP (30 nM - 1mM) to oocytes injected with hP2X receptor RNA evoked inward currents (Figures 9, 10 and 11). Responses to low concentrations of ATP (30 - 300 nM) developed over 3-5 s. Higher concentrations of ATP (1 μM) evoked responses which peaked within 1 - 1.5 s and then declined during the continued application of ATP (40 - 60% of the peak amplitude after 5 s). The current returned to control values on washout of ATP. The peak amplitude of the inward current evoked by ATP was concentration-dependent (Figures 9, 10 and 11) and could be fitted by a curve with a Hill slope of 1 with a EC50 of 0.82 μM. When ATP (100 μM) was applied for 5 s every 10 min, reproducible inward currents were recorded. This is in contrast to the responses of the P2X receptor clone from rat vas deferens where a second application of ATP (> 1 μM) applied 10 mins after the first, evoked an inward current that was ~50% of the initial peak amplitude.

Concentration-response curves were constructed for a number of other P2 purinoceptor agonists (Figures 9, 10 and 11). 2meSATP, 2-chloro-ATP, α,β-meATP and ADP were full agonists. BzATP, AP5A and ATP-γ-S produced maximal responses of about 65% of the maximal ATP response. The
maximal responses to d and 1-β,γ-meATP were not
determined. Adenosine, AMP and UTP (100 μM) evoked small
inward currents (2.3 ± 1.5, 6.08 ± 2, and 3.7 ± 1.8% of
the response to 100 μM ATP respectively). The EEC₅₀
values and relative potencies of purinceptor analogs are
summarised in Table 1 below.

Table 1

<table>
<thead>
<tr>
<th>agonist</th>
<th>EEC₅₀(μM)</th>
<th>relative potency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>0.82</td>
<td>1</td>
</tr>
<tr>
<td>2MeSATP</td>
<td>0.6 ± 0.1</td>
<td>1.36</td>
</tr>
<tr>
<td>2chloroATP</td>
<td>0.76 ± 0.1</td>
<td>1.08</td>
</tr>
<tr>
<td>AP5A</td>
<td>2 ± 0.2</td>
<td>0.41</td>
</tr>
<tr>
<td>α,β-meATP</td>
<td>3.6 ± 1.6</td>
<td>0.23</td>
</tr>
<tr>
<td>BzATP</td>
<td>4.2 ± 2.2</td>
<td>0.20</td>
</tr>
<tr>
<td>ATP-γ-S</td>
<td>10.6 ± 3.8</td>
<td>0.077</td>
</tr>
<tr>
<td>d,β,γ-meATP</td>
<td>24.1 ± 1.6</td>
<td>0.034</td>
</tr>
<tr>
<td>ADP</td>
<td>34.3 ± 16</td>
<td>0.024</td>
</tr>
</tbody>
</table>

EEC₅₀: Equi-effective concentrations producing an inward
current equivalent to 50% of the peak response to 100 μM
ATP. EEC₅₀ taken from individual fitted concentration
response curves with a Hill slope of 1. EEC₅₀ for ATP
from mean data from all experiments. (n = 3-4).

EXAMPLE 10 Antagonist studies

The P2-purinceptor antagonist suramin (1 - 100 μM)
shifted the concentration-response curve for ATP to the
right. At 1 μM suramin the shift was almost parallel.
The dissociation equilibrium constant (Kᵩ) estimated from
Kᵩ = 1/(DR-1) where DR is the dose ratio was 130 nM. With
higher concentrations of suramin the inhibition did not
appear to be competitive. Under the present experimental conditions this $K_B$ estimate is higher than those reported previously for suramin (pA2 5.9, Trezise et al, Br J Pharmacol 112: 282-288 (1994)) $pK_B$ 5.2, von Kugelgen et al, Interaction of adenine nucleotides, UTP and suramin in mouse vas deferens: suramin-sensitive and suramin-insensitive components in the contractile effect of ATP Naunyn Schmiedeberg's Arch Pharmacol 342: 198-205 (1990)). The antagonism by suramin was fully reversed after 10 mins wash and indicates that the non-competitive antagonism at high concentrations is not due to irreversible binding of the antagonist to the receptor.

The putative $P_{2X}$ purinoceptor antagonists PPADS, DIDS and pyridoxal 5 phosphate (Ziganshin et al, Selective antagonism by PPADS at $P_{2X}$ purinoceptors in rabbit isolated blood vessels Br J Pharmacol 111: 923-929 (1994), Bultmann & Starke, Blockade by 4,4'-diisothiocyanatostilben-2,2'-disulphonate (DIDS) of $P_{2X}$ purinoceptors in rat vas deferens Br J Pharmacol 112: 690-694 (1994), Trezise et al, Eur J Pharmacol 259: 295-300 (1994)) inhibited inward currents evoked by 10 $\mu$M ATP (approximately EC$_{90}$ concentration) in a concentration dependent manner (Figures 12 and 13). Suramin PPADS and DIDS were equally effective in inhibiting ATP evoked currents (IC$_{50}$ ~ 1 $\mu$M). The IC 50 for P5P was ~ 20 $\mu$M. PPADS and P5P antagonism was readily reversible on washout. In contrast, inhibitory effects of DIDS (100 $\mu$M) were very slow to reverse on washout.
(iii) RAT SUPERIOR CERVICAL GANGLION P$_{2x}$ RECEPTOR

Example 11 Isolation and functional expression of a cDNA encoding a P$_{2x}$ receptor from rat superior cervical ganglion (referred to herein as clone 3)

A 440 bp fragment was amplified by polymerase chain reaction (PCR) from rat testis cDNA, using degenerate primers based on conserved nucleotide sequences within the rat vas deferens P$_{2x}$ receptor cDNA and on the sequence of PC12 cDNA (Ehrlich H A (ed) PCR Technology MacMillan, Basingstoke (1989)). The primers used are given below:

Sense

5' T G T/C G A A/G A/G T I T T/C I G G/C I T G G T G T/C C C 3' (SEQ ID NO 2)

Antisense

5' G C A/G A A T/C C T A/G A A A/G T T A/G T/A A I C C 3' (SEQ ID NO 3)

(wherein I = Inosine and "T/C" indicates that either T or C is present at the position indicated (this applies mutatis mutandis to the other alternatives given).

The cloned PCR fragment was labelled and used as a hybridization probe for screening a rat testis cDNA bank in λZAP. One recombinant phage was positive, and its insert was excised and transferred to a plasmid (#432). This cDNA was 1500 bp with a single EcoRI site (at position 1000, still in the open reading frame). The 5' end of the cDNA was too short to encode the entire N terminus.
Internal primers specific to the new sequence were made and the tissue distribution was tested by PCR. The candidate was present in mRNA prepared from phaeochromocytoma (PC12) cells, intestine and superior cervical ganglion (scg). The hybridization probe was therefore used to screen a rat scg cDNA bank in λgt10. From 30 initial positives, 20 pure phage DNA stocks were prepared; 19 were various portions of the candidate sequence, and the insert from one was transferred to plasmid (p457) and sequenced. The insert appeared to be a full length cDNA; it has a single open reading frame of 388 amino acids (Fig. 2). The insert from p457 was subcloned into pcDNA3 (p464) and used to transfect human embryonic kidney (HEK293) cells.

The functional characterisation of the clone illustrated in Fig 2 (referred to herein as clone 3) was carried out by electrical recordings from transfected HEK293 cells and from oocytes injected with the in vitro transcribed RNA, as described in Example 4 for the rat vas deferens P2X receptor. Table A summarizes the main properties of clone 3 as compared to those of rat vas/human bladder cDNA clone, and the PC12 cDNA clone (provided by David Julius and Tony Brake of the University of California at San Francisco).
<table>
<thead>
<tr>
<th></th>
<th>bladder</th>
<th>clone 3</th>
<th>PC12</th>
</tr>
</thead>
<tbody>
<tr>
<td>kinetics</td>
<td>very strong</td>
<td>very little</td>
<td>very little</td>
</tr>
<tr>
<td>desensitization</td>
<td>profound</td>
<td>very little</td>
<td>very little</td>
</tr>
<tr>
<td>rundown</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                |                  |                  |                  |
| ionic permeability |                  |                  |                  |
| monovalent      | no differences   | no differences   | no differences   |
| divalent (Ca++) | high permeability| high permeability| high permeability|
| Ca++ block      | none             | intermediate     | very strong      |

|                |                  |                  |                  |
| agonist profile|                  |                  |                  |
| ATP            | 0.7 μM           | 11 μM            | 8 μM             |
| α,β-meATP      | 3 μM             | >>100 mM         | >>100 μM         |

|                |                  |                  |                  |
| antagonist profile |                 |                  |                  |
| suramin        | 1 μM             | < 40% block       | 6 μM             |
| PPADS          | 1 μM             | < 30% block       | 1 μM             |
| P-5-P          | 6 μM             | < 40% block       | 6 μM             |
| DIDS           | 1 μM             | > 100 μM          |                  |
The main functional properties of clone 3 are as follows. (a) The currents evoked by ATP show little or no decline during applications of several seconds; that is, there is little desensitisation (Fig. 14). (b) The relative permeabilities of the ionic pore to sodium, potassium, cesium, tetraethylammonium and to calcium are not different to those observed for the rat vas deferens/human bladder or the PC12 forms of the receptor. (c) Extracellular calcium (30 mM) inhibits the inward current through the P2X receptor channel of the PC12 form whereas it does not block current through the rat vas deferens/human bladder form; clone 3 is intermediate in sensitivity. (d) The effectiveness of agonists that are structurally related to ATP is the same as that found for the PC12 form; most notably, α,α-methylene ATP has little or no agonist action (Fig. 14). (e) Currents activated by ATP at the clone 3 receptor were much less sensitive to antagonism by suramin., pyridoxal 5'-phosphate and pyridoxal-6-azophenyl-2',4'-disulphonic acid (PPADS) than were similar current mediated by the other two forms (rat vas deferens/human bladder; PC12) (Fig. 15).
(iv) RAT DORSAL ROOT GANGLION P$_{2X}$-RECEPTOR

Example 12 Isolation of a cDNA encoding a P$_{2X}$ receptor from a rat dorsal root ganglion

By using PCR with the same primers as used in Example 11 above, but using different cDNA sources, further P$_{2X}$ family members can be found.

Using this method, rat dorsal root ganglion P$_{2X}$ receptor cDNA was isolated. Fig. 1B shows the cDNA sequence of this clone (referred to herein as clone 6), together with the putative amino acid sequence. The portions underlined in this figure correspond to the PCR primers initially used.

A similar procedure to that described in Example 11 was then used to isolate the full length cDNA.
1

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: GLAXO GROUP LIMITED
(B) STREET: GLAXO HOUSE. BERKELEY AVENUE
(C) CITY: GREENFORD
(D) STATE: MIDDLESEX
(E) COUNTRY: UNITED KINGDOM
(F) POSTAL CODE (ZIP): UB6 0NN

(ii) TITLE OF INVENTION: DNA AND PROTEIN SEQUENCES

(iii) NUMBER OF SEQUENCES: 11

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0. Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
GAGAGAGA GCGGCCACTT TTTTTTTTTTTTTTTTTT

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 3
   (D) OTHER INFORMATION:/mod_base= OTHER
      /note= "T or C"

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 6
   (D) OTHER INFORMATION:/mod_base= OTHER
      /note= "A or G"

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 7
   (D) OTHER INFORMATION:/mod_base= OTHER
      /note= "A or G"

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 9
   (D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 11
   (D) OTHER INFORMATION:/mod_base= OTHER
      /note= "T or C"

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 12
   (D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 14
   (D) OTHER INFORMATION:/mod_base= OTHER
      /note= "G or C"

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 15
   (D) OTHER INFORMATION:/mod_base= i

(ix) FEATURE:
   (A) NAME/KEY: modified_base
   (B) LOCATION: 21
   (D) OTHER INFORMATION:/mod_base= OTHER
      /note= "T or C"

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
TGNGANTNT NIGNNTGGTG NCC
(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 3
(D) OTHER INFORMATION:/mod_base= OTHER
   /note= "A or G"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION:/mod_base= OTHER
   /note= "T or C"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION:/mod_base= OTHER
   /note= "A or G"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 12
(D) OTHER INFORMATION:/mod_base= OTHER
   /note= "A or G"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION:/mod_base= OTHER
   /note= "A or G"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 16
(D) OTHER INFORMATION:/mod_base= OTHER
   /note= "T or A"

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 18
(D) OTHER INFORMATION:/mod_base= i

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

GCNAANCTNA ANTNNANCC
(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1387 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(B) CLONE: rat P2x from vas deferens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:210..1406

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
GCCAAGAGT GCTCTTACCA CCCCAGTTT TCTCTCCAA CCCCCCCCC ACCCTGACAC 60
CTCCAAGTG TGGCCACCT AGGCCTGCTC GTCTCTTCAG GGGCGGGGAG CCCGAGTCAC 120
TCCACTGCTA TTGTAGATGC AGATGGTGAC CTGCCCCGAG CCATAGAGCC CTGTCCTGGG 180
GTTCATCTCT GAGCCCCCTTC TGCCCCACC ATG GCT CGG GCG CTG CAA GAT GAG 233

CTG TCA GCC TTC TTC TTT GAA TAT GAC ACT CCC CGG ATG GTG CTG GTA 281
Leu Ser Ala Phe Phe Phe Glu Tyr Asp Thr Pro Arg Met Val Leu Val
10 15 20

CGA AAC AAG GTG GGA GTC ATT TTC CTG CTG ATC CAG TG TGG GTT 329
Arg Asn Lys Lys Val Gly Val Val Ile Phe Arg Leu Ile Gln Leu Val
25 30 35 40

CTG GTC TAC TGT ATT GGG TGG GTT TTT GTG TAT GAA AAA GGA TAC CAG 377
Leu Val Tyr Val Ile Gly Trp Val Phe Val Tyr Gly Lys Trp Gly
45 50 55

ACC TCA AGT GAC CTC ATC AGC AGT GTG TCC GTG AAG CTC AAG GGC TTG 425
Thr Ser Ser Asp Leu Ile Ser Ser Val Ser Val Lys Leu Gly Leu
60 65 70

GCT GTG ACC CAG CTC CAG GGC GCA GCA CCC CAG GTC TGG GAC GTG GCT 473
Ala Val Thr Gln Leu Gln Gly Leu Gly Pro Gln Val Trp Asp Val Ala
75 80 85

GAC TAT GTC TTC CCA GCA CAC GGG GAC AQC CCC TTT GTA GTT ATG ACC 521
Asp Tyr Val Phe Pro Ala His Gly Asp Ser Ser Phe Val Val Met Thr
90 95 100

AAC TTC ATG ACC CCT CAG ACT CAA GGC CAT TGT GCA GAG AAC 569
Asn Phe Ile Val Thr Pro Gln Gln Thr Gln Gly His Cys Ala Glu Asn
105 110 115 120
```
CCA GAA GGT GCC ATA TGC CAG GAT GAC AGT GGC TGC ACT CCA GGA AAA
       Pro Glu Gly Gly Ile Cys Gin Asp Ser Gly Cys Thr Pro Gly Lys
       125          130                     135

GCA GAA AGG AAA GCC CAA GGT ATT CGC ACA GGC AAC TGT GTG CCC TTC
       Ala Glu Arg Lys Ala Gin Gly Ile Arg Thr Gly Asn Cys Val Pro Phe
       140                     145                     150

AAT GGC ACT GTG AAG ACA TGT GAG ATC TTC GTG TGG TGT CCT GTA GAG
       Asn Gly Thr Val Lys Thr Cys Ile Phe Gly Trp Cys Pro Val Glu
       155                     160                     165

GTG GAT GAC AAG ATC CCA AGC CCT GCT CTT CTG GAG GCT GAG AAC
       Val Asp Asp Lys Ile Pro Ser Pro Ala Leu Arg Glu Ala Glu Asn
       170                     175                     180

TTC ACC CTC TTC ATC AAA AAC AGC ATC AGC TTT CCA CGC TTC AAG GT
       Phe Thr Leu Phe Ile Lys Asn Ser Ile Ser Phe Pro Arg Phe Lys Val
       185                     190                     195                     200

AAC AGG CGC AAC CTG GTA GAG GAG GTG AAC GCC ACC ACG ATG AAG AAC
       Asn Arg Arg Asn Leu Val Glu Glu Val Asn Gly Thr Tyr Met Lys Lys
       205                     210                     215

TGC CTC TAT CAC AAG ATT CAA CAC CCC CTG TGC CCA GTT ACG CCT
       Cys Leu Tyr His Lys Ile Gin His Pro Leu Cys Pro Val Phe Asn Leu
       220                     225                     230

GCC TAT GTG GTG CGA GAG TCA GCC CAG GAC TTC CGC AGC CCT GCT GAG
       Gly Tyr Val Val Arg Glu Ser Gly Gin Asp Phe Arg Ser Leu Ala Glu
       235                     240                     245

AAG GGT GGG GTG GTT GGT ATC ACC ATT GAC TGG AAG TGT GAT CTG GAC
       Lys Gly Gly Val Val Gly Ile Thr Ile Asp Trp Lys Cys Asp Leu Asp
       250                     255                     260

TGG CAC GTT CGG CAC TGC AAA CCC ATC TAC CAG TTC CAC GGA CTG TAT
       Thr His Val Arg His Cys Lys Pro Ile Tyr Gin Phe His Gly Leu Tyr
       265                     270                     275                     280

GGG GAG AAG AAC CTG TCT CCA GCC TTC AAG TCC GTA TCC GCC AGG CAT
       Gly Glu Lys Asn Leu Ser Pro Gly Phe Asp Phe Arg Phe Ala Arg His
       285                     290                     295

TTC GTG CAG AAT GGG ACA AAC CGT CTG CAC CTC TTC AAG GTG TTT GGG
       Phe Val Gin Asn Thr Asn Arg Gin Phe Asp Leu Val Val Phe Gly
       300                     305                     310

ATT CAC TTT GAT ATC CTT GTG GAT GGC AAG GCT GGG AAG TTT GAC ATC
       Ile His Phe Asp Ile Leu Val Asp Gly Lys Ala Gly Asp Ile
       315                     320                     325

ATC CCT ACT ATG ACT ACT ATC GTC TCT GGG ATT GCC AGC TTC TTT GGA GTG
       Ile Pro Thr Met Thr Thr Ile Gly Ser Gly Ile Gly Phe Gly Val
       330                     335                     340

GCC ACA GTT TGT GAT CTC TTA TGT CTC CAC ATG CTG CAC TTG AAG
       Ala Thr Val Leu Cys Asp Leu Leu Leu His Ile Leu Pro Lys Arg
       345                     350                     355                     360
CAC TAC TAC AAG CAG AAG AAG TTC AAA TAT GCC GAG GAC ATG GGG CCG 1337
His Tyr Tyr Lys Gin Lys Lys Phe Lys Tyr Ala Glu Asp Met Gly Pro
365

GGA GAG GGT GAA CAT GAC CCC GTG GCC ACC AGC TCC ACT CTG GSC CTG 1385
Gly Glu Gly Glu His Asp Pro Val Ala Thr Ser Ser Thr Leu Gly Leu
390

CAG GAG AAC ATG AGG ACC TCC TGACCTTAGT CTTGAGATCC GGACTTGAGC 1436
Gln Glu Asn Met Arg Thr Ser
395

CAGTGTGTGG CTTCCGGCAA GGGCGATGG CTTTGAAGCCA GGCCAGAGG CATTCCAGA 1496
GGCTTTCCTG CAAGGCAGAC ACCAGTGSCC CTCGCTTCA GCATGAAGAC AGGCAAGACT
1556
TTGGATTTCA GAGCTCTTGT TTCGTTCCA CATGCTCCTT CCTGAGGAGT GCCTCTCCA
1616
GTTCACCA ATTTGGGTTTC ATATGATCGG GCCCCTCACA CATCTTACT CTAGCTTTGT 1676
GCCACGGCT CAGGCTGTCA TTGCTTTTCC CACACCTTTA CCTGCTTAA ATATGCTCT
1736
TCCACATGCT AGCCACTGAC CAGATGTGCT AGTATTGACT TTAATAAAA TATAAAAAA
1796
AAAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA A
1837
(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

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

Met Ala Arg Arg Leu Gln Asp Glu Leu Ser Ala Phe Phe Gly Tyr
1  5 10
Asp Thr Pro Arg Met Val Leu Val Arg Asn Lys Lys Val Gly Val Ile
20 25 30
Phe Arg Leu Ile Gin Leu Val Val Leu Val Tyr Val Ile Gly Trp Val
35 40 45
Phe Val Tyr Glu Lys Gly Tyr Gin Thr Ser Ser Asp Leu Ile Ser Ser
50 55 60
Val Ser Val Lys Leu Lys Gly Leu Ala Val Thr Gin Leu Gin Gly Leu
65 70 75 80
Gly Pro Gin Val Trp Asp Val Ala Asp Tyr Val Phe Pro Ala His Gly
85 90 95
Asp Ser Ser Phe Val Val Met Thr Asn Phe Ile Val Thr Pro Gin Gin
100 105 110
Thr Gin Gly His Cys Ala Glu Asp Pro Gly Gly Gly Ile Cys Gin Asp
115 120
Asp Ser Gly Cys Thr Pro Gly Lys Ala Glu Arg Lys Ala Gin Gly Ile
130 135 140
Arg Thr Gly Asn Cys Val Pro Phe Asn Gly Thr Val Lys Thr Cys GLu
145 150 155 160
Ile Phe Gly Trp Cys Pro Val Glu Val Asp Asp Lys Ile Pro Ser Pro
165 170 175
Ala Leu Leu Arg Glu Ala Glu Asp Phe Thr Leu Phe Ile Asn Ser
180 185 190
Ile Ser Phe Pro Arg Phe Lys Val Asn Arg Arg Asn Leu Val Glu Glu
195 200 205
Val Asn Gly Thr Tyr Met Lys Lys Cys Leu Tyr His Lys Ile Gin His
210 215 220
Pro Leu Cys Pro Val Phe Asn Leu Gly Tyr Val Val Arg Glu Ser Gly
225 230 235 240
Gln Asp Phe Arg Ser Leu Ala Glu Lys Gly Gly Val Val Gly Ile Thr
245 250 255
Ile Asp Trp Lys Cys Asp Leu Asp Trp His Val Arg His Cys Lys Pro
260 265 270
Ile Tyr Gin Phe His Gly Leu Tyr Gly Glu Lys Asn Leu Ser Pro Gly
275 280 285
Phe Asn Phe Arg Phe Ala Arg His Phe Val Gin Asn Gly Thr Asn Arg
290 295 300
Arg His Leu Phe Lys Val Phe Gly Ile His Phe Asp Ile Leu Val Asp
305 310 315 320
Gly Lys Ala Gly Lys Phe Asp Ile Ile Pro Thr Met Thr Thr Ile Gly
325 330 335
Ser Gly Ile Gly Ile Phe Gly Val Ala Thr Val Leu Cys Asp Leu Leu
340 345 350
Leu Leu His Ile Leu Pro Lys Arg His Tyr Tyr Lys Gin Lys Lys Phe
355 360 365
Lys Tyr Ala Glu Asp Met Gly Pro Gly Glu Gly Glu His Asp Pro Val
370 375 380
Ala Thr Ser Ser Thr Leu Gly Leu Gin Glu Asn Met Arg Thr Ser
385 390 395
(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1997 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(B) CLONE: rat P2x clone 3

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 101..1264

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

```
GCCAGCCGAGC TGCCGGGAGC TGTTGCGTGGG AGCTAGCACC CGGAGCCGAC GGGGCGGAGG
GGACCCACAG TGTCGCAAGGC CGGAGGCGGT CGGCGGGAGCC ATG GCG GGC TGC TGC
 Met Ala Gly Ser Cys
400

TCC GTG CTC GGG TCC TTC CTG TGG TAC GAC ACG CGG ATC GTG
Ser Val Leu Gly Ser Phe Leu Phe Glu Tyr Asp Thr Pro Arg Ile Val
405 410 415 420

CTC ATC CGC AGC GTG AAA GTG GGG CTC ATG AAG CGC GGC GTG CAG CTG
Leu Ile Arg Ser Arg Lys Val Gly Leu Met Asn Arg Ala Val Gln Leu
425 430 435

CTC ATC CTG GCT TAC GTC ATC GGG TGG GTG TTC GTG TGG GAA AAG GGC
Leu Ile Leu Ala Tyr Val Ile Gly Trp Val Phe Val Trp Glu Gly
440 445 450

TAC CAG GAA ACG GAC TCC GTG GTG ATG GCC TCG GTG ACA ACC AAA GCC AAA
Tyr Gin Glu Thr Asp Ser Val Val Ser Ser Val Thr Thr Lys Ala Lys
455 460 465

GST GTG CCT GTG ACC AAC ACC TCT CAG CCT GGA TTC CGG ATG TGG GAC
Gly Val Ala Thr Asn Ser Gln Leu Gly Phe Arg Ile Trp Asp
470 475 480

GTG GCC GAC TAT GTG ATT CCA GCT CAG GAG GAA AAC TCC CTC TTC ATT
Val Ala Asp Tyr Val Ile Pro Ala Gln Glu Asn Ser Leu Phe Ile
485 490 495 500

ATG ACC AAC ATG ATT GTG ACC GTG AAC CAG ACA CAG ACG ACC TGT CCA
Met Thr Asn Met Ile Val Thr Val Asn Gin Thr Gin Ser Thr Cys Pro
505 510 515

GAG ATT CCT GAT AAG ACC AGC ATT TGT AAT TCA GAC GCC GAC TGC ACT
Glu Ile Pro Asp Lys Thr Ser Ile Cys Asn Ser Asp Ala Asp Cys Thr
520 525 530
```
CCT GCC TCC GTG GAC ACC CAC AGC AGT GGA GTT GCG ACT GGA AGA TGT
Pro Gly Ser Val Asp Thr His Ser Ser Gly Val Ala Thr Gly Arg Cys
535
540
545

GTT CCT TTC AAT GAG TCT GTG AAG ACC TGT GAG GTG GCT GCA TGS TGC
Val Pro Phe Asn Glu Ser Val Lys Thr Cys Glu Val Ala Ala Trp Cys
550
555
560

CCG GTG GAG AAC GAC GTT GCC GTG CCA ACG CCG GCT TTC TTA AAG GCT
Pro Val Glu Asp Ser Gly Val Gly Val Pro Thr Pro Ala Phe Leu Lys Ala
565
570
575
580

GCA GAA AAC TTC ACC CTC TTG GTA AAG AAC AAG ATC TGG TAC CCC AAG
Ala Glu Asn Phe Thr Leu Leu Val Lys Asn Asn Ile Trp Tyr Pro Lys
585
590
595

TTT AAC TTC AGC AAG AGG AAC ATC CTC CCC AAC ATC ACC ACG TTC TAC
Phe Asn Phe Ser Lys Arg Asn Ser Leu Pro Asn Thr Ser Tyr
600
605
610

CTC AAA TCG TGC ATT TAC AAT GCT CAA ACG GAT CCC TTC TGC CCC ATA
Leu Lys Ser Cys Ile Tyr Asn Ala Gln Thr Asp Pro Phe Cys Pro Ile
615
620
625

TTC GGT CTG GGC ACA ATC GTG GGG GAC GCC GCA CAT AGC TTC CAG GAG
Phe Arg Leu Gly Thr Ile Val Gly Asp Ala Gly His Ser Phe Gln Glu
630
635
640

ATG GCA GTT GAG GGA GGC ATC ATG GGT ATC CAG ATC AAG TGG GAC TGC
Met Ala Val Glu Gly Gly Ile Met Gly Ile Gln Ile Lys Trp Asp Cys
645
650
655
660

AAC CTG GAT AGA GCC GCC TTC TTT GGC CTG CCC AGA TAT TTC CCG
Asn Leu Asp Arg Ala Ser Leu Cys Leu Pro Arg Tyr Ser Phe Arg
665
670
675

CGC CTG GAC ACC CGG GAC CTG GAA CAC ATG GTG TCT CCT GGC TAC AAT
Arg Leu Asp Thr Arg Asp Leu Glu His Asn Val Ser Pro Gly Tyr Asn
680
685
690

TTC AGG TTT GGC AAC AAG TAC TAC AAG GAC CTG GCC GCC AAA GAG CAG CGG
Phe Arg Phe Ala Lys Tyr Tyr Arg Asp Leu Ala Gly Lys Glu Gin Arg
695
700
705

ACA CTC ACC AAG GCG TAC GCC GTC TTT GAC ATC ATG GTG TTT GGA
Thr Leu Thr Lys Ala Tyr Gln Ile Arg Phe Asp Ile Ile Val Phe Gly
710
715
720

AAG GCT GGG AAG TTT GAC ATC ATC CTT ACC ATG AAG GTC TCT
Lys Ala Gly Lys Phe Asp Ile Ile Pro Thr Met Ile Asn Val Gly Ser
725
730
735
740

GCC GTG GCC CTC GCC GTC GCC GCG ATG GTG TCT GAC ATC GTC
Gly Leu Ala Leu Leu Gly Val Ala Thr Val Leu Cys Asp Val Ile Val
745
750
755

CTC TAC TGC ATG AAG AAG AAA TAC TAC TAC CGG GAC AAG AAA TAT AAG
Leu Tyr Cys Met Lys Lys Tyr Tyr Tyr Arg Asp Lys Lys Tyr Lys
760
765
770
TAT GTG GAA GAC TAC GAG CAG GGT CTT TCG GGG GAG ATG AAC CAG
Tyr Val Glu Asp Tyr Glu Gln Gly Leu Ser Gly Glu Met Asn Gln
775
780
TGACGCTTAA AGTATATTCC ACCCGCCT ACGCCCGCG AGCAGAAAAG TGGGAGAGA
TGCTACTGC GTGGTGCACT CTAGAAGAAG CTCCAGAGTT TCACGCAGT TCTCCACTCC
ACAATAACTG AGGTGGCGCA AGGCACATCTT GTGGAGCCTT GGCCTTTGCT CTGGCTGCTCA
GATGGGCTTC CAGATACAGA AATCTCCTGT CTTCGTGGCCTC TAGAAATGCT GGGATCAAC
ATGTCACTTG CAATTCCCAAT TTCCCATGGG GAGTTCGCGCA TTTTTATCAT TTTACCCATT
CTTTTTGAT ACATCTAAGG CTGCCCTCAG ACCGAAGACG TTTCTACCAC CTATACACCC
TTTTATCTC ACTGTGTTGT GGGGGGGGGT GTTTTGACCA CGACGCAAGG TGGATGTCTG
GTGTGCTGGT GGCTGGGCG CACTGGCTTT ATACAGTGTG AGGSTATGSA GGTAGGAAGG
GTGGGGCGC AGGACACTC GTGGTCCTTA CGACAGGGCC CAGGCTGCTG CACGCACTTC
TATTCTATAG AAGAGGGGCT CTCTCGAGTG CTGTCAGCAG GCCTGGGACA CCACTTCTCT
TCCCTATAAT CAGAACAGTT GTCCTTGATGAGGCGAGG GTTAGCTTTCA GCTTTTTAAA
GGGCTGTGTT AAAAGTGCAT AGAGCCAAAAC ATAAAGAAA ATAATTATTTT AAAAAAAA
AAA
1264
1324
1384
1444
1504
1564
1624
1684
1744
1804
1864
1924
1984
1997
(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 388 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Ala Gly Cys Cys Ser Val Leu Gly Ser Phe Leu Phe Glu Tyr Asp
1 5 10 15
Thr Pro Arg Ile Val Leu Ile Arg Ser Arg Lys Val Gly Leu Met Asn
20 25 30
Arg Ala Val Gin Leu Leu Ile Leu Ala Tyr Val Ile Gin Trp Val Phe
35 40 45
Val Trp Glu Lys Gly Tyr Gin Glu Thr Asp Ser Val Val Ser Ser Val
50 55 60
Thr Thr Lys Ala Lys Gly Val Ala Val Thr Asn Thr Ser Gin Leu Gly
65 70 75 80
Phe Arg Ile Trp Asp Val Ala Asp Tyr Val Ile Pro Ala Gin Gly Gin
85 90 95
Asn Ser Leu Phe Ile Met Thr Asn Met Ile Val Thr Val Gin Gin Thr
100 105 110
Gln Ser Thr Cys Pro Glu Ile Pro Asp Lys Thr Ser Ile Cys Asn Ser
115 120 125
Asp Ala Asp Cys Thr Pro Gly Ser Val Asp Thr His Ser Gin Gly Val
130 135 140
Ala Thr Gly Arg Cys Val Pro Phe Asn Glu Ser Val Lys Thr Cys Glu
145 150 155 160
Val Ala Ala Trp Cys Pro Val Glu Asp Val Gly Val Pro Thr Pro
165 170 175
 Ala Phe Leu Lys Ala Ala Gin Phe Thr Leu Val Lys Asn
180 185 190
Ile Trp Tyr Pro Lys Phe Asn Phe Ser Lys Arg Asn Ile Leu Pro Asn
195 200 205
Ile Thr Thr Ser Tyr Leu Lys Ser Cys Ile Tyr Asn Ala Gin Thr Asp
210 215 220
Pro Phe Cys Pro Ile Phe Arg Leu Gly Thr Ile Val Gly Asp Ala Gly
225 230 235 240
His Ser Phe Gin Glu Met Ala Val Glu Gly Gly Ile Met Gly Ile Gin
245 250 255
Ile Lys Trp Asp Cys Asn Leu Asp Arg Ala Ala Ser Leu Cys Leu Pro
260 265 270
Arg Tyr Ser Phe Arg Arg Leu Asp Thr Arg Asp Leu Glu His Asn Val
275 280 285
Ser Pro Gly Tyr Asn Phe Arg Phe Ala Lys Tyr Tyr Arg Asp Leu Ala
290 295 300
Gly Lys Glu Gln Arg Thr Leu Thr Lys Ala Tyr Gly Ile Arg Phe Asp
305 310 315 320
Ile Ile Val Phe Gly Lys Ala Gly Lys Phe Asp Ile Ile Pro Thr Met
325 330 335
Ile Asn Val Gly Ser Gly Leu Ala Leu Gly Val Ala Thr Val Leu
340 345 350
Cys Asp Val Ile Val Leu Tyr Cys Met Lys Lys Tyr Tyr Arg
355 360 365
Asp Lys Tyr Lys Tyr Val Glu Asp Tyr Glu Gln Gly Leu Ser Gly
370 375 380
Glu Met Asn Gin
385
(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1753 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(B) CLONE: rat P2x clone 6

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 163..1353

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

CACTGCGGCTA CAGTGGCGTG GCCTACAGGA ACTGGCTCTT TTCTCTCAA GCCTTAAAC GCCTTTACAG 60
AGCCGACTCC AGTTCTTGAT CTCTGTCTCC CAGTCTGAA GCTCCTTTCTC TCTTACGGCT 120
GCATCCACAG CCCCTCTAAG TGCCGCTGAG CAGTTTCTCA GT ATG AAG TGT ATA 174
Met Asn Cys Ile 390
TCA GAC TCT TTC ACC TAC GAG ACT AAG TCG GTG GTG AAG AGC 222
Ser Asp Phe Phe Thr Tyr Glu Thr Thr Lys Ser Val Val Lys Ser 395
395
TGG ACC ATT GGG ATC ATC AAC CGA GCC GTC CAG CTG ATG ATT ATC TCC 400
Trp Thr Ile Gly Ile Ile Asn Arg Ala Val Gln Leu Leu Ile Ile Ser 405
Trp Thr Ile Gly Ile Ile Asn Arg Ala Val Gln Leu Leu Ile Ile Ser 410
TGG ACC ATT GGG ATC ATC AAC CGA GCC GTC CAG CTG ATG ATT ATC TCC 415
Trp Thr Ile Gly Ile Ile Asn Arg Ala Val Gln Leu Leu Ile Ile Ser 420
410
TAC TTT GTG GGG TGG GTT TTC TTG CAT GAG AAG GCC TAC CAA GTG AGS 318
Tyr Phe Val Gly Trp Val Phe Leu His Glu Lys Ala Tyr Gln Val Arg
425
GAC ACC GCC ATT GAG TCC TCA GTA GTT ACA AAG GTG AAA GGC TTC GGG 366
Asp Thr Ala Ile Glu Ser Ser Val Val Thr Lys Val Lys Gly Phe Gly
445
GAC ACC GCC ATT GAG TCC TCA GTA GTT ACA AAG GTG AAA GGC TTC GGG 450
Asp Thr Ala Ile Glu Ser Ser Val Val Thr Lys Val Lys Gly Phe Gly
455
CGC TAT GCC AAC AGA GTC ATG GAC GTG TCG GAT TAT GTG ACC CCA CCC 414
Arg Tyr Ala Asn Arg Val Met Asp Val Ser Asp Tyr Val Thr Pro Pro
460
465
470
CGC TAT GCC AAC AGA GTC ATG GAC GTG TCG GAT TAT GTG ACC CCA CCC 462
Arg Tyr Ala Asn Arg Val Met Asp Val Ser Asp Tyr Val Thr Pro Pro
Gln Gly Thr Ser Val Phe Val Ile Ile Thr Lys Met Ile Val Thr Gly 475
AAT CAA ATG CAA GGA TTC TGT CCA GAG AAT GAA GAG AAG TAC CGC TGT 510
Asn Gln Met Glu Gly Phe Cys Pro Glu Asn Glu Lys Tyr Arg Cys
490
495
500
GTG TCT GAG AGC CAG TGT GGG CCT GAA GCC TCC CCA GST GGG GGG ATC 558
Val Ser Asp Ser Glu Cys Gly Pro Glu Arg Phe Pro Gly Gly Gly Ile
505
510
515
520
AAC CCA GTG TTC GCC AGT GAC CAG GCC ACT GTG GAG AAG CAG TCT ACA 1326
Asn Pro Val Phe Ala Ser Asp Gin Ala Thr Val Glu Lys Gin Ser Thr 765 770 775

GAC TCA GGG GCC TAT TCT ATT GGT CAC TAGGGCCTCT TCCCCAGGTT 1373
Asp Ser Gly Ala Tyr Ser Ile Gly His 780 785

CCATGGCTAC CCGTAGGCTG CAGAACCTGC AAACAGGCCA CTCTATCTAA GCAGTCAGGG 1433
GTGGGAGGG GAGAGAAGGG CTCGCTATTT CGCTGCTTCA CCCCCAAAGAC TAGATCCAGA 1493
TATCTAGGCC CTCACTGTTTC AACAGATTGG CAAGTTTCC CACTAAAGCT TGAATCCTTGC 1553
CTTTACCCCT TGCATGCGCT CCACTGGCTT CCCTGGATCC CAGSACAGCA GCATCCACCC 1613
CTTTCCAAAG GATTGAGAAA ATGCTGCTA AGTTAACCC CATAGGACCT ACCACGTACC 1673
AAGCAGCTCC ACACATAATTA TCCCTTTTCA CCCCTAAAAT AATCCCTATAA GCTAGAAAA 1733
AAAAAAAAAA AAAAAAAAA 1753
(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 397 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Asn Cys Ile Ser Asp Phe Thr Tyr Glu Thr Thr Lys Ser Val
1   5  10  15
Val Val Lys Ser Trp Thr Ile Gly Ile Ile Asn Arg Ala Val Gln Leu
20  25  30
Leu Ile Ile Ser Tyr Phe Val Gly Trp Val Phe Leu His Glu Lys Ala
35  40  45
Tyr Gln Val Arg Asp Thr Ala Ile Gln Ser Ser Val Val Thr Lys Val
50  55  60
Lys Gly Phe Gly Arg Tyr Ala Asn Arg Val Met Asp Val Ser Asp Tyr
65  70  75  80
Val Thr Pro Pro Gln Gly Thr Ser Val Phe Val Ile Ile Thr Lys Met
85  90  95
Ile Val Thr Glu Asn Gln Met Gln Gly Phe Cys Pro Glu Asn Glu Glu
100 105 110
Lys Tyr Arg Cys Val Ser Asp Ser Gln Cys Gly Pro Glu Arg Phe Pro
115 120 125
Gly Gly Gly Ile Leu Thr Gly Arg Cys Val Asn Tyr Ser Ser Val Leu
130 135 140
Arg Thr Cys Glu Ile Gln Gly Trp Cys Pro Thr Glu Val Asp Thr Val
145 150 155 160
Glu Met Pro Ile Met Met Glu Ala Glu Asn Phe Thr Ile Phe Ile Lys
165 170 175
Asn Ser Ile Arg Phe Pro Leu Phe Asp Gln Asp Phe Glu Lys Gly Asn Leu Leu
180 185 190
Pro Asn Leu Thr Asp Lys Asp Ile Lys Arg Cys Arg Phe His Pro Glu
195 200 205
Lys Ala Pro Phe Cys Pro Ile Leu Arg Val Gly Asp Val Val Lys Phe
210 215 220
Ala Gly Glu Asp Phe Ala Lys Leu Ala Arg Thr Gly Gly Val Leu Gly
225 230 235 240
Ile Lys Ile Gly Trp Val Cys Asp Leu Asp Lys Ala Trp Asp Gln Cys
245 250 255
Ile Pro Lys Tyr Ser Phe Thr Arg Leu Asp Gly Val Ser Glu Lys Ser
260 265 270
Ser Val Ser Pro Gly Tyr Asn Phe Arg Phe Ala Lys Tyr Tyr Lys Met
   275  280  285
Glu Asn Gly Ser Glu Tyr Arg Thr Leu Leu Lys Ala Phe Gly Ile Arg
   290  295  300
Phe Asp Val Leu Val Tyr Gly Asn Ala Gly Lys Phe Asn Ile Ile Pro
   305  310  315  320
Thr Ile Ile Ser Ser Val Ala Ala Phe Thr Ser Val Gly Val Gly Thr
   325  330  335
Val Leu Cys Asp Ile Ile Leu Leu Asn Phe Leu Lys Gly Ala Asp His
   340  345  350
Tyr Lys Ala Arg Lys Phe Glu Val Thr Glu Thr Thr Leu Lys Gly
   355  360  365
Thr Ala Ser Thr Asn Pro Val Phe Ala Ser Asp Gln Ala Thr Val Glu
   370  375  380
Lys Gln Ser Thr Asp Ser Gly Ala Tyr Ser Ile Gly His
   385  390  395
(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2643 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: human P2x

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 174..1370

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

```
GCTTCAGCT GACCTCTGAC TCCTGCTCTC TGGCTCCACC TGCACGGCC TGCCTCTCCT
60
AAGGGGCCGG AGAGCCCCCA GAAGGCTTAC CATGAGCTG GTGAGGGGCA CCGGCTTCAC
120
CCTGAGAGCA GAGGCGGTGC AGGGGGCTCA GTTCTGAGCC CAGCCGGCCC ACC ATG
176
Met

GCA CGG CGG TTC CAG GAG GAG CTG GCC GCC TTC CTC TTC GAG TAT GAC
Ala Arg Arg Phe Gin Gin Leu Ala Ala Phe Leu Phe Glu Tyr Asp
400 405 410

ACC CCC CGC ATG GTG CTG GTG CGT AAT AAG AAG GTG GCC GTT ATC TTC
Thr Pro Arg Met Val Leu Val Arg Asn Lys Val Gly Val Ile Phe
415 420 425 430

CGA CTG ATC CAG CTG CTG GTC ATC TAC GTC ATC GCC TGG GTG TTT
Arg Leu Ile Gin Leu Val Val Tyr Val Ile Gly Trp Val Phe
435 440 445

CTC TAT GAG AAG GCC TAC ACC TCG AGC GCC CTC AGC AGT GTC
Leu Tyr Glu Gin Asp Arg Gin Arg Gin Leu Tyr Ser Ser Gin Ile Ser Ser Val
450 455 460

TCT GTT AAA CTC AAG GCC CTG GTG GCC GTG ACC CTC CCT GCC CTC GCC
Ser Val Leu Leu Leu Ala Val Thr Gin Leu Pro Gin Leu Gly
465 470 475

CCC CAG GTG TGG GAT GTG GTC GAC TAC TCT ACC CCA GCC CAG GGG GAC
Pro Gin Val Trp Asp Val Ala Asp Tyr Val Phe Pro Ala Gin Gly Asp
480 485 490

AAC TCC TTC GTG GTG ATG ACC AAT TTC ATC GTG ACC CCG AAG CAG ACT
Asn Ser Phe Val Val Met Thr Asn Phe Ile Val Thr Pro Lys Gin Thr
495 500 505 510
```
(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 399 amino acids
  (B) TYPE: amino acid
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

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

1. A recombinant or isolated DNA molecule encoding a $P_{2x}$ receptor, wherein the receptor:
   (a) has the amino sequence shown in Figure 1, Figure 2, Figure 3 or Figure 4; or
   (b) is substantially homologous to the sequence shown in Figure 1, Figure 2, Figure 3 or Figure 4;

or a fragment of such a DNA molecule, which fragment includes at least 15 nucleotides taken from nucleotides 1 to 813 shown in Figure 1, from the full nucleotide sequences shown in Figures 2 and 3, or from nucleotides 1 to 1744 shown in Figure 4.

2. A recombinant or isolated DNA molecule encoding a $P_{2x}$ receptor, wherein the receptor:
   (a) has the amino sequence shown in Figure 1 or Figure 4; or
   (b) is substantially homologous to the sequence shown in Figure 1 or Figure 4;

or a fragment of such a DNA molecule, which fragment includes at least 15 nucleotides taken from nucleotides 1 to 813 shown in Figure 1 or from nucleotides 1 to 777 shown in Figure 4.

3. A recombinant or isolated DNA molecule encoding a $P_{2x}$ receptor, wherein the receptor:
   (a) has the amino sequence shown in Figure 1; or
   (b) is substantially homologous to the sequence shown in Figure 1;

or a fragment of such a DNA molecule, which fragment includes at least 15 nucleotides taken from nucleotides 1 to 813 shown in Figure 1.
4. A DNA molecule as claimed in any of claims 1 to 3, which encodes a human $P_{2\times}$ receptor.

5. A DNA molecule as claimed in any of claims 1 to 4, which is cDNA.

6. A DNA molecule as claimed in any of claims 1 to 5, which is in the form of a vector.

7. A host cell transformed or transfected with a vector as described in claim 6.

8. A host cell as claimed in claim 7 which is a stably transfected mammalian cell which expresses a $P_{2\times}$ receptor.

9. A preparation of $P_{2\times}$ receptor which is free of protein with which it is naturally associated.

10. A preparation of $P_{2\times}$ receptor which is free of $P_{2\times}$ receptor.

11. $P_{2\times}$ receptor as prepared by recombinant DNA technology.

12. A peptide fragment of $P_{2\times}$ receptor which includes an epitope which is immunologically non-cross reactive with the RP-2 polypeptide disclosed in Owens et al. (loc. cit.).

13. An antibody which is specific for an epitope of $P_{2\times}$ receptor which is immunologically non-cross reactive with the RP-2 polypeptide disclosed in Owens et al. (loc. cit.).
14. An antibody as claimed in claim 13, which is a monoclonal antibody.

15. A cell expressing an antibody as claimed in claim 14.

16. The use of a P2X receptor or a preparation thereof, as claimed in claim 7, 8 or 9, as a screen for compounds useful in the treatment or prophylaxis of a human or non-human animal disease or condition.

17. The use of a P2X receptor or a preparation thereof as claimed in claim 9, 10 or 11 as a screen for identifying a P2X agonist or a P2X antagonist.

18. A P2X agonist or a P2X antagonist identified by a screen as described in claim 17.

19. A method for obtaining a DNA molecule according to claim 1, wherein the molecule is obtained by chemical synthesis or by using recombinant DNA technology.

20. A method for obtaining a P2X receptor comprising expressing the P2X receptor using a host cell according to claim 8 and, optionally, purifying the P2X receptor.

21. A DNA molecule, a P2X receptor, a P2X agonist or a P2X antagonist, a method, or a use, substantially as hereinbefore described, with reference to the accompanying examples.
FIGURE 1

P2xα1 cDNA from rat vas deferens

```
1  gccaaagcgttctgtacccaccccaggttttttctccccaccccagcccaccccaccccacccagtgaatccctctgtgct
81  agccctgctctcctattagggcggcagccccagctacccactgcatattgtgagctgagctggctgccccctgta
161  ccatagaggccgttggtctctttgctagccccctcttgccagcacc gct cgg cgg cttg caa gat
1
231  gag ctg tca gcc ttc ttc ttt gaa tat gac act ccc cgg atg tgg ctt gta cga aac aac
290  be l s a f f f e y d t p r h v l v r n k
27
291  aag gtt gga gtc att ttc cgt atc cag tgt gtt ctt gtc tgc ttc att ggg tgg
350  gk v g v i f r l q o l v v y v i g w
47
351  gtt ttt gtc tat gaa aag gaa tac cag acc tca aat gac ctc atc aac aat gtt tcc gtt
410  ve v v y e k g y q t s s d l i s s s v s v
67
411  aag ctc aag ggc tgg gct gtt gct acc ctt cag gcc tgt gca gtc cgg ctt gca gac gtt
470  k l g l a v t q l q g l g p q v w d v
87
471  gct gac tat gtc ttc cca gca cac ggg gac aag tcc ttt gta gtt att gcc aac ttc ttc
530  a y v f p a h g d s s f v v m t n f i
107
531  gtt gcc cct cag cag act cag gac cat tgt gca gac aac cca gag ggt gcc ata tgc cag
590  v t p q q t q g h c a e n p e g g i c q
127
591  gat gac agt gcc tgc act cca gaa gca gag aag aag cac gct atg acc atg acg aca ggc
650  d d s c g c t p g k a e r k a q g i r t g
147
651  aac tgt gtc ccc ttc aat ggc act tgg aag aca tgg aat gat ctt gtt tgt tgg ctt gta
710  n c v p f n g t v k t c i f g w c p y
167
711  gag gtt gat gac aag atc cca aag cct cgt ctt cgg gct gag gac aac tgt acc ctc
770  e v d d k i s p a l l r e a e n f t l
187
```

CCC CAG CTG GCA CAT GCC TGC TAC CCA TGC CCT CCA CAC A G
P Q L A H G C Y P C P F H R sequence of RP-2
rat P2X clone 3

1  cgcagcagacccgtccgagctgggtttgagctacgccggagctgggttggaggtct
2  gagcagcagggctccgagcc ATG GCG GGC TGC TGC CTG GTG CTC GGC TTC TTC GAG TAC 145
3  1 MAG C C S V L G S F L P E Y 15
4  146 GAC ACG CCG CCG ATC GTG CTC ATC CCG AGC AGG CTG AAA GCG GTG GTC ATG AAC CCG GCG GTG 205
5  16 D T P R I V L I R S R K V G L M N R A V 35
6  206 CAG GTG CTC ATC GTG GCT GCT ATC GGG TCG GTG TCG GTG TCG GAA AAG GGC TAC CAG 265
7  36 Q L L I L A Y V I G W V F V W E K G Y Q 55
8  266 GAA ACG GAC TTC GTG GTC AGC TCG GTG ACA ACC AAA GCC AAA GGT GTG GCT GTC ACC AAC 325
9  56 E T D S V U V S S V T K A K G V A V T N 75
10 326 ACC TCT CAG CTT GGA TTC CCG ATC TCG GGC GTG TCG GAG TAT GTG ATT CCA GCT CAG GAG 385
11 76 T S Q L G F R I W D V A D Y V I P A Q E 95
12 386 GAA AAC TTC CTC TTC ATT ATG ACC ACC ATG ATT GTC ACC GTG AAC CAG ACA CAG AGC ACC 445
13 96 E N S L F I M T N H I V T T N Q T S T 115
14 446 TGT CCA GAG ATT CCT GAT AAG ACC AGC ATT TGT AAT TCA GAC GCC GAC TGC ACT CCT GGC 505
15 11 C F E I D P K T S I C N S D A D C T P G 135
16 506 TTC GTG GAC ACC CAC AGC AGT GGA GTT GCG ACT GGA AGA TGT GTT CTT TTC AAT GAG CTC 565
17 135 S V D T H S S G V A T G R C V P F N E S 155
18 566 GTG AAG ACC TGT GAG GTG GCT GCA TGG TCG CCC GTG GAG AAC GAC GTT GGC GTG CCA ACG 625
19 156 V K T C E V A W C P V E N D V G V P T 175
20 626 CCG GCT TCC TTA AAG GTC GCA GAA AAC TTC ACC CTC TGG GTA AAG AAC ATC TGG TAC 685
21 176 P A F L K A A E N F T L L V K N N I W Y 195
22 686 CCC AAG TTT AAC TCC AGC AAG AGG AAC AAC ATC CCC AAC ATC ACC AGG TCC TAC CTC AAA 745
23 196 P K F H F S K R N I L P N I T T S Y L K 215
24 746 TCG TGC ATT TAC AAT GCT CAA AGC GAT CCC TCC CCC CTA TCC CTC GGT GTC ACA ATC 805
25 216 S C I Y N A Q T D P F C P I F R L G T I 235

3/15
rat P2X clone 6

1 cactgggtcacagtgcctggtcatggtctttctcagccctcattaaagcagcccactccagtctctgtat 80
81 ctttgctccagctctgctctagttttctcctttcttaggtctgcatccagcccttctcattaatggttgcttgagcagcagttctctca 160

161 gt ATG AAC TGT ATA TCA GAC TTC TTC ACC TAC GAG ACT ACC AAC TCG G TG GTG GTG AAG 219
1 MNCISDFFTTYETTKSVVK 19

220 AGC TGG ACC ATT GGC ATC ATC AAC CGA GCC GTC CAG CTG GTG ATT ATC TCC TAC TAC TGG GTG 279
20SWTGIGINRAVQLLIISYFV 39

280 GGG TGG GTG TTC TTG CAT GAG AAG GCC TAC CAA GTG AGG GAC ACC GCC ATT GAG TCC TCA 339
40GWVFLHEKAYQVDRTAIES 59

340 GTA GTT ACA AAG GTG AAA GGC TTC GGG CGC ATG ACC AAG GTA ATG GAC GTG TCG GAT 399
60VVTKTGVKGFGRYANRVMDS 79

400 TAT GTG ACC CCA CCC CAG GCC ACC TCT GTG TTC GTG ATC ATC ACC ACC AAA ATG ATC GTT ACT 459
80YVTTPQGTSVFIVITKMIVT 99

460 GAA AAT CAA ATG CAA GGA TTC TGT CCA GAG AAT GAA GAG AAG TAC CGG TGT GTG TCT GAC 519
100ENQHGFPCPENEEKYRCVSD 119

520 AGC CAG TGT GGG CCT GAA CGC TTC CCA GGT GGG GGG ATC CTG GAC GCC TGG AAG GAC 579
120SQCPFRPGGGILTGRCV 139

580 TAC AGC TCT GGT CCA CGG ACC TGT GAG ATC CAG GCC TGG TCC ACT GAG GTG GAC ACC 639
140YSVSLRTECATIEQGWCPTEVDT 159

640 GTG GAG ATG CCT ATG ATG GAG GCT GAG AAC TTC ACC ATT TCC ATC AAG AAG AAC ATC 699
160VEMPIMEAENFTIPIKNSI 179

700 CGT TTC CCT TTC TTC AAC TTT GAG AAG GGA AAC ATC CTG CCT ATT AAC ACC ACC GAC AAG GAC 759
180RFPLGNFEKGNLPLNPNTDKD 199

760 ATA AAG AGG TGC CGC TTC CAC CCA CAA AAG GCC CCA TTT TGC CCC ACC TTG AAG GTA CGG 819
200IKRCKRFHPKEKAPFCPILRVG 219
1  gcctccagctgccctttgcttcctgtctttgctctcactgccaccgcctgtctttcataaagggccagggcagccccaa  80
81  gaagctctatcccgacgtggttggtgcacccggctcaccactgagagcagaggggttgctcagttcagc  160
161  cagccggcccacc ATG GCA CGG CGG TTC CAG GAG GAG CTG GCC GCC TTC CTC TTC GAG TAT  221
1  MARRFQEEELAAFLFEY  16
17  DTPRHLVVRNKKVGVIVIFRLI  36
222  CAG CTG GTG GTG GTG TAC GTC ATC GGC TGG GTG TTT CTC TAT GAG AAG GGC TAC CAG  331
37  QLVQLVLYVIWGWLFLYEGKYQ  56
282  CAG CTG GTG GTG GTG TAC GTC ATC GGC TGG GTG TTT CTC TAT GAG AAG GGC TAC CAG  341
37  QLVQLVLYVIWGWLFLYEGKYQ  56
342  ACC TCG AGC GCC CTC ATC AGC ATG GTG TCT AGA TCT AAT GCA AAG GTC GCC GCC GTG ACC CAG  401
57  TSSGLISSSVSVKLKGALAVTG  76
402  CTC CCT GCC CTC GCC CCC CAG GTG TGC TGG GAT GTG GCT GAC TAC GTC TTCTCCA GCC CAG GGG  461
77  LPGLGQPQWVDVAHDYVFQAPQG  96
462  GAC AAC TCC TTC GTG GTG ATG ACC AAT TTC ATC GTG ACC CGG AAG CAG ACT CAA GCC TAC  521
97  DINSFVMTNFIVTPKQTQGY  116
522  TGC GCA GAG CAC CCA GAA GGG GCC ATG TGC AAG GAA GAT GGC TGT ACC CCT GGG AAG  581
117  CAEHPPEGGCICKEDSGCTPGK  136
582  GCC AAG AGG AAG GCC CAA GCC ATC CGC AGC GCC AAG TGT GTG GCC TTC AAC GAC ACT GTG  641
137  AKRKAQGIRTGKCVAFNDTV  156
642  AAG AGG TGT GAG ATC TTT GCC TGG TGC CCC GTG GAG GTG GAT GAC GAC ATC CGG CGC CCT  701
157  KTCEIFGWCPEVEVDDDIDPRP  176
1362 AGG ACA TCC TGA tgtcgggcccccaactcttgactgggtcacagtgaaggttcc 1437
397 R T S

1438 cagccaggcccagagggcctccccaggaagtctctctaccctctcagccagccggcagagctcggtatctg 1517
1518 gcatagtaggagagacactgtgcaaatctctgagctcggctccgacccccacacacccttgagggagggctacacttagctcgtag 1597
1598 cgcctctctggtgggggaaagctggggtggggtggcaggaccccacacaccctgcacccctgctctcggttttctctctcc 1677
1678 ggacgtctcattatcctcaaccggtcctccccatctctctctactgatctgctcctcgggtgtagtggcagtcagtagataggtgcc 1757
1758 taaaatgunaattacaaaaaaaaaataagatgaacacaaacacaaaatctaatctctctgtggactgtcactgaccacacactttcatcactgtcagta 1837
1838 gatacgtgtgggtagtggcttcgctcttgacactggacagtggggcctttttctgactgtcaaaagaaatctgctgcatatgagacgcccct 1917
1918 cgggggtcctctggtctggctcttccccctggggaaccttactgctctgggcccaccccaagccccac 1997
1998 tccccaaacacaagtccagtgggtctggcagttgactctgtgtgtgagtggaatccctgccccagcacactttgaggaacctaagtt 2077
2078 aagagtggtgcgtctcgatcagctctttgtactgctgccagggctagcaacgccactgtgacgctgctcgactgctctcggtctctcgttt 2157
2158 caggtgcacacactgtcagaggtctcaaaccacctccagcatgtgttgtgtgtgtctgctcactggtgcaacacdgagtgcactgtcatgca 2237
2238 cacaacccagctgcacacacaacctacgtgcacacagctcctaaccacactgtgtgtgtatactgtgaaacgctgcggcctgcaacaac 2317
2318 acaaacatctacaacacatatcaaccccaagcacaacacacatgtgtctcgtgcctcatgctcaacctccatagggaaggggctttc 2397
2398 tccaggttgctggcaggccagggaggctccccagccatgaatcttactctcagctacctcgggtggggtgggagccccagcc 2477
2478 cacaattctgtggcctcctgctgtggtcagccccagctccagaagcctcctggctgtctctgtaacagaggtctgagg 2557
2558 aagcggggggtggagtacaattaaagggatagggccacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2637
2638 aaaaaa 2643